



BNL-221919-2021-JAAM

# Phylogenetic occurrence of the phenylpropanoid pathway and lignin biosynthesis in plants

T. Yao, M. Xie

To be published in "Frontiers in Plant Science"

July 2021

Biology Department  
**Brookhaven National Laboratory**

**U.S. Department of Energy**

USDOE Office of Science (SC), Biological and Environmental Research (BER) (SC-23)

Notice: This manuscript has been authored by employees of Brookhaven Science Associates, LLC under Contract No. DE-SC0012704 with the U.S. Department of Energy. The publisher by accepting the manuscript for publication acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

## **DISCLAIMER**

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party's use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

# Phylogenetic occurrence of the phenylpropanoid pathway and lignin biosynthesis in plants

Tao Yao<sup>1</sup>, Kai Feng<sup>1</sup>, Meng Xie<sup>2</sup>, Jaime Barros<sup>3</sup>, Timothy Tschaplinski<sup>1</sup>, Gerald A. Tuskan<sup>1</sup>, Wellington Muchero<sup>1</sup>, Jin-Gui Chen<sup>1\*</sup>

<sup>1</sup>Oak Ridge National Laboratory (DOE), United States, <sup>2</sup>Brookhaven National Laboratory (DOE), United States, <sup>3</sup>University of North Texas, United States

*Submitted to Journal:*  
Frontiers in Plant Science

*Specialty Section:*  
Plant Metabolism and Chemodiversity

*Article type:*  
Review Article

*Manuscript ID:*  
704697

*Received on:*  
03 May 2021

*Revised on:*  
11 Jul 2021

*Journal website link:*  
[www.frontiersin.org](http://www.frontiersin.org)

### *Conflict of interest statement*

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

### *Author contribution statement*

T.Y. and K.F. drafted the manuscript. M.X., J.B., T.J.T., G.A.T., W.M. and J.-G.C. revised the manuscript.

### *Keywords*

Lignin biosynthesis, Tracheophytes, transcription factor, EPSP synthase, phylogenetic occurrence, lignin utilization.

### *Abstract*

Word count: 152

The phenylpropanoid pathway serves as a rich source of metabolites in plants and provides precursors for lignin biosynthesis. Lignin first appeared in tracheophytes and has been hypothesized to have played pivotal roles in land plant colonization. In this review, we summarize recent progress in defining the lignin biosynthetic pathway in lycophytes, monilophytes, gymnosperms, and angiosperms. In particular, we review the key structural genes involved in p-hydroxyphenyl-, guaiacyl- and syringyl-lignin biosynthesis across plant taxa and consider and integrate new insights on major transcription factors, such as NACs and MYBs. We also review insight regarding a new transcriptional regulator, 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase, canonically identified as a key enzyme in the shikimate pathway. We use several case studies, including EPSP synthase, to illustrate the evolution processes of gene duplication and neo-functionalization in lignin biosynthesis. This review provides new insights into the genetic engineering of the lignin biosynthetic pathway to overcome biomass recalcitrance in bioenergy crops.

### *Contribution to the field*

May 3, 2021 Editor Frontiers in Plant Science Section: Plant Metabolism and Chemodiversity Special Issue: Phenylpropanoid Systems Biology and Biotechnology Dear Editor, We would like to submit our manuscript entitled "Phylogenetic occurrence of the phenylpropanoid pathway and lignin biosynthesis in plants" By Yao et al. for your consideration for publication in the Special Issue "Phenylpropanoid Systems Biology and Biotechnology" of the Plant Metabolism and Chemodiversity section. We experienced some delays in preparing manuscript for submission to this special issue. If this special issue no longer accepts manuscript, we are OK to transfer our manuscript to other appropriate sections. The phenylpropanoid pathway serves as a rich source of metabolites in plants and provides precursors for lignin biosynthesis. In this review, we analyze the phylogenetic occurrence of the phenylpropanoid and lignin biosynthetic pathways and transcriptional regulations in lycophytes, monilophytes, gymnosperms, and angiosperms. We also illustrate the evolution processes of gene duplication and neo-functionalization in lignin biosynthesis. We believe that this study provides new insights into the genetic engineering of the lignin biosynthetic pathway. Thanks for your consideration. If you require further information, please don't hesitate to contact me. Sincerely, Jin-Gui Chen On behalf of all authors

### *Funding statement*

This work was supported by the Center for Bioenergy Innovation by the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

1 **Phylogenetic occurrence of the phenylpropanoid pathway and lignin**  
2 **biosynthesis in plants**

3  
4 **Tao Yao<sup>1,2</sup> †, Kai Feng<sup>1,2</sup> †, Meng Xie<sup>3</sup>, Jaime Barros<sup>2,4</sup>, Timothy J. Tschaplinski<sup>1,2</sup>, Gerald**  
5 **A. Tuskan<sup>1,2</sup>, Wellington Muchero<sup>1,2</sup>\* and Jin-Gui Chen<sup>1,2</sup>\***

6 <sup>1</sup> Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States

7 <sup>2</sup> Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, United  
8 States

9 <sup>3</sup> Biology Department, Brookhaven National Laboratory, Upton, NY, United States

10 <sup>4</sup> BioDiscovery Institute and Department of Biological Sciences, University of North Texas,  
11 Denton, TX, United States

12  
13 **\* Correspondence:**

14 Wellington Muchero (mucherow@ornl.gov)

15 Jin-Gui Chen (chenj@ornl.gov)

16  
17 †These authors contributed equally to this work  
18

19 Notice: This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-  
20 00OR22725 with the U.S. Department of Energy. The United States Government retains and the  
21 publisher, by accepting the article for publication, acknowledges that the United States  
22 Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or  
23 reproduce the published form of this manuscript, or allow others to do so, for United States  
24 Government purposes. The Department of Energy will provide public access to these results of  
25 federally sponsored research in accordance with the DOE Public Access Plan  
26 (<http://energy.gov/downloads/doe-public-access-plan>).

27 **Abstract**

28

29 The phenylpropanoid pathway serves as a rich source of metabolites in plants and provides  
30 precursors for lignin biosynthesis. Lignin first appeared in tracheophytes and has been  
31 hypothesized to have played pivotal roles in land plant colonization. In this review, we summarize  
32 recent progress **in defining** the lignin biosynthetic pathway in lycophytes, monilophytes,  
33 gymnosperms, and angiosperms. In particular, we review the key structural genes involved in *p*-  
34 hydroxyphenyl-, guaiacyl- and syringyl-lignin biosynthesis across **plant taxa and** consider and  
35 **integrate new insights** on major transcription factors, such as NACs and MYBs. **We also review**  
36 **insight regarding** a new transcriptional regulator, 5-enolpyruvylshikimate-3-phosphate (EPSP)  
37 synthase, canonically identified as a key enzyme in the shikimate pathway. We use several case  
38 studies, including EPSP synthase, to illustrate the evolution processes of gene duplication and neo-  
39 functionalization in lignin biosynthesis. This **review** provides new insights into the genetic  
40 engineering of the lignin biosynthetic pathway to overcome biomass recalcitrance in bioenergy  
41 crops.

42 **Keywords: lignin biosynthesis, tracheophytes, transcription factor, EPSP synthase,**  
43 **phylogenetic occurrence, lignin utilization.**

44

In review

45 It is hypothesized that the first land plants possessed adaptive metabolic, physiologic, and  
46 morphologic changes as a means of coping with abiotic stresses, such as UV-B irradiation and  
47 desiccation (Niklas et al., 2017). In this scenario the phenylpropanoid pathway played a pivotal  
48 role in land colonization of early plants by yielding protective secondary metabolites including  
49 flavonoids and lignin. Many flavonoids bestowed land plants with the ability to absorb UV-B,  
50 while lignin, as the cell wall component, provided mechanical support and facilitated water  
51 transport for the vascular plants (Rensing, 2018). Recently several comparative genomics,  
52 phylogenetics, and evolutionary genetics approaches have been employed to illustrate the  
53 evolution of phenylpropanoid biosynthesis pathway (Ma and Constabel, 2019; Davies et al., 2020).  
54 In this review, we unite these current outcomes and provide a comprehensive overview of the  
55 phylogenetic occurrence of phenylpropanoid biosynthesis and lignin biosynthesis pathways and  
56 showcase the role of gene duplication and neo-functionalization contributing to land plant  
57 evolution.

58 To aid our understanding the phylogenetic occurrence of the phenylpropanoid pathway and lignin  
59 biosynthesis in plants we offer a primer on lignin biosynthesis. Lignin is derived from three major  
60 hydroxycinnamyl alcohols, including *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol  
61 by radical coupling (Weng and Chapple, 2010). As such, *p*-hydroxyphenyl (H), guaiacyl (G) and  
62 syringyl (S) monolignols are the main units for lignin polymerization. In addition, two additional  
63 non-canonical monolignols, caffeoyl alcohol (C) and 5-hydroxyconiferyl (5HG) alcohol, have been  
64 found naturally in some species or can be introduced via genetic engineering (Dixon and Barros,  
65 2019; Wang et al., 2020).

66 The lignin biosynthesis pathway has been refined and re-envisioned by several research groups  
67 over the past two decades. Based on recent studies in the model herbaceous plant *Arabidopsis* and  
68 the model woody plant *Populus*, eleven core structural enzymes of the lignin biosynthesis pathway  
69 have been identified (Boerjan et al., 2003; Vanholme et al., 2013; Zhang et al., 2020). L-  
70 phenylalanine ammonia-lyase (PAL), 4-hydroxycinnamate CoA ligase (4CL) and cinnamate 4-  
71 hydroxylase (C4H) are the three enzymes that belong to the general phenylpropanoid pathway  
72 shared by the biosynthesis of lignin and flavonoids. Generally, the initial substrate of the  
73 phenylpropanoid pathway, phenylalanine, is converted into cinnamate by PAL, C4H converts  
74 cinnamate into *p*-coumarate, and *p*-coumarate is then activated by 4CL to form *p*-coumaroyl CoA.

75 The other eight enzymes belong to lignin-specific pathway (Figure 1), including cinnamoyl CoA  
76 reductase (CCR), cinnamyl alcohol dehydrogenase (CAD), coumarate 3-hydroxylase (C3H),  
77 coumaroyl shikimate 3'-hydroxylase (C3'H), ferulate/coniferaldehyde 5-hydroxylase (F5H),  
78 caffeate/5-hydroxy-coniferaldehyde 3/5-O-methyltransferase (COMT), caffeoyl CoA 3-O-  
79 methyltransferase (CCoAOMT), hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl  
80 transferase (HCT) and caffeoyl shikimate esterase (CSE). *p*-coumaroyl CoA is converted into the  
81 simplest H-lignin monomer by a reductase CCR and a dehydrogenase CAD. In addition to CAD  
82 and CCR, G-lignin biosynthesis starting from *p*-coumarate requires C3H, COMT and 4CL; or 4CL,  
83 HCT, C3'H, CSE, and CCoAOMT. F5H and COMT are crucial for S-lignin biosynthesis.  
84 Noticeably, aldehyde dehydrogenase (ALDH) catalyzes the opposite direction of reactions in  
85 lignin biosynthesis, which is required for ferulate and sinapate biosynthesis from coniferaldehyde  
86 and sinapaldehyde, respectively (Nair et al., 2004).

## 87 I. Lignin biosynthesis pathway in tracheophytes

## 88 1. The origin of lignin biosynthesis pathway

89 Although lignin has not been discovered in bryophytes, nine structural gene families that are  
90 responsible for the biosynthesis of H- and G-lignin monomers occur in moss genomes (Xu et al.,  
91 2009) (Table 1). Studies using the model plant *Physcomitrella patens* shed light on the biosynthesis  
92 pathway of phenylpropanoids and lignin. Knock-out of the *CYP98* gene in *P. patens*, which  
93 encodes a P450 oxygenase, blocks the biosynthesis of the moss cuticle, thus affecting gametophore  
94 formation and organ fusion. C3'H is a homolog of CYP98 in higher plants. However, CYP98 in  
95 moss uses the *p*-coumaroyl-threonate as substrate, whereas C3'H in higher plants uses *p*-  
96 coumaroyl-shikimate as substrate (Schoch et al., 2001), leading to the distinct biosynthesis  
97 pathways for cuticle (Renault et al., 2017). Interestingly, no phenylpropanoid genes have been  
98 found in red algae genomes, but trace amounts of lignin have been reported in red algae, and as  
99 such, indicates that the lignin biochemical machinery preexisted the evolution of land plants  
100 (Martone et al., 2009; Brawley et al., 2017). The extant presence of lignin in red algae may also  
101 represent convergent evolution independent of lignin biosynthesis in bryophytes.

## 102 2. H-lignin biosynthesis in seedless vascular plants

103 During land plant evolution, lignin appeared first in lycophytes (Renault et al., 2019) in the form  
104 of H-lignin. Interestingly, there are only low levels of H-lignin in gymnosperms and traces of H-  
105 lignin in angiosperms. In contrast, H-lignin is highly abundant in seedless vascular plants,  
106 including lycophytes and pteridophytes. Lignin is found between the cellulose matrix and forms a  
107 rigid cell wall in these plants (Espineira et al., 2011; Ralph et al., 2019). In gymnosperms and  
108 angiosperms, H-lignin can be enriched by down-regulation of *C3'H*, *HCT*, and *CSE* genes (Franke  
109 et al., 2002a; Wagner et al., 2007; Coleman et al., 2008; Li et al., 2010, Vanholme et al, 2013,  
110 Fornale et al., 2015), though in many cases growth was negatively impacted. Interestingly, the  
111 *Arabidopsis C3'H* mutant *ref8* showed severe growth defect that was rescued by disruption of the  
112 mediator complex units MED5a and MED5b (Bonawitz et al., 2014). These results indicate that  
113 H-lignin may represent one of the earliest forms of lignin.

## 114 3. G- and S-lignin biosynthesis in pteridophytes

115 G-lignin biosynthesis in pteridophytes is evolutionarily conserved. The *Df4CL2* gene is a 4-  
116 coumarate:coenzyme A ligase coding gene identified from the fern species *Dryopteris fragrans*.  
117 Heterologous expression of this gene in tobacco increased the synthesis of lignin, demonstrating  
118 the conserved function of *4CL* in *D. fragrans* and tobacco (*Nicotiana tabacum*) (Li et al., 2020).  
119 Similarly, two caffeoyl coenzyme A 3-O-methyltransferases (CCoAOMT) have been cloned from  
120 the fern species *Polypodiodes amoena*, and their functions in lignin biosynthesis have been  
121 confirmed via heterologous expression in *Arabidopsis* (Zhang et al., 2019).

122 S-lignin has been identified in lycophytes such as *Selaginella mellendorffii*; however, its  
123 biosynthetic pathway is different from that in angiosperms (Renault et al., 2019). In angiosperm,  
124 both *C3H* and *F5H* are involved in S-lignin biosynthesis. In contrast, in lycophytes, *SmF5H* has  
125 dual functions that enables S-lignin to be synthesized directly from *p*-coumaraldehyde and *p*-  
126 coumaryl alcohol. Here, *SmF5H* and *SmCOMT* form a scaffold and are responsible for S-lignin  
127 biosynthesis. Phylogenetic analyses suggests that these two genes were independently evolved  
128 from their counterparts in angiosperm (Weng et al., 2008b; Weng et al., 2011). Besides the well-  
129 known S-lignin biosynthesis in *Selaginella*, several ferns, such as *Dennstaedtia bipinata*, also

130 contain a large amount of S-lignin in the sclerotic sheaths. However, the biosynthetic pathway has  
131 not been elucidated (Logan and Thomas, 1985; Weng and Chapple, 2010). Further studies of the  
132 lignin biosynthesis-related genes in these fern species and comparison with what we have known  
133 in other species in the lineage is needed to provide insights on the evolution of the S-lignin  
134 biosynthetic pathway.

#### 135 4. G-lignin biosynthesis in gymnosperms

136 Gymnosperms diverged from angiosperms 300 Mya (De La Torre et al., 2020). In general,  
137 gymnosperms lack the *F5H* gene, and therefore, gymnosperm lignin mainly contains G-  
138 monolignol and contains no or little S-monolignol (Li et al., 2001; Weng and Chapple, 2010).  
139 When *Cf4CL* and *CfCCoAOMT* were cloned from *Cryptomeria fortunei*, a gymnosperm, and  
140 heterologously expressed in tobacco, an angiosperm, G-lignin biosynthesis was increased,  
141 indicating that these two lignin genes can function equally well in both gymnosperms and  
142 angiosperms (Guo et al., 2019). Similarly, a caffeoyl shikimate esterase (CSE), LkCSE, from *Larix*  
143 *kaempferi*, can convert caffeoyl shikimate to caffeate and shikimate, supporting the conserved  
144 function of CSE between gymnosperms and angiosperms (Wang et al., 2019). Gymnosperms also  
145 produce a compression layer within xylem that enriched in H-lignin in tracheid. A recent study  
146 showed that spatial patterning of H- and G-lignin during wood formation is related to different  
147 localizations and enzyme activities of lignin polymerization enzymes, laccases (Hiraide et al.,  
148 2021). Interestingly, some gymnosperm species, such as *Gentales*, can also synthesize S-lignin  
149 (Renault et al., 2019). *Gnetum genmon* contains angiosperm-like vessels as well as tracheids and  
150 fiber tracheids (Tomlinson, 2001), and it shares the chemical compositions of lignin with  
151 angiosperms (Nawawi et al., 2016). These results suggest that the biosynthetic pathway for G-  
152 lignin is shared between gymnosperms and angiosperms. Based on these results we are left with  
153 two alternate hypotheses; ancient gymnosperms were able to produce S-lignin which was  
154 subsequently lost in modern gymnosperms or the occurrence of S-lignin in *Gentales* is a recent  
155 convergent evolutionary event. As an ancient gymnosperm, further systematic studies of lignin  
156 biosynthesis in *Gentales* are needed to definitively describe the evolution trajectory in  
157 gymnosperms.

#### 158 5. G- and S-lignin biosynthesis in angiosperms

159 Angiosperms contain the lignin composed of G-, S- and H-lignin monomers in various ratios  
160 (Mansfield et al., 2012). The lignin biosynthesis pathways of angiosperms have been characterized  
161 using the model plants, *Arabidopsis*, *Populus*, and *Brachypodium*, among others. Xu et al. (2009)  
162 analyzed ten of eleven lignin biosynthetic gene families (without CSE) across fourteen plant  
163 species and one symbiotic fungal species using comparative genomics. The analysis revealed that  
164 the rapid expansion of these gene families occurred after the divergence between dicots and  
165 monocots 140-150 million years ago (Xu et al., 2009; Rao and Dixon, 2018).

166 Although the lignin biosynthetic pathways are generally conserved among angiosperms,  
167 alternative pathways have evolved in monocots. In dicots, the first enzyme in the phenylpropanoid  
168 pathway, PAL, converts phenylalanine (Phe) to cinnamate. Cinnamate is then converted to *p*-  
169 coumarate by the second enzyme, C4H. However, a bypass route has been discovered in monocots.  
170 PTAL was identified as a bifunctional enzyme that recognizes tyrosine (Tyr) as the substrate and  
171 converts it to *p*-coumarate directly in *Brachypodium distachyon* (Barros et al., 2016). <sup>13</sup>C isotope  
172 feeding with *BdPTAL1-RNAi* transgenic plants revealed that BdPTAL1-mediated lignin

173 biosynthesis contributed to half of the total lignin content in *B. distachyon* (Barros et al., 2016).  
174 Another grass-specific enzyme is *p*-coumaroyl-CoA:monolignol transferase (PMT) that catalyzes  
175 the incorporation of *p*-coumarate into the lignin polymer backbone typically found in the *Poaceae*  
176 family (Withers et al., 2012; Petrik et al., 2014). These findings suggest that lineage-specific lignin  
177 biosynthetic pathways have evolved independently in dicots and monocots and highlight the need  
178 to study species-specific branches in the lignin biosynthetic pathway.

179 The *C4H* gene progenitor appears to have duplicated in early seed plants, yielding two clades that  
180 are preserved in Taxaceae and most angiosperms. A second duplication event happened after the  
181 divergence of dicots and monocots. By analyzing the protein structure and function of  
182 *Brachypodium C4H*, it was found that each of *Brachypodium C4H* paralog genes can rescue the  
183 growth defect of the *Arabidopsis c4h* mutant, indicating that the *C4H* in monocots preserved the  
184 canonical function in lignin biosynthesis. However, the protein structures of *C4Hs* in *B. distachyon*  
185 differ from that in *Arabidopsis*. This newly derived *C4H* type in monocots has an elongated N-  
186 terminus, which alters the subcellular localization and allows the orientation of *C4H* to the lumen  
187 of endoplasmic reticulum (ER) through a double-spanning hairpin structure. Therefore, it is  
188 possible that an alternate *C4H* exists within the ER (Renault et al., 2017).

189 *C3H* and *C3'H* catalyze the conversion of *p*-coumarate and *p*-coumaroyl shikimate into caffeate  
190 (via a bifunctional cytosolic ascorbate peroxidase, Barros et al., 2019) and caffeoyl shikimate (via  
191 a cytochrome P450 monooxygenase, Schoch et al. 2001), respectively. These enzymes play  
192 important roles in G-lignin and S-lignin biosynthesis. There is only one member of the *C3'H*  
193 family in *A. thaliana* and two members of the cytosolic *C3H* family in *A. thaliana* and *B.*  
194 *distachyon* (Franke et al., 2002b; Barros et al., 2019). *PtrC3'H3* was recognized as the homolog  
195 of *Arabidopsis C3H*. However, it was proposed that *PtrC3'H3* requires *PtrC4H1* or *PtrC4H2* to  
196 form a complex to enhance its enzymatic activity in *Populus trichocarpa* (Chen et al., 2011)  
197 (Figure 1). Recent study showed that triple knocking-down *PtrC4H1/PtrC4H2/PtrC3'H3* causes  
198 monolignol benzoate (ML-BL) conjugation and significantly reduces lignin biosynthesis while  
199 increasing H-lignin for about 70-fold (Kim et al., 2020). These findings suggest that simultaneous  
200 modification of *C4H* and *C3H* could be used for reducing biomass recalcitrance in bioenergy crops.

201 Phylogenetic analysis of 192 *4CLs* across land plants suggested that a duplication of the *4CL* gene  
202 family occurred prior to the split of gymnosperms and angiosperms (Li et al., 2015). Functional  
203 divergence of the *4CL* gene family, post duplication, has been broadly found in angiosperms. In  
204 fact, four members of the *4CL* gene family have been reported in *P. patens*, but only three of them  
205 were expressed under tested conditions (Silber et al., 2008). There are four *4CL* genes in  
206 *Arabidopsis*, five in rice, and seven in *Populus*. Functional analysis of these gene families revealed  
207 that only one subgroup of this gene families is involved in lignin biosynthesis, while other  
208 subgroups are involved in the biosynthesis of flavonoids or phenolics via neofunctionalization  
209 (Ehlting et al., 1999; Gui et al., 2011; Li et al., 2015; Rao et al., 2015) (Table 1). Loss-of-function  
210 mutation of *4CL* genes in herbaceous species causes reductions in G-lignin and increase of S/G  
211 ratios. However, knock-out *4CL1* gene in *Populus* led to reduction of S-lignin and decrease of S/G  
212 ratio, and the homeostasis of G-lignin was maintained by *4CL5* in *4cl1* mutant. These findings  
213 point towards a functional divergence of *4CLs* between herbaceous and woody species (Xiong et  
214 al., 2019; Tsai et al., 2020).

215 HCT catalyzes the conversion of caffeoyl shikimate to caffeoyl-CoA. Down-regulation of *AtHCT*  
216 caused the reduction of S-lignin content in *Arabidopsis* (Hoffmann et al., 2004). The orthologs of  
217 *HCTs* are present among all the land plants, which suggests that this enzyme evolved before the

218 occurrence of lignin. A recent study showed that *P. patens HCT* and *M. polymorpha HCT* can  
219 complement the deficiency of *Arabidopsis hct* mutant in terms of morphology and metabolite  
220 levels, suggesting that the function of *HCT* is likely conserved in all embryophytes (Kriegshauser  
221 et al., 2021). It appears that gene duplication of *HCT* occurred in dicots that produced the *HQT*  
222 gene. Despite the sequence similarity between *HCT* and *HQT*, the latter is required for  
223 biosynthesizing chlorogenic acid rather than lignin in *Cynara cardunculus* (Sonnante et al., 2010).  
224 Knock-down of *HCT* led to increase of G-lignin and decrease of S-lignin and S/G ratio in *Populus*  
225 (Zhou et al., 2020). However, knock-down of both *HCT1* and *HCT2* did not drastically change  
226 lignin content or composition in *B. distachyon*. Meanwhile, the saccharification efficiency was  
227 greatly enhanced in the double knock-down line (Serrani-Yarce et al., 2021). These findings  
228 suggest *HCT* genes play different roles in some monocots compared to that of dicots.

229 CSE is a newly discovered enzyme involved in monolignol biosynthesis. Together with 4CL, these  
230 two enzymes form a bypass pathway of monolignol biosynthesis in *Arabidopsis* (Vanholme et al.,  
231 2013). *CSE* genes cloned from *Medicago truncatula* and *Populus deltoides* have been shown to be  
232 functionally conserved with their *Arabidopsis* homolog (Ha et al., 2016; Saleme et al., 2017).  
233 However, the homolog of *CSE* gene has not been identified in most monocots, including maize  
234 and *Brachypodium*. Recently, the generation of *cse1*, *cse2* single mutant and *cse1/cse2* double  
235 mutant in *Populus* further confirmed their partial redundant roles in lignin biosynthesis. In addition  
236 to causing a 35% reduction in lignin content, the *cse1/cse2* double mutant significantly improved  
237 cellulose-to-glucose transformation efficiency. As such, CSEs in *Populus* could be promising  
238 target genes in biorefinery although their growth penalty should be managed to avoid (de Vries et  
239 al., 2021). Noticeably, *CSE* has also been shown to be functional in gymnosperms, such as *Larix*  
240 *kaempferi* (Wang et al., 2019). These findings suggest that *CSE* may be evolved prior to the  
241 divergence of gymnosperms and angiosperms, but was lost in many monocots (Wang et al., 2019;  
242 Serrani-Yarce et al., 2021).

243 COMT and F5H are two key enzymes required for catalyzing the intermediates in G-lignin  
244 biosynthesis into S-lignin biosynthesis. It has been reported that simultaneously manipulating  
245 *COMT* and *F5H* resulted in a dramatic change of S-lignin biosynthesis (Wu et al., 2019). *COMT*  
246 and *F5H* in *S. moellendorffii* appears to have an independent origin compared to that of  
247 angiosperms. There are two *F5H* genes in *Arabidopsis* (*AtF5H1/CYP84A1* and *AtF5H2/CYP84A2*),  
248 and only *AtF5H1* has been confirmed to be involved in lignin biosynthesis (Meyer et al., 1998).  
249 Similarly, there is one functional *COMT* gene identified among 13 homologous genes in  
250 *Arabidopsis* (Raes et al., 2003). In *Populus*, five *F5H* genes have been cloned, and two of them,  
251 *PtrF5H1* and *PtrF5H2*, were reported to be involved in lignin biosynthesis. Thirteen members of  
252 *COMT* gene family were identified in *P. trichocarpa*, but only *PtrCOMT2* is highly expressed in  
253 xylem (Shi et al., 2009) (Table 1). The function of *F5H* was shown to be conserved in monocots,  
254 such as *Oryza sativa*. One of three *F5H* genes, *OsCald5H1*, was reported to greatly affect the S/G-  
255 lignin composition via over-expression or knock-out (Takeda et al., 2017; 2019). *OsCaldOMT1*  
256 has been proven to be a functional *COMT* in rice (Lam et al., 2019). Noticeably, it not only  
257 regulates S-lignin biosynthesis, but also control triclin-lignin biosynthesis. The dual functions of  
258 *OsCaldOMT1* seems to be specific in grass species (Lam et al., 2019). *CCoAOMT*, another O-  
259 methyltransferase, converts feruloyl CoA to sinapoyl CoA and is required for the conversion of  
260 G-lignin into S-lignin. Genetic engineering of this enzyme led to change in G-lignin biosynthesis  
261 in *Populus*, alfalfa, *Pinus radiata*, maize and tobacco (Zhong et al., 2000; Guo et al., 2001; Wagner  
262 et al., 2011; Li et al., 2013; Xiao et al., 2020). These studies suggest the function of *CCoAOMT* is  
263 likely to be conserved among all angiosperms and occurred with the advent of the angiosperms.

264 CCR recognizes four types of cinnamoyl-CoAs, including *p*-coumaroyl CoA, caffeoyl CoA,  
265 feruloyl CoA and sinapoyl CoA, and converts them into cinnamaldehydes. Phylogenetic analysis  
266 of 146 CCR genes of various land plants revealed that CCR family contains three classes: CCR,  
267 CCR-like, and DFR, and that only the CCR class contains *bona fide* lignin biosynthetic genes. All  
268 these three classes are distributed across land plants, including *P. patens*, which contains a **single**  
269 functional CCR gene. These results suggested that the progenitor CCR gene evolved after the  
270 advent of lycophytes (Barakat et al., 2011). Still, functional divergence within the CCR family has  
271 **arisen** in several species. For example, in *Arabidopsis*, *AtCCR1* is involved in lignin biosynthesis,  
272 whereas *AtCCR2* is involved in pathogen response (Lauvergeat et al., 2001; Ruel et al., 2009).  
273 Downregulation of a CCR gene, *CCR2*, reduces lignin biosynthesis and increases saccharification  
274 efficiency in *Populus*. However, it also causes severe biomass penalty (Van Arker et al., 2014).  
275 Recently, a *ccr2* mutant was generated by the CRISPR/Cas9 approach that contain a null and  
276 haplo-insufficient allele in *Populus*. This mutant line does not have growth penalty, but still has  
277 low lignin content and improved saccharification efficiency (De Meester et al., 2020). Therefore,  
278 *CCR2* gene could be a useful target that can be deployed in genetic engineering of bioenergy  
279 woody crops.

280 CAD catalyzes the final step of monolignol biosynthesis **leading to** **compositional differences in**  
281 **lignin forms**. Guo et al. (2010) performed phylogenetic analysis of the CAD gene family from 52  
282 species **and** classified them into three classes. Class I comprises *bona fide* CADs which are only  
283 present in vascular plants, suggesting their co-occurrence with the advent of lignin. The functional  
284 characterizations of Class II and Class III CADs remain unclear (Guo et al., 2010). **Within the large**  
285 **gene families, CADC and CADD, PtrCAD1 and OsCAD2 have been reported to be functional CAD**  
286 **genes involved in lignin biosynthesis in Arabidopsis, rice, and Populus.** Knock-down or knock-  
287 out of these genes resulted in reduced lignin content as well as altered lignin structures (Anderson  
288 et al., 2015; Van Acker et al., 2017; Matin et al., 2019). Finally, it was reported that CAD and  
289 CCR form an enzyme complex that regulates monolignol biosynthesis in *P. trichocarpa* (Yan et  
290 al., 2019).

291 **In summary, as an important branch of the phenylpropanoid pathway, structural genes of the lignin**  
292 **biosynthesis pathway are conserved in most embryophytes. F5H and COMT contribute to S-lignin**  
293 **biosynthesis and have been hypothesized to have independent origins in S. moellendorffii and**  
294 **angiosperms. Gene duplications and gene family expansion of lignin biosynthetic genes in**  
295 **angiosperms have given rise to sub-functionalization and neo-functionalization of the various**  
296 **members, which is consistent with their morphological and functional changes compared with**  
297 **lower plants.**

## 298 **II. Transcriptional regulations of lignin biosynthetic pathway**

299 The lignin biosynthetic pathway includes both structural genes and regulatory proteins.  
300 Transcriptional regulation, controlling the gene expression of structural genes, plays important  
301 **roles in lignin biosynthesis. Such genes reflect the phylogenetic occurrence of the phenylpropanoid**  
302 **pathway and evolutionary trajectory of lignin biosynthesis in plants.** MYBs and NACs are two  
303 major transcription factor families, comprising three layers of the hierarchical transcriptional  
304 regulatory network (Ohtani and Demura, 2019). Therefore, we focus on analyzing these two  
305 families of transcription factors to illustrate the evolutionary divergence of transcriptional  
306 regulation in lignin biosynthesis.

## 307 1. MYB46-mediated transcriptional regulation of lignin biosynthesis

308 Transcription factor MYB46, is a central regulator in secondary cell wall formation (Zhong et al.,  
309 2007). MYB46 and MYB83 are two functionally redundant *Arabidopsis thaliana* MYB  
310 transcription factors that act as master switches of lignin biosynthesis regulating nine out of 11  
311 monolignol biosynthetic genes (*PAL*, *C4H*, *4CL*, *HCT*, *C3'H*, *CCoAOMT*, *F5H*, *CCR* and *CAD*)  
312 (Kim et al., 2014). Besides lignin, the biosynthesis of other secondary cell wall components,  
313 including xylan and cellulose are also regulated by MYB46/MYB83 (McCarthy et al., 2009;  
314 Zhong and Ye, 2012; Kim et al., 2013). Several MYB46 orthologs from other plant species have  
315 also been shown to function as key regulators for secondary cell wall biosynthesis, including  
316 PtMYB4 from pine, EgMYB2 from *Eucalyptus*, OsMYB46 from rice, PtrMYB2, PtrMYB3,  
317 PtrMYB20, PtrMYB21 from *Populus* and ZmMYB46 from maize (Patzlaff et al., 2003;  
318 Goicoechea et al., 2005; Zhong et al., 2011; Zhong et al., 2013). The functions of MYB46 and  
319 MYB83 in lignin biosynthesis are well-conserved in angiosperms.

320 The phylogenetic history of lignin related *MYBs* appears to coincide with the advent of the lignin  
321 biosynthetic genes, which first emerged in early land plants (Xu et al., 2014; Bowman et al., 2017).  
322 Homologs of *MYB46* and *MYB83* have been found in *P. patens* and *S. moellendorffii* (Zhong et al.,  
323 2010). Functional conservation of their homologs via transgenic validation has also been  
324 demonstrated in vascular plants, including gymnosperms and angiosperms (Zhao and Bartley,  
325 2014). We hypothesis that *MYB46* and *MYB83* might be required for phenylpropanoid biosynthesis  
326 outside of the lignin biosynthetic pathway in non-vascular plants while playing core roles in lignin  
327 biosynthesis in all vascular plants.

## 328 2. Upstream regulators of MYB46/MYB83

329 Major transcription factors regulating *MYB46/MYB83* are the NAC TF family proteins. NAC TF  
330 family proteins share a conserved NAC domain located at the N-terminal region and a highly  
331 divergent C-terminal activation domain (Olsen et al., 2005). These TFs are specific to plants and  
332 play diverse roles in plant defense, growth, and development (Olsen et al., 2005). *NAC*  
333 *SECONDARY WALL THICKENING PROMOTING FACTOR1 (NST1)* and *NST2* are redundantly  
334 responsible for secondary wall thickening in anther endothecium (Mitsuda et al., 2005). A MYB  
335 family protein, MYB26, localized in the nucleus, was found to be an upstream positive regulator  
336 of *NST1* and *NST2*. Overexpression of *MYB26* was found to increase lignin deposition and the  
337 expression of *NST1* and *NST2* (Yang et al., 2007). Recent study shows that Xylem NAC Domain  
338 1 (*XND1*) interacts with *NST1* and inhibits the transcriptional activity of *NST1*, thus repressing  
339 secondary cell formation (Zhang et al., 2020). In addition, VASCULAR-RELATED NAC-  
340 DOMAIN 6 (*VND6*) and *VND7* directly regulate *MYB46* and *MYB83* expression (Zhong et al.,  
341 2008; McCarthy et al., 2009; Ohashi-Ito et al., 2010; Yamaguchi et al., 2011). Overexpression of  
342 *VND6* and *VND7* can induce the ectopic differentiation of metaxylem-like vessels and protoxylem-  
343 like vessels, respectively (Kubo et al., 2005). The functional suppression of *VND6* and *VND7*  
344 caused defects in the formation of vessel elements (Kubo et al., 2005; Yamaguchi et al., 2008). In  
345 *Arabidopsis*, there are seven *VND* genes (*VND1-VND7*). Similar to *VND6* and *VND7*,  
346 overexpression of *VND1* to *VND5* also induces ectopic secondary cell wall deposition, suggesting  
347 that all *VND* members contribute to lignin biosynthesis during xylem vessel development (Zhou  
348 et al., 2014; Endo et al., 2014).

349 A third class of TFs involved in lignin biosynthesis include the WRKY gene family. Mutation of  
350 the *Arabidopsis WRKY12* gene caused secondary cell wall thickening in pith cells that is associated

351 with ectopic deposition of lignin, xylan, and cellulose. *WRKY12* mutation upregulated the  
352 transcription of downstream genes encoding the NAC domain TF *NST2* and the zinc finger TF  
353 *C3H14*, which activate secondary wall synthesis (Wang et al., 2010). Direct binding of *WRKY12*  
354 to the *NST2* gene promoter led to repression of *NST2* and *C3H14*, as defined by *in vitro* assays and  
355 *in planta* transgenic experiments (Wang et al., 2010). Interestingly, *WRKY12* gene is expressed in  
356 both pith and cortex that do not have secondary wall thickening, suggesting that *WRKY12* may  
357 control the parenchymatous nature of pith cells by acting as a negative regulator of secondary cell  
358 wall NACs (Wang et al., 2010). *WRKY15* was reported to repress the expression of *VND7* and  
359 suppress TE differentiation through indirect regulation (Ge et al., 2020). Based on our current  
360 understanding, *WRKY* TFs act upstream of NACs to regulate secondary cell wall biosynthesis.

361 Two members of the ASYMMETRIC LEAVES2-LIKE/LATERAL ORGAN BOUNDARIES  
362 DOMAIN (ASL/LBD) *ASL19/LBD30*, *ASL20/LBD18* were identified to be involved in a positive  
363 feedback loop for *VND7* expression that regulates tracheary elements (TE) differentiation-related  
364 genes (Soyano et al., 2008). Overexpression of *ASL19* and *ASL20* induced trans-differentiation of  
365 cells from nonvascular tissues into TE-like cells, similar to those induced by *VND6* or *VND7*  
366 overexpression. Expression of both *ASL19/LBD30* and *ASL20/LBD18* are dependent on *VND6*  
367 and *VND7* (Soyano et al., 2008). *XND1* has been reported to inactivate *VND6* by physically  
368 interacting with *VND6* and directing *VND6* from the nucleus to the cytoplasm (Zhong et al., 2020).  
369 Another NAC transcriptional factor, *VND-INTERACTING2* (*VNI2*), can bind to *VND* proteins  
370 and has been shown to function as a transcriptional repressor of *VND7*-mediated gene transcription  
371 (Yamaguchi et al., 2010). Recent studies show that *E2Fc* is a key upstream regulator of *VND6* and  
372 *VND7*, directly targeting the genomic loci of *VND6* and *VND7*. *E2Fc* is a transcriptional repressor,  
373 and transcript abundance of *VND6* and *VND7* were significantly increased in *E2Fc* knockdown  
374 *Arabidopsis* lines (Taylor-Teeple et al., 2015). Taken together, *VND6* and *VND7* represent key  
375 regulators in lignin biosynthesis whose functions are tightly regulated by various TFs (Ko et al.,  
376 2012; Schuetz et al., 2013). Phylogenetic analysis discovered close homologs of *VND6* and *VND7*  
377 in all vascular plants, whose functions were demonstrated to be conserved in *P. trichocarpa*, *Zea*  
378 *mays*, *Oryza sativa*, and *B. distachyon* (Zhong et al., 2010; Zhong et al., 2011; Valdivia et al.,  
379 2013).

380 *SND1/NST3* and *NST1* are required for secondary wall thickening in stem fibers (Mitsuda et al.,  
381 2007). When these genes were expressed constitutively in *Arabidopsis*, ectopic secondary wall  
382 thickening in various tissues was induced (Mitsuda et al., 2005; Mitsuda et al., 2007). Putative  
383 orthologs of *NST1*, *NST2*, and *SND1/NST3* are present in the genome of *Populus* and are expressed  
384 in developing xylem (Mitsuda et al., 2007), implicating a role in lignin biosynthesis. The function  
385 of *NST* homologs in lignin biosynthesis has been confirmed in *Medicago truncatula* and cotton  
386 (Zhao et al., 2010; Fang et al., 2020). However, the homologs of *NST* proteins have not been  
387 identified in gymnosperms or earlier species, implying that these proteins may not have evolved  
388 until the appearance of angiosperms (Nakano et al., 2015).

### 389 3. Downstream targets of MYB46/MYB83

390 Three MYB family proteins, *MYB58*, *MYB63*, and *MYB85*, whose coding genes are direct targets  
391 of *MYB46*, have been shown to function as direct transcriptional activators of lignin biosynthesis  
392 during secondary wall formation in *Arabidopsis* (Zhong et al., 2008; Ko et al., 2009; Demura and  
393 Ye, 2010; Zhou et al., 2020). All three MYBs cause ectopic lignin deposition when overexpressed.

394 The coding genes of three other MYB family proteins, MYB32, MYB4, and MYB7, are also  
395 directly activated by MYB46 (Ko et al., 2009). These three MYBs, sharing high sequence  
396 similarity with a conserved EAR motif, have been shown to be transcriptional repressors (Dubos  
397 et al., 2010). Trans-activation assays showed that these MYB transcription factors directly repress  
398 the expression of *SND1*, forming a feedback regulatory loop to maintain the abundance of *SND1*  
399 (Wang et al., 2011).

400 **KNOTTED ARABIDOPSIS THALIANA7 (KNAT7) and BEL1-LIKE HOMEODOMAIN6**  
401 **(BLH6) belong to knotted-like homeobox proteins and bell-like homeodomain proteins,**  
402 **respectively. KNAT7 and BLH6 were reported to be direct targets of MYB46 and MYB83 (Zhong**  
403 **and Ye, 2012). KNAT7 and BLH6 interact with each other and negatively regulate lignin**  
404 **biosynthesis while KNAT3 was reported to form heterodimer with KNAT7 to synergistically**  
405 **regulate lignin content and composition (Liu et al., 2014; Qin et al., 2020; Wang et al., 2020).**

406 Although the first and second layers of master switches of lignin biosynthesis were shown to be  
407 conserved in vascular plants even in early land plants, the targets of MYB46/83 have not been  
408 **shown** to be functionally conserved in lower plants. For example, the close homologs of *MYB58*  
409 and *MYB63* failed to be identified in lower tracheophytes (Zhong et al., 2010). In addition, their  
410 homologs in switchgrass were found to be mainly involved in flavonoids biosynthesis rather than  
411 lignin biosynthesis. **A plausible explanation is that wide expansion, promiscuous functionality,**  
412 **and functional diversification of the MYB family across different species have made it difficult to**  
413 **identify the genuine orthologs responsible for lignin biosynthesis regulation (Zhao and Bartley,**  
414 **2014; Nakano et al., 2015). Furthermore, lineage-specific MYBs may contribute to lignin**  
415 **biosynthesis in different tracheophytes. For instance, MYB75 was found to repress secondary cell**  
416 **wall biosynthesis and activate anthocyanin biosynthesis in dicots but not in monocots (Zhao and**  
417 **Bartley, 2014).**

#### 418 **4. EPSP as a transcriptional repressor**

419 EPSP synthase is a key enzyme in shikimate pathway, which is present in both plants and many  
420 prokaryotes. EPSP synthase has been a well-known herbicide target, which has been widely used  
421 in agriculture (Sammons and Gaines, 2014). Noticeably, there is only one copy of an EPSP  
422 synthase coding gene in green algae, lycophytes, and bryophytes, but duplicated genes were found  
423 in angiosperms, such as *Arabidopsis* and *Populus* (Tohge et al., 2013; Yang et al., 2017; Xie et al.,  
424 2018) (Figure 3). **The gene duplication in angiosperms may have given rise to neo-**  
425 **functionalization for the additional gene copy.**

426 A recent study in *P. trichocarpa* discovered the transcriptional regulatory function of one EPSP  
427 synthase gene (*PtrEPSP-TF*) (Xie et al., 2018). Overexpression of *PtrEPSP-TF* led to ectopic  
428 deposition of lignin, accumulation of phenylpropanoid metabolites and differential expression of  
429 secondary cell wall biosynthetic genes. It was shown that *PtrEPSP-TF* accumulates in the nucleus  
430 and acts as a transcriptional repressor by directly binding to the promoter element of a *hAT*  
431 transposase family gene (*PtrhAT*). *PtrhAT* is also located in the nucleus and serves as a  
432 transcriptional repressor. The direct target of *PtrhAT* is *PtrMYB021*, which is a homolog of *MYB46*  
433 in *Arabidopsis* that acts as a master switch for secondary cell wall biosynthesis, as described above.  
434 By repressing the expression of *PtrhAT*, *PtrEPSP-TF* activates the expression of *PtrMYB021* and  
435 the phenylpropanoid pathway (Xie et al., 2018). **In conclusion, *PtrEPSP-TF/PtrhAT/PtrMYB021***  
436 **form an additional regulatory loop in lignin biosynthesis in *Populus*.**

437 PtrEPSP-TF distinguishes itself from ancestral EPSP synthases by carrying an additional helix-  
438 turn-helix (HTH) motif in the N-terminus (Xie et al., 2018). HTH motifs are commonly found in  
439 transcription factors as nucleic acid binding domains (Aravind et al., 2005). With the addition of  
440 the N-terminal HTH DNA binding motif, PtrEPSP-TF exhibited nuclear accumulation and  
441 functioned as a transcriptional repressor. By comparing 57 EPSP synthase isoforms from 42 plant  
442 genomes, the HTH motif was found to be almost entirely missing in EPSP synthases in  
443 nonvascular, algal, and monocots, but was found in many dicots (Xie et al., 2018). The presence  
444 of secondary cell wall is a key distinguishing feature separating dicots from algae and mosses. It  
445 is intriguing that this shikimate pathway derived-EPSP synthase isoform appears to have obtained  
446 a regulatory function modulating the expression of processes that are ubiquitous in dicots relative  
447 to other plants. With this in mind, we hypothesized that domain co-option may have occurred  
448 during the course of evolution when early dicotyledonous plants attained complex cell wall  
449 structure (Weng et al., 2008a; Tohge et al., 2013). The discovery of the additional regulatory loop  
450 of *MYB46* in *Populus* also supports the existence of woody plant-specific regulatory mechanisms  
451 in lignin biosynthesis.

### 452 **III. Perspectives on the origin and evolution of lignin biosynthesis in plants**

453 The phenylpropanoid pathway produced thousands of metabolites which are essential for plant  
454 terrestrialization and subsequent radiation. Lignins appeared as specialized metabolites with the  
455 evolution of tracheophytes. The identification of progenitors of lignin biosynthetic genes in  
456 bryophytes provides new insights into the origin of lignin biosynthesis (Kreegshauser et al., 2021).  
457 The recent progress on genome sequencing of Charophyte algae, bryophytes, lycophytes, and ferns  
458 have also provided unprecedented opportunities to study the origin of phenylpropanoid  
459 biosynthetic pathway (Szovenyi et al., 2021).

460 Based on current knowledge of lignin biosynthetic pathways across tracheophytes, we conclude  
461 that most lignin biosynthetic genes experienced expansions and neofunctionalization. As a result,  
462 lignin biosynthetic pathway has become increasingly complex evidenced by the existence of many  
463 alternate pathways and regulatory hierarchies. In support of this hypothesis many of the alternative  
464 pathways have been shown to be lineage specific. Lignin biosynthesis in monocots served an  
465 example of diversification. For example, PTAL-mediated by-pass route in lignin biosynthesis and  
466 PMT-mediated lignin modification are specific to monocots (Barros et al., 2016; Petrik et al.,  
467 2014). Equally, S-lignin biosynthesis in *S. mellendorffii* suggested that S-lignin biosynthetic  
468 pathway may be evolved multiple times or lost in gymnosperms and other pteridophytes (Weng  
469 and Chapple, 2010).

470 Transcriptional regulatory modules have been shown to be generally conserved for  
471 phenylpropanoid and lignin biosynthesis; however, a third layer of MYB TFs are not  
472 evolutionarily conserved and have witnessed a wide expansion of family members. Finally, newly  
473 identified TFs, such as EPSP-TF, have been shown to regulate lignin biosynthesis specifically in  
474 woody plants (Xie et al., 2018). The studies on transcriptional regulation of lignin biosynthesis  
475 represents an emerging opportunity to understand the phylogenetic occurrence of the  
476 phenylpropanoid pathway and lignin biosynthesis in plants.

### 477 **Concluding remarks and future directions**

478 In this review, we summarized the phylogenetic occurrence of lignin biosynthetic genes and  
479 related transcriptional regulation across different plant species. Comprehensively, the core

480 enzymes in lignin biosynthesis and basal transcriptional regulatory module **are** conserved among  
481 embryophytes, although bryophytes do not produce lignin. With evolutionary time, lignin  
482 composition diversity **has** increased and **has been** associated with gene duplication, functional gene  
483 co-option, and neo- and sub-functionalization, which involved many structural genes and  
484 transcriptional regulators. In addition, concomitant with the increase of lignin biosynthetic  
485 complexity, is the increase in functional diversity, e.g., water conductivity and defense. As most  
486 of the current knowledge of lignin biosynthesis is based on the study of **a few** angiosperms,  
487 identification and functional characterization of the lignin biosynthetic pathways and their  
488 regulation in lower plants will provide a comprehensive view of their evolutionary history and lead  
489 to new insights in lignin biosynthesis.

490

#### 491 **Conflict of Interest**

492 The authors declare no competing financial interests.

#### 493 **Author Contributions**

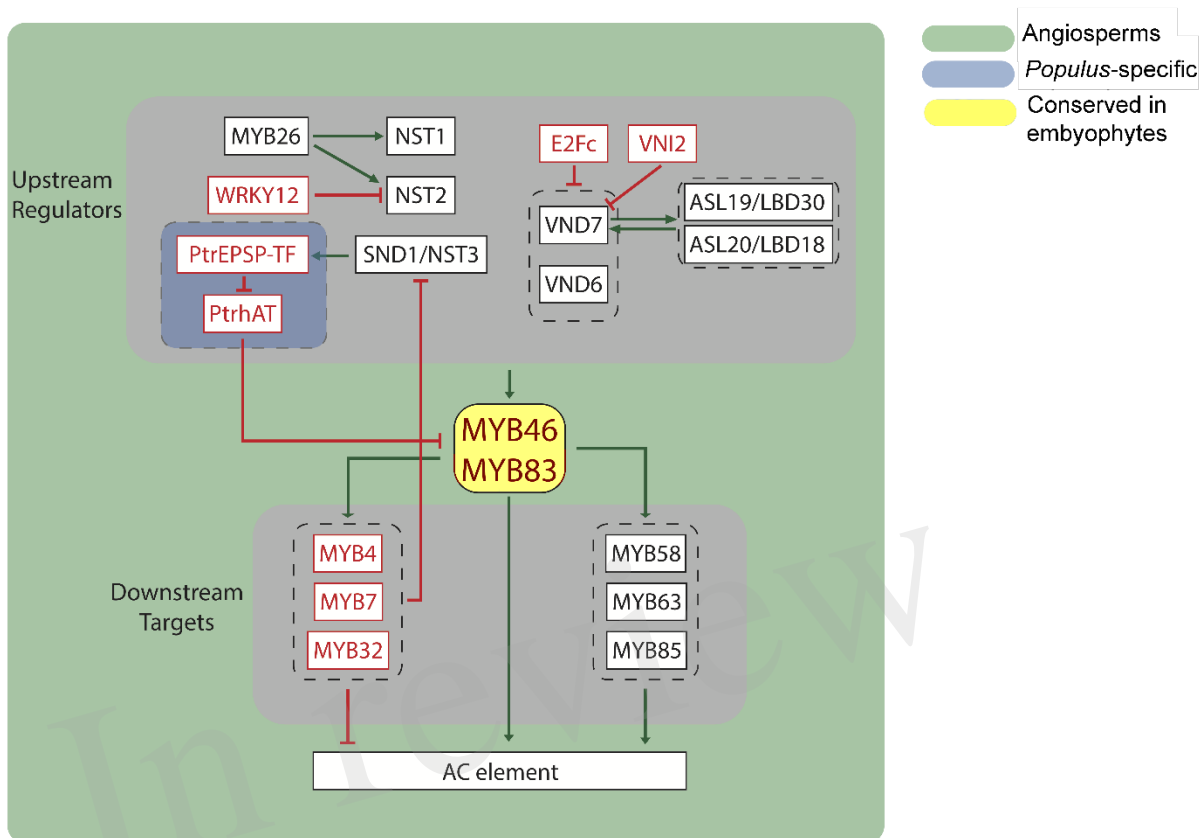
494 T.Y. and K.F. drafted the manuscript. M.X., J.B., T.J.T., G.A.T., W.M. and J.-G.C. revised the  
495 manuscript.

#### 496 **Acknowledgements**

497 This work was supported by the Center for Bioenergy Innovation by the Office of Biological and  
498 Environmental Research in the U.S. Department of Energy Office of Science. Oak Ridge National  
499 Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract  
500 DE-AC05-00OR22725.

501

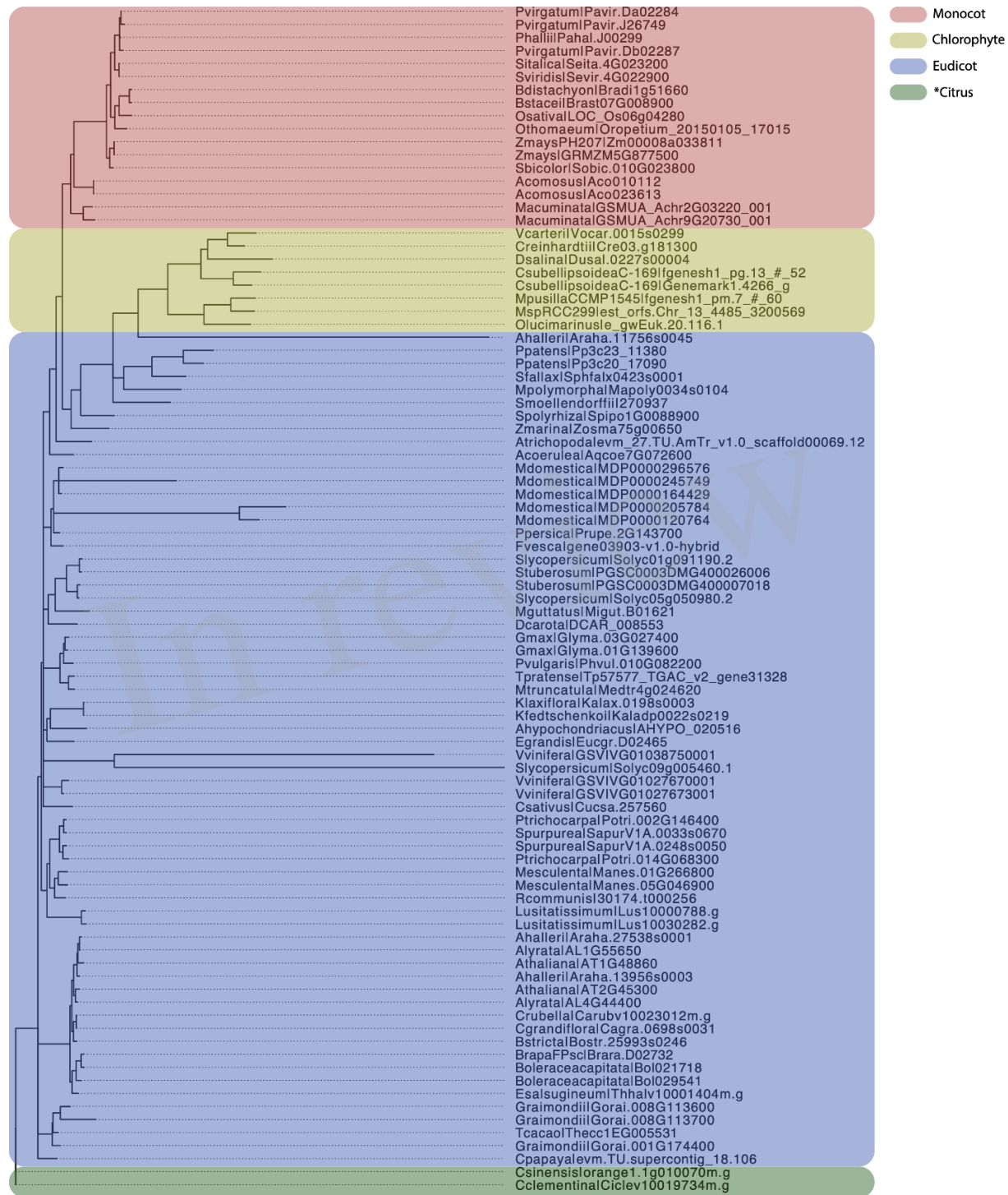




514  
515  
516  
517  
518  
519  
520

**Figure 2.** Transcriptional network of lignin biosynthesis in plants

Black box indicates transcriptional activator, and red box indicates transcriptional repressor. Green arrows indicate transcriptional activation, and red blunt arrows indicate transcriptional repression. AC element, recognized by MYBs, were found in most of lignin biosynthetic genes.



521  
 522 **Figure 3.** Molecular dating of EPSPs.

523 A total of 91 EPSPs, identified by searching against PtrEPSP-TF in phytozome, were used for  
 524 molecular dating analysis. We first used MUSCLE (Edgar, 2004) to perform multiple alignments  
 525 of EPSP proteins, an in-house python script was then used to convert the amino acid alignment to  
 526 nucleotide alignment, and finally TrimAL (Capella-Gutiérrez et al., 2009) was used to trim the

527 alignment using parameters “-gt 0.8 -st 0.001”, which indicate the tolerating gaps of no more than  
528 20% and similarity score less than 0.001. Mrbayes (Huelsenbeck et al., 2001) was used to conduct  
529 molecular dating with parameters “lset nst=6 rates=invgamma” using the “GTR + I +  $\Gamma$ ” model.  
530 A total of 10,000,000 mcmc generations were run after the standard deviation of split frequencies  
531 falls under 0.05. FigTree (Rambaut, 2012) was used to visualize the phylogenetic tree. Number of  
532 each node indicates the posterior probabilities. Pink, yellow, blue and green colors separate  
533 monocot, Chlorpphyte, Eudicot and Citrus, respectively.  
534

In review

535 **Table1** Copy number variation of lignin biosynthetic genes in selected model species

Gene family	<i>C. reinhardtii</i> (green algae)	<i>P. patens</i> (moss)	<i>A. thaliana</i>	<i>P.</i> <i>trichocarpa</i>	<i>O. sativa</i>
<i>PAL</i>	0	26	4	7	9
<i>4CL</i>	0	4	4	7	5
<i>C4H</i>	0	6	1	4	4
<i>CCR</i>	0	1	4	13	14
<i>CAD</i>	1	4	9	26	12
<i>C3H</i>	0	3	3	3	3
<i>F5H</i>	0	0	1	4	3
<i>COMT</i>	0	4	1	13	9
<i>CCoAOMT</i>	0	5	2	2	5
<i>HCT</i>	0	12	1	9	22
<i>CSE</i>	0	1	1	2	1
<i>ALDH</i>	4	37	14	30	21

536

537 **References**

538

539 Anderson, N.A., Tobimatsu, Y., Ciesielski, P.N., Ximenes, E., Ralph, J., Donohoe, B.S., et al.  
540 (2015). Manipulation of guaiacyl and syringyl monomer biosynthesis in an *Arabidopsis*  
541 cinnamyl alcohol dehydrogenase mutant results in atypical lignin biosynthesis and  
542 modified cell wall structure. *Plant Cell* 27, 2195-2209. doi: 10.1105/tpc.15.00373.

543 Aravind, L., Anantharaman, V., Balaji, S., Babu, M.M., and Iyer, L.M. (2005). The many faces of  
544 the helix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol. Rev.*  
545 29, 231-262. doi: 10.1016/j.fmre.2004.12.008.

546 Barakat, A., Yassin, N.B., Park, J.S., Choi, A., Herr, J., and Carlson, J.E. (2011). Comparative and  
547 phylogenomic analyses of cinnamoyl-CoA reductase and cinnamoyl-CoA-reductase-like  
548 gene family in land plants. *Plant Sci.* 181, 249-257. doi: 10.1016/j.plantsci.2011.05.012.

549 Barros, J., Escamilla-Trevino, L., Song, L., Rao, X., Serrani-Yarce, J.C., Palacios, M.D., et al.  
550 (2019). 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic  
551 ascorbate peroxidase. *Nat. Commun.* 10, 1994. doi: 10.1038/s41467-019-10082-7.

552 Barros, J., Serrani-Yarce, J.C., Chen, F., Baxter, D., Venables, B.J., and Dixon, R.A. (2016). Role  
553 of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nat. Plants* 2, 16050. doi:  
554 10.1038/nplants.2016.50.

555 Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin Biosynthesis. *Annu. Rev. Plant Biol.* 54,  
556 519-546. doi: 10.1146/annurev.arplant.54.031902.134938.

557 Bonawitz, N.D., Kim, J.I., Tobimatsu, Y., Ciesielski, P.N., Anderson, N.A., Ximenes, E., et al.  
558 (2014). Disruption of Mediator rescues the stunted growth of a lignin-deficient *Arabidopsis*  
559 mutant. *Nature* 509, 376-380. doi: 10.1038/nature13084.

560 Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., et al. (2017). Insights  
561 into Land Plant Evolution Garnered from the *Marchantia polymorpha* Genome. *Cell* 171,  
562 287-304.e215. doi: <https://doi.org/10.1016/j.cell.2017.09.030>.

563 Brawley, S.H., Blouin, N.A., Ficko-Blean, E., Wheeler, G.L., Lohr, M., Goodson, H.V., et al.  
564 (2017). Insights into the red algae and eukaryotic evolution from the genome of *Porphyra*  
565 *umbilicalis* (Bangiophyceae, Rhodophyta). *Proc. Natl. Acad. Sci. U. S. A.* 114, E6361-  
566 E6370. doi: 10.1073/pnas.1703088114.

567 Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated  
568 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972-1973.  
569 doi: 10.1093/bioinformatics/btp348.

570 Chen, H.-C., Li, Q., Shuford, C.M., Liu, J., Muddiman, D.C., Sederoff, R.R., et al. (2011).  
571 Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid  
572 derivatives in monolignol biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 108, 21253-21258.  
573 doi: 10.1073/pnas.1116416109.

574 Coleman, H.D., Park, J.-Y., Nair, R., Chapple, C., and Mansfield, S.D. (2008). RNAi-mediated  
575 suppression of p-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin  
576 deposition and soluble secondary metabolism. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4501-  
577 4506. doi: 10.1073/pnas.0706537105.

- 578 Davies, K.M., Jibrán, R., Zhou, Y., Albert, N.W., Brummell, D.A., Jordan, B.R., et al. (2020). The  
579 Evolution of Flavonoid Biosynthesis: A Bryophyte Perspective. *Front. Plant Sci.* 11. doi:  
580 10.3389/fpls.2020.00007.
- 581 De La Torre, A.R., Piot, A., Liu, B., Wilhite, B., Weiss, M., and Porth, I. (2020). Functional and  
582 morphological evolution in gymnosperms: A portrait of implicated gene families. *Evol.*  
583 *Appl.* 13, 210-227. doi: <https://doi.org/10.1111/eva.12839>.
- 584 De Meester, B., Madariaga Calderón, B., de Vries, L., Pollier, J., Goeminne, G., Van Doorselaere,  
585 J., et al. (2020). Tailoring poplar lignin without yield penalty by combining a null and  
586 haploinsufficient CINNAMOYL-CoA REDUCTASE2 allele. *Nat. Commun.* 11, 5020. doi:  
587 10.1038/s41467-020-18822-w.
- 588 de Vries, L., Brouckaert, M., Chanoca, A., Kim, H., Regner, M.R., Timokhin, V.I., et al. (2021)  
589 CRISPR-Cas9 editing of CAFFEOYL SHIKIMATE ESTERASE 1 and 2 shows their  
590 importance and partial redundancy in lignification in *Populus tremula* x *P. alba*. *Plant*  
591 *Biotechnol. J.* n/a(n/a). doi: <https://doi.org/10.1111/pbi.13651>.
- 592 Demura, T., and Ye, Z.-H. (2010). Regulation of plant biomass production. *Curr. Opin. Plant Biol.*  
593 13, 298-303. doi: <https://doi.org/10.1016/j.pbi.2010.03.002>.
- 594 Dixon, R.A., and Barros, J. (2019). Lignin biosynthesis: old roads revisited and new roads explored.  
595 *Open Biol.* 9, 190215. doi: 10.1098/rsob.190215.
- 596 Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB  
597 transcription factors in Arabidopsis. *Trends Plant Sci.* 15, 573-581. doi:  
598 <https://doi.org/10.1016/j.tplants.2010.06.005>.
- 599 Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high  
600 throughput. *Nucleic Acids Res.* 32, 1792-1797. doi: 10.1093/nar/gkh340.
- 601 Ehrling, J., Büttner, D., Wang, Q., Douglas, C.J., Somssich, I.E., and Kombrink, E. (1999). Three  
602 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily  
603 divergent classes in angiosperms. *Plant J.* 19, 9-20. doi: 10.1046/j.1365-  
604 313X.1999.00491.x.
- 605 Endo, H., Yamaguchi, M., Tamura, T., Nakano, Y., Nishikubo, N., Yoneda, A., et al. (2014).  
606 Multiple classes of transcription factors regulate the expression of VASCULAR-  
607 RELATED NAC-DOMAIN7, a master switch of xylem vessel differentiation. *Plant Cell*  
608 *Physiol.* 56, 242-254. doi: 10.1093/pcp/pcu134.
- 609 Espiñeira, J.M., Novo Uzal, E., Gómez Ros, L.V., Carrión, J.S., Merino, F., Ros Barceló, A., et al.  
610 (2011). Distribution of lignin monomers and the evolution of lignification among lower  
611 plants. *Plant Biol.* 13, 59-68. doi: 10.1111/j.1438-8677.2010.00345.x.
- 612 Fang, S., Shang, X., Yao, Y., Li, W., and Guo, W. (2020). NST- and SND-subgroup NAC proteins  
613 coordinately act to regulate secondary cell wall formation in cotton. *Plant Sci.* 301, 110657.  
614 doi: <https://doi.org/10.1016/j.plantsci.2020.110657>.
- 615 Fornalé, S., Rencoret, J., Garcia-Calvo, L., Capellades, M., Encina, A., Santiago, R., et al. (2015).  
616 Cell wall modifications triggered by the down-regulation of Coumarate 3-hydroxylase-1  
617 in maize. *Plant Sci.* 236, 272-282. doi: <https://doi.org/10.1016/j.plantsci.2015.04.007>.

- 618 Franke, R., Hemm, M.R., Denault, J.W., Ruegger, M.O., Humphreys, J.M., and Chapple, C.  
619 (2002a). Changes in secondary metabolism and deposition of an unusual lignin in the ref8  
620 mutant of Arabidopsis. *Plant J.* 30, 47-59. doi: <https://doi.org/10.1046/j.1365-313X.2002.01267.x>  
621
- 622 Franke, R., Humphreys, J.M., Hemm, M.R., Denault, J.W., Ruegger, M.O., Cusumano, J.C., et al.  
623 (2002b). The ArabidopsisREF8 gene encodes the 3-hydroxylase of phenylpropanoid  
624 metabolism. *Plant J.* 30, 33-45. doi: [10.1046/j.1365-313X.2002.01266.x](https://doi.org/10.1046/j.1365-313X.2002.01266.x).
- 625 Ge, S., Han, X., Xu, X., Shao, Y., Zhu, Q., Liu, Y., et al. (2020). WRKY15 suppresses tracheary  
626 element differentiation upstream of VND7 during xylem formation. *Plant Cell* 32, 2307-  
627 2324. doi: [10.1105/tpc.19.00689](https://doi.org/10.1105/tpc.19.00689).
- 628 Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., et al. (2005).  
629 EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell  
630 wall formation and lignin biosynthesis. *Plant J.* 43, 553-567. doi:  
631 <https://doi.org/10.1111/j.1365-313X.2005.02480.x>.
- 632 Gui, J., Shen, J., and Li, L. (2011). Functional characterization of evolutionarily divergent 4-  
633 coumarate: coenzyme A ligases in rice. *Plant Physiol.* 157, 574-586. doi:  
634 [10.1104/pp.111.178301](https://doi.org/10.1104/pp.111.178301).
- 635 Guo, D., Chen, F., Inoue, K., Blount, J.W., and Dixon, R.A. (2001). Downregulation of caffeic  
636 acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa.  
637 impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant*  
638 *Cell* 13, 73-88. doi: [10.1105/tpc.13.1.73](https://doi.org/10.1105/tpc.13.1.73).
- 639 Guo, D.-M., Ran, J.-H., and Wang, X.-Q. (2010). Evolution of the cinnamyl/sinapyl alcohol  
640 dehydrogenase (CAD/SAD) gene family: the emergence of real lignin is associated with  
641 the origin of bona fide CAD. *J. Mol. Evol.* 71, 202-218. doi: [10.1007/s00239-010-9378-3](https://doi.org/10.1007/s00239-010-9378-3).
- 642 Guo, Z., Hua, H., Xu, J., Mo, J., Zhao, H., and Yang, J. (2019). Cloning and functional analysis of  
643 lignin biosynthesis genes Cf4CL and CfCCoAOMT in *Cryptomeria fortunei*. *Genes (Basel)*  
644 10. doi: [10.3390/genes10080619](https://doi.org/10.3390/genes10080619).
- 645 Ha, C.M., Escamilla-Trevino, L., Yance, J.C.S., Kim, H., Ralph, J., Chen, F., et al. (2016). An  
646 essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago*  
647 *truncatula*. *Plant J.* 86, 363-375. doi: [10.1111/tpj.13177](https://doi.org/10.1111/tpj.13177).
- 648 Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapiere, C., et al. (2004).  
649 Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinic  
650 hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *Plant Cell* 16, 1446-  
651 1465. doi: [10.1105/tpc.020297](https://doi.org/10.1105/tpc.020297).
- 652 Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees.  
653 *Bioinformatics* 17, 754-755. doi: [10.1093/bioinformatics/17.8.754](https://doi.org/10.1093/bioinformatics/17.8.754).
- 654 Kim, H., Li, Q., Karlen, S.D., Smith, R.A., Shi, R., Liu, J., et al. (2020). Monolignol benzoates  
655 incorporate into the lignin of transgenic *Populus trichocarpa* depleted in C3H and C4H.  
656 *ACS Sustain. Chem. Eng.* 8, 3644-3654. doi: [10.1021/acssuschemeng.9b06389](https://doi.org/10.1021/acssuschemeng.9b06389).

- 657 Kim, W.-C., Ko, J.-H., Kim, J.-Y., Kim, J., Bae, H.-J., and Han, K.-H. (2013). MYB46 directly  
658 regulates the gene expression of secondary wall-associated cellulose synthases in  
659 Arabidopsis. *Plant J.* 73, 26-36. doi: <https://doi.org/10.1111/j.1365-313x.2012.05124.x>.
- 660 Kim, W.-C., Kim, J.-Y., Ko, J.-H., Kang, H., and Han, K.-H. (2014). Identification of direct targets  
661 of transcription factor MYB46 provides insights into the transcriptional regulation of  
662 secondary wall biosynthesis. *Plant Mol. Biol.* 85, 589-599. doi: 10.1007/s11103-014-0205-  
663 x.
- 664 Ko, J.-H., Kim, W.-C., and Han, K.-H. (2009). Ectopic expression of MYB46 identifies  
665 transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis.  
666 *Plant J.* 60, 649-665. doi: <https://doi.org/10.1111/j.1365-313x.2009.03989.x>.
- 667 Ko, J.H., Kim, W.C., Kim, J.Y., Ahn, S.J., and Han, K.H. (2012). MYB46-mediated transcriptional  
668 regulation of secondary wall biosynthesis. *Mol. Plant* 5, 961-963. doi: 10.1093/mp/sss076.
- 669 Kriegshauser, L., Knosp, S., Grienberger, E., Tatsumi, K., Gütle, D.D., Sørensen, I., et al. (2021).  
670 Function of the HYDROXYCINNAMOYL-CoA:SHIKIMATE  
671 HYDROXYCINNAMOYL TRANSFERASE is evolutionarily conserved in  
672 embryophytes. *Plant Cell* 33, 1472-1491. doi: 10.1093/plcell/koab044.
- 673 Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005).  
674 Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19,  
675 1855-1860. doi: 10.1101/gad.1331305.
- 676 Lam, P.Y., Tobimatsu, Y., Matsumoto, N., Suzuki, S., Lan, W., Takeda, Y., et al. (2019).  
677 OsCALDOMT1 is a bifunctional O-methyltransferase involved in the biosynthesis of tricinn-  
678 lignins in rice cell walls. *Sci. Rep.* 9, 11597. doi: 10.1038/s41598-019-47957-0.
- 679 Lauvergeat, V., Lacomme, C., Lacombe, E., Lasserre, E., Roby, D., and Grima-Pettenati, J. (2001).  
680 Two cinnamoyl-CoA reductase (CCR) genes from Arabidopsis thaliana are differentially  
681 expressed during development and in response to infection with pathogenic bacteria.  
682 *Phytochemistry* 57, 1187-1195. doi: [https://doi.org/10.1016/S0031-9422\(01\)00053-X](https://doi.org/10.1016/S0031-9422(01)00053-X).
- 683 Li, L., Cheng, X.F., Leshkevich, J., Umezawa, T., Harding, S.A., and Chiang, V.L. (2001). The  
684 last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene  
685 encoding sinapyl alcohol dehydrogenase. *Plant Cell* 13, 1567-1586. doi:  
686 10.1105/tpc.010111.
- 687 Li, S.S., Chang, Y., Li, B., Shao, S.L., and Zhen-Zhu, Z. (2020). Functional analysis of 4-  
688 coumarate: CoA ligase from Dryopteris fragrans in transgenic tobacco enhances lignin and  
689 flavonoids. *Genet. Mol. Biol.* 43, e20180355. doi: 10.1590/1678-4685-gmb-2018-0355.
- 690 Li, X., Bonawitz, N.D., Weng, J.-K., and Chapple, C. (2010). The Growth Reduction Associated  
691 with Repressed Lignin Biosynthesis in Arabidopsis thaliana Is Independent of Flavonoids.  
692 *Plant Cell* 22, 1620-1632. doi: 10.1105/tpc.110.074161.
- 693 Li, X., Chen, W., Zhao, Y., Xiang, Y., Jiang, H., Zhu, S., et al. (2013). Downregulation of caffeoyl-  
694 CoA O-methyltransferase (CCoAOMT) by RNA interference leads to reduced lignin  
695 production in maize straw. *Genet. Mol. Biol.* 36, 540-546. doi: 10.1590/s1415-  
696 47572013005000039.

- 697 Li, Y., Kim, J.I., Pysh, L., and Chapple, C. (2015). Four isoforms of Arabidopsis 4-coumarate:  
698 CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant*  
699 *Physiol.* 169(4), 2409-2421. doi: 10.1104/pp.15.00838.
- 700 Liu, Y., You, S., Taylor-Teeples, M., Li, W.L., Schuetz, M., Brady, S.M., et al. (2014). BEL1-  
701 LIKE HOMEODOMAIN6 and KNOTTED ARABIDOPSIS THALIANA7 interact and  
702 regulate secondary cell wall formation via repression of REVOLUTA. *Plant Cell* 26, 4843-  
703 4861. doi: 10.1105/tpc.114.128322.
- 704 Logan, K.J., and Thomas, B.A. (1985). Distribution of Lignin Derivatives in Plants. *New Phytol.*  
705 99, 571-585. doi: <https://doi.org/10.1111/j.1469-8137.1985.tb03685.x>.
- 706 Ma, D., and Constabel, C.P. (2019). MYB Repressors as Regulators of Phenylpropanoid  
707 Metabolism in Plants. *Trends Plant Sci.* 24, 275-289. doi:  
708 <https://doi.org/10.1016/j.tplants.2018.12.003>.
- 709 Mansfield, S.D., Kim, H., Lu, F., and Ralph, J. (2012). Whole plant cell wall characterization using  
710 solution-state 2D NMR. *Nat. Protoc.* 7, 1579-1589. doi: 10.1038/nprot.2012.064.
- 711 Martin, A.F., Tobimatsu, Y., Kusumi, R., Matsumoto, N., Miyamoto, T., Lam, P.Y., et al. (2019).  
712 Altered lignocellulose chemical structure and molecular assembly in CINNAMYL  
713 ALCOHOL DEHYDROGENASE-deficient rice. *Sci. Rep.* 9, 17153. doi: 10.1038/s41598-  
714 019-53156-8.
- 715 Martone, P.T., Estevez, J.M., Lu, F., Ruel, K., Denny, M.W., Somerville, C., et al. (2009).  
716 Discovery of Lignin in Seaweed Reveals Convergent Evolution of Cell-Wall Architecture.  
717 *Curr. Biol.* 19, 169-175. doi: <https://doi.org/10.1016/j.cub.2008.12.031>.
- 718 McCarthy, R.L., Zhong, R., and Ye, Z.-H. (2009). MYB83 Is a Direct Target of SND1 and Acts  
719 Redundantly with MYB46 in the Regulation of Secondary Cell Wall Biosynthesis in  
720 Arabidopsis. *Plant Cell Physiol.* 50, 1950-1964. doi: 10.1093/pcp/pcp139.
- 721 Meyer, K., Shirley, A.M., Cusumano, J.C., Bell-Lelong, D.A., and Chapple, C. (1998). Lignin  
722 monomer composition is determined by the expression of a cytochrome P450-dependent  
723 monooxygenase in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 95, 6619-6623. doi:  
724 10.1073/pnas.95.12.6619.
- 725 Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (2007). NAC  
726 Transcription Factors, NST1 and NST3, Are Key Regulators of the Formation of  
727 Secondary Walls in Woody Tissues of Arabidopsis. *Plant Cell* 19, 270-280. doi:  
728 10.1105/tpc.106.047043.
- 729 Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2005). The NAC transcription factors  
730 NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for  
731 anther dehiscence. *Plant Cell* 17, 2993-3006. doi: 10.1105/tpc.105.036004.
- 732 Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N.A., and Ohtani, M. (2015). NAC-MYB-based  
733 transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant*  
734 *Sci.* 6. doi: 10.3389/fpls.2015.00288.
- 735 Nawawi, D.S., Syafii, W., Akiyama, T., and Matsumoto, Y. (2016). Characteristics of guaiacyl-  
736 syringyl lignin in reaction wood in the gymnosperm *Gnetum gnemon* L. *Holzforchung* 70,  
737 593-602. doi: <https://doi.org/10.1515/hf-2015-0107>.

- 738 Niklas, K.J., Cobb, E.D., and Matas, A.J. (2017). The evolution of hydrophobic cell wall  
739 biopolymers: from algae to angiosperms. *J. Exp. Bot.* 68, 5261-5269. doi:  
740 10.1093/jxb/erx215.
- 741 Ohashi-Ito, K., Oda, Y., and Fukuda, H. (2010). Arabidopsis VASCULAR-RELATED NAC-  
742 DOMAIN6 directly regulates the genes that govern programmed cell death and secondary  
743 wall formation during xylem differentiation. *Plant Cell* 22, 3461-3473. doi:  
744 10.1105/tpc.110.075036.
- 745 Ohtani, M., and Demura, T. (2019). The quest for transcriptional hubs of lignin biosynthesis:  
746 beyond the NAC-MYB-gene regulatory network model. *Curr. Opin. Biotechnol.* 56, 82-  
747 87. doi: <https://doi.org/10.1016/j.copbio.2018.10.002>.
- 748 Olsen, A.N., Ernst, H.A., Leggio, L.L., and Skriver, K. (2005). NAC transcription factors:  
749 structurally distinct, functionally diverse. *Trends Plant Sci.* 10, 79-87. doi:  
750 <https://doi.org/10.1016/j.tplants.2004.12.010>.
- 751 Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L.J., Smith, C., et al. (2003).  
752 Characterisation of a pine MYB that regulates lignification. *Plant J.* 36, 743-754. doi:  
753 <https://doi.org/10.1046/j.1365-313X.2003.01916.x>.
- 754 Petrik, D.L., Karlen, S.D., Cass, C.L., Padmakshan, D., Lu, F., Liu, S., et al. (2014). p-Coumaroyl-  
755 CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in  
756 *Brachypodium distachyon*. *Plant J.* 77, 713-726. doi: <https://doi.org/10.1111/tpj.12420>.
- 757 Qin, W., Yin, Q., Chen, J., Zhao, X., Yue, F., He, J., et al. (2020). The class II KNOX transcription  
758 factors KNAT3 and KNAT7 synergistically regulate monolignol biosynthesis in  
759 Arabidopsis. *J. Exp. Bot.* 71, 5469-5483. doi: 10.1093/jxb/eraa266.
- 760 Raes, J., Rohde, A., Christensen, J.H., Van de Peer, Y., and Boerjan, W. (2003). Genome-Wide  
761 Characterization of the Lignification Toolbox in Arabidopsis. *Plant Physiol.* 133, 1051-  
762 1071. doi: 10.1104/pp.103.026484.
- 763 Ralph, J., Lapierre, C., and Boerjan, W. (2019). Lignin structure and its engineering. *Curr. Opin.*  
764 *Biotechnol.* 56, 240-249. doi: 10.1016/j.copbio.2019.02.019.
- 765 Rambaut A. (2012) *Figtree v1.4*. <http://tree.bio.ed.ac.uk/software/figtree/>
- 766 Rao, G., Pan, X., Xu, F., Zhang, Y., Cao, S., Jiang, X., et al. (2015). Divergent and overlapping  
767 function of five 4-coumarate/coenzyme A ligases from *Populus tomentosa*. *Plant Mol. Biol.*  
768 *Rep.* 33, 841-854. doi: 10.1007/s11105-014-0803-4.
- 769 Rao, X., Chen, X., Shen, H., Ma, Q., Li, G., Tang, Y., et al. (2019). Gene regulatory networks for  
770 lignin biosynthesis in switchgrass (*Panicum virgatum*). *Plant Biotechnol J.* 17, 580-593.  
771 doi: <https://doi.org/10.1111/pbi.13000>.
- 772 Rao, X., and Dixon, R.A. (2018). Current models for transcriptional regulation of secondary cell  
773 wall biosynthesis in grasses. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00399.
- 774 Renault, H., De Marothy, M., Jonasson, G., Lara, P., Nelson, D.R., Nilsson, I., et al. (2017). Gene  
775 duplication leads to altered membrane topology of a cytochrome P450 enzyme in seed  
776 plants. *Mol. Biol. Evol.* 34, 2041-2056. doi: 10.1093/molbev/msx160.

- 777 Renault, H., Werck-Reichhart, D., and Weng, J.-K. (2019). Harnessing lignin evolution for  
778 biotechnological applications. *Curr. Opin. Biotechnol.* 56, 105-111. doi:  
779 <https://doi.org/10.1016/j.copbio.2018.10.011>.
- 780 Rensing, S.A. (2018). Great moments in evolution: the conquest of land by plants. *Curr. Opin.*  
781 *Plant Biol.* 42, 49-54. doi: <https://doi.org/10.1016/j.pbi.2018.02.006>.
- 782 Ruel, K., Berrio-Sierra, J., Derikvand, M.M., Pollet, B., Thévenin, J., Lapiere, C., et al. (2009).  
783 Impact of CCR1 silencing on the assembly of lignified secondary walls in *Arabidopsis*  
784 *thaliana*. *New Phytol.* 184, 99-113. doi: 10.1111/j.1469-8137.2009.02951.x.
- 785 Saleme, M.d.L.S., Cesarino, I., Vargas, L., Kim, H., Vanholme, R., Goeminne, G., et al. (2017).  
786 Silencing CAFFEYOYL SHIKIMATE ESTERASE Affects Lignification and Improves  
787 Saccharification in Poplar. *Plant Physiol.* 175, 1040-1057. doi: 10.1104/pp.17.00920.
- 788 Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P., et al. (2001). CYP98A3  
789 from *Arabidopsis thaliana* Is a 3'-Hydroxylase of Phenolic Esters, a Missing Link in the  
790 Phenylpropanoid Pathway. *J. Biol. Chem.* 276, 36566-36574. doi:  
791 <https://doi.org/10.1074/jbc.M104047200>.
- 792 Schuetz, M., Smith, R., and Ellis, B. (2013). Xylem tissue specification, patterning, and  
793 differentiation mechanisms. *J. Exp. Bot.* 64, 11-31. doi: 10.1093/jxb/ers287.
- 794 Serrani-Yarce, J.C., Escamilla-Trevino, L., Barros, J., Gallego-Giraldo, L., Pu, Y., Ragauskas, A.,  
795 et al. (2021). Targeting hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase  
796 for lignin modification in *Brachypodium distachyon*. *Biotechnol. Biofuels.* 14, 50. doi:  
797 10.1186/s13068-021-01905-1.
- 798 Shi, R., Sun, Y.-H., Li, Q., Heber, S., Sederoff, R., and Chiang, V.L. (2009). Towards a Systems  
799 Approach for Lignin Biosynthesis in *Populus trichocarpa*: Transcript Abundance and  
800 Specificity of the Monolignol Biosynthetic Genes. *Plant Cell Physiol.* 51, 144-163. doi:  
801 10.1093/pcp/pcp175.
- 802 Silber, M.V., Meimberg, H., and Ebel, J. (2008). Identification of a 4-coumarate:CoA ligase gene  
803 family in the moss, *Physcomitrella patens*. *Phytochemistry* 69, 2449-2456. doi:  
804 <https://doi.org/10.1016/j.phytochem.2008.06.014>.
- 805 Sonnante, G., D'Amore, R., Blanco, E., Pierri, C.L., De Palma, M., Luo, J., et al. (2010). Novel  
806 hydroxycinnamoyl-coenzyme A quinate transferase genes from artichoke are involved in  
807 the synthesis of chlorogenic acid. *Plant Physiol.* 153, 1224-1238. doi:  
808 10.1104/pp.109.150144.
- 809 Soyano, T., Thitamadee, S., Machida, Y., and Chua, N.H. (2008). ASYMMETRIC LEAVES2-  
810 LIKE19/LATERAL ORGAN BOUNDARIES DOMAIN30 and ASL20/LBD18 regulate  
811 tracheary element differentiation in *Arabidopsis*. *Plant Cell* 20, 3359-3373. doi:  
812 10.1105/tpc.108.061796.
- 813 Szövényi, P., Gunadi, A., and Li, F.-W. (2021). Charting the genomic landscape of seed-free plants.  
814 *Nat. Plants* 7, 554-565. doi: 10.1038/s41477-021-00888-z.
- 815 Taylor-Teeple, M., Lin, L., de Lucas, M., Turco, G., Toal, T.W., Gaudinier, A., et al. (2015). An  
816 *Arabidopsis* gene regulatory network for secondary cell wall synthesis. *Nature* 517, 571-  
817 575. doi: 10.1038/nature14099.

- 818 Takeda, Y., Koshiba, T., Tobimatsu, Y., Suzuki, S., Murakami, S., Yamamura, M., et al. (2017).  
819 Regulation of CONIFERALDEHYDE 5-HYDROXYLASE expression to modulate cell  
820 wall lignin structure in rice. *Planta* 246, 337-349. doi: 10.1007/s00425-017-2692-x.
- 821 Takeda, Y., Suzuki, S., Tobimatsu, Y., Osakabe, K., Osakabe, Y., Ragamustari, S.K., et al. (2019).  
822 Lignin characterization of rice CONIFERALDEHYDE 5-HYDROXYLASE loss-of-  
823 function mutants generated with the CRISPR/Cas9 system. *Plant J* 97, 543-554. doi:  
824 10.1111/tpj.14141.
- 825 Tohge, T., Watanabe, M., Hoefgen, R., and Fernie, A.R. (2013). The evolution of phenylpropanoid  
826 metabolism in the green lineage. *Crit. Rev. Biochem. Mol. Biol.* 48, 123-152. doi:  
827 10.3109/10409238.2012.758083.
- 828 Tomlinson, P.B. (2001). Reaction tissues in *Gnetum gnemon*—a preliminary report. *IAWA J.* 22,  
829 401-413. doi: <https://doi.org/10.1163/22941932-90000385>.
- 830 Tsai, C.-J., Xu, P., Xue, L.-J., Hu, H., Nyamdari, B., Naran, R., et al. (2020). Compensatory  
831 Guaiacyl Lignin Biosynthesis at the Expense of Syringyl Lignin in *4CL1*-Knockout Poplar.  
832 *Plant Physiol.* 183, 123-136. doi: 10.1104/pp.19.01550.
- 833 Valdivia, E.R., Herrera, M.T., Gianzo, C., Fidalgo, J., Revilla, G., Zarra, I., et al. (2013).  
834 Regulation of secondary wall synthesis and cell death by NAC transcription factors in the  
835 monocot *Brachypodium distachyon*. *J. Exp. Bot.* 64, 1333-1343. doi: 10.1093/jxb/ers394.
- 836 Van Acker, R., Leplé, J.-C., Aerts, D., Storme, V., Goeminne, G., Ivens, B., et al. (2014). Improved  
837 saccharification and ethanol yield from field-grown transgenic poplar deficient in  
838 cinnamoyl-CoA reductase. *Proc. Natl. Acad. Sci. U. S. A.* 111, 845-850. doi:  
839 10.1073/pnas.1321673111.
- 840 Van Acker, R., Déjardin, A., Desmet, S., Hoengenaert, L., Vanholme, R., Morreel, K., et al. (2017).  
841 Different routes for conifer-and sinapaldehyde and higher saccharification upon deficiency  
842 in the dehydrogenase *CAD1*. *Plant Physiol.* 175, 1018-1039. doi: 10.1104/pp.17.00834.
- 843 Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., et al. (2013). Caffeoyl  
844 shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*.  
845 *Science* 341, 1103-1106. doi: 10.1126/science.1241602.
- 846 Wagner, A., Ralph, J., Akiyama, T., Flint, H., Phillips, L., Torr, K., et al. (2007). Exploring  
847 lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate  
848 hydroxycinnamoyltransferase in *Pinus radiata*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 11856-  
849 11861. doi: 10.1073/pnas.0701428104.
- 850 Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Torr, K., Donaldson, L., et al. (2011).  
851 CCoAOMT suppression modifies lignin composition in *Pinus radiata*. *Plant J.* 67, 119-129.  
852 doi: 10.1111/j.1365-313X.2011.04580.x.
- 853 Wang, H., Avci, U., Nakashima, J., Hahn, M.G., Chen, F., and Dixon, R.A. (2010). Mutation of  
854 WRKY transcription factors initiates pith secondary wall formation and increases stem  
855 biomass in dicotyledonous plants. *Proc. Natl. Acad. Sci. U. S. A.* 107, 22338-22343. doi:  
856 10.1073/pnas.1016436107.

- 857 Wang, H., Zhao, Q., Chen, F., Wang, M., and Dixon, R.A. (2011). NAC domain function and  
858 transcriptional control of a secondary cell wall master switch. *Plant J.* 68, 1104-1114. doi:  
859 10.1111/j.1365-313X.2011.04764.x.
- 860 Wang, S., Yamaguchi, M., Grienberger, E., Martone, P.T., Samuels, A.L., and Mansfield, S.D.  
861 (2020). The Class II KNOX genes KNAT3 and KNAT7 work cooperatively to influence  
862 deposition of secondary cell walls that provide mechanical support to Arabidopsis stems.  
863 *Plant J.* 101, 293-309. doi: <https://doi.org/10.1111/tpj.14541>.
- 864 Wang, X., Chao, N., Zhang, M., Jiang, X., and Gai, Y. (2019). Functional Characteristics of  
865 Caffeoyl Shikimate Esterase in *Larix Kaempferi* and Monolignol Biosynthesis in  
866 Gymnosperms. *Int. J. Mol. Sci.* 20. doi: 10.3390/ijms20236071.
- 867 Wang, X., Zhuo, C., Xiao, X., Wang, X., Docampo-Palacios, M.L., Chen, F., et al. (2020).  
868 Substrate-Specificity of LACCASE 8 Facilitates Polymerization of Caffeoyl Alcohol for C-  
869 Lignin Biosynthesis in the Seed Coat of *Cleome hassleriana*. *Plant Cell*, tpc.00598.02020.  
870 doi: 10.1105/tpc.20.00598.
- 871 Welker, C.M., Balasubramanian, V.K., Petti, C., Rai, K.M., DeBolt, S., and Mendu, V. (2015).  
872 Engineering plant biomass lignin content and composition for biofuels and bioproducts.  
873 *Energies* 8, 7654-7676.
- 874 Weng, J.-K., Akiyama, T., Ralph, J., and Chapple, C. (2011). Independent recruitment of an O-  
875 methyltransferase for syringyl lignin biosynthesis in *Selaginella moellendorffii*. *Plant Cell*  
876 23, 2708-2724. doi: 10.1105/tpc.110.081547.
- 877 Weng, J.-K., and Chapple, C. (2010). The origin and evolution of lignin biosynthesis. *New Phytol.*  
878 187, 273-285. doi: <https://doi.org/10.1111/j.1469-8137.2010.03327.x>.
- 879 Weng, J.K., Li, X., Bonawitz, N.D., and Chapple, C. (2008a). Emerging strategies of lignin  
880 engineering and degradation for cellulosic biofuel production. *Curr. Opin. Biotechnol.* 19,  
881 166-172. doi: 10.1016/j.copbio.2008.02.014.
- 882 Weng, J.K., Li, X., Stout, J., and Chapple, C. (2008b). Independent origins of syringyl lignin in  
883 vascular plants. *Proc. Natl. Acad. Sci. U. S. A.* 105, 7887-7892. doi:  
884 10.1073/pnas.0801696105.
- 885 Withers, S., Lu, F., Kim, H., Zhu, Y., Ralph, J., and Wilkerson, C.G. (2012). Identification of  
886 Grass-specific Enzyme That Acylates Monolignols with *p*-Coumarate. *J. Biol. Chem.* 287,  
887 8347-8355. doi: <https://doi.org/10.1074/jbc.M111.284497>.
- 888 Wu, Z., Wang, N., Hisano, H., Cao, Y., Wu, F., Liu, W., et al. (2019). Simultaneous regulation of  
889 F5H in COMT-RNAi transgenic switchgrass alters effects of COMT suppression on  
890 syringyl lignin biosynthesis. *Plant Biotechnol. J.* 17, 836-845. doi: 10.1111/pbi.13019.
- 891 Xiao, Y., Li, J., Liu, H., Zhang, Y., Zhang, X., Qin, Z., et al. (2020). The Effect of Co-  
892 Transforming *Eucalyptus urophylla* Catechol-O-methyltransferase and Caffeoyl-CoA O-  
893 methyltransferase on the Biosynthesis of Lignin Monomers in Transgenic Tobacco. *Russ.*  
894 *J. Plant Physiol.* 67, 879-887. doi: 10.1134/S1021443720050180.
- 895 Xie, M., Muchero, W., Bryan, A.C., Yee, K., Guo, H.B., Zhang, J., et al. (2018). A 5-  
896 Enolpyruvylshikimate 3-Phosphate Synthase Functions as a Transcriptional Repressor in  
897 *Populus*. *Plant Cell* 30, 1645-1660. doi: 10.1105/tpc.18.00168.

- 898 Xiong, W., Wu, Z., Liu, Y., Li, Y., Su, K., Bai, Z., et al. (2019). Mutation of 4-coumarate:  
899 coenzyme A ligase 1 gene affects lignin biosynthesis and increases the cell wall  
900 digestibility in maize brown *midrib5* mutants. *Biotechnol. Biofuels.* 12, 82. doi:  
901 10.1186/s13068-019-1421-z.
- 902 Xu, B., Ohtani, M., Yamaguchi, M., Toyooka, K., Wakazaki, M., Sato, M., et al. (2014).  
903 Contribution of NAC Transcription Factors to Plant Adaptation to Land. *Science* 343,  
904 1505-1508. doi: 10.1126/science.1248417.
- 905 Xu, Z., Zhang, D., Hu, J., Zhou, X., Ye, X., Reichel, K.L., et al. (2009). Comparative genome  
906 analysis of lignin biosynthesis gene families across the plant kingdom. *BMC*  
907 *Bioinformatics* 10, S3. doi: 10.1186/1471-2105-10-S11-S3.
- 908 Yamaguchi, M., Kubo, M., Fukuda, H., and Demura, T. (2008). Vascular-related NAC-DOMAIN7  
909 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and  
910 shoots. *Plant J.* 55, 652-664. doi: 10.1111/j.1365-313X.2008.03533.x.
- 911 Yamaguchi, M., Mitsuda, N., Ohtani, M., Ohme-Takagi, M., Kato, K., and Demura, T. (2011).  
912 VASCULAR-RELATED NAC-DOMAIN7 directly regulates the expression of a broad  
913 range of genes for xylem vessel formation. *Plant J.* 66, 579-590. doi: 10.1111/j.1365-  
914 313X.2011.04514.x.
- 915 Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., et al. (2010).  
916 VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem  
917 vessel formation in Arabidopsis. *Plant Cell* 22, 1249-1263. doi: 10.1105/tpc.108.064048.
- 918 Yan, X., Liu, J., Kim, H., Liu, B., Huang, X., Yang, Z., et al. (2019). CAD1 and CCR2 protein  
919 complex formation in monolignol biosynthesis in *Populus trichocarpa*. *New Phytol.* 222,  
920 244-260. doi: 10.1111/nph.15505.
- 921 Yang, C., Xu, Z., Song, J., Conner, K., Vizcay Barrena, G., and Wilson, Z.A. (2007). Arabidopsis  
922 MYB26/MALE STERILE35 Regulates Secondary Thickening in the Endothecium and Is  
923 Essential for Anther Dehiscence. *Plant Cell* 19, 534-548. doi: 10.1105/tpc.106.046391.
- 924 Yang, X., Beres, Z.T., Jin, L., Parrish, J.T., Zhao, W., Mackey, D., et al. (2017). Effects of over-  
925 expressing a native gene encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS)  
926 on glyphosate resistance in Arabidopsis thaliana. *PLoS One* 12, e0175820. doi:  
927 10.1371/journal.pone.0175820.
- 928 Zhang, J., Tuskan, G.A., Tschaplinski, T.J., Muchero, W., and Chen, J.-G. (2020). Transcriptional  
929 and Post-transcriptional Regulation of Lignin Biosynthesis Pathway Genes in *Populus*.  
930 *Front. Plant Sci.* 11, 652-652. doi: 10.3389/fpls.2020.00652.
- 931 Zhang, Q., Luo, F., Zhong, Y., He, J., and Li, L. (2019). Modulation of NAC transcription factor  
932 NST1 activity by XYLEM NAC DOMAIN1 regulates secondary cell wall formation in  
933 Arabidopsis. *J. Exp. Bot.* 71, 1449-1458. doi: 10.1093/jxb/erz513.
- 934 Zhang, X.-S., Ni, R., Wang, P.-Y., Zhu, T.-T., Sun, C.-J., Lou, H.-X., et al. (2019). Isolation and  
935 functional characterization of two Caffeoyl Coenzyme A 3-O-methyltransferases from the  
936 fern species *Polypodiodes amoena*. *Plant Physiol. Biochem.* 136, 169-177. doi:  
937 <https://doi.org/10.1016/j.plaphy.2019.01.021>.

- 938 Zhao, K., and Bartley, L.E. (2014). Comparative genomic analysis of the R2R3 MYB secondary  
939 cell wall regulators of Arabidopsis, poplar, rice, maize, and switchgrass. *BMC Plant Biol.*  
940 14, 135. doi: 10.1186/1471-2229-14-135.
- 941 Zhao, Q., Gallego-Giraldo, L., Wang, H., Zeng, Y., Ding, S.-Y., Chen, F., et al. (2010). An NAC  
942 transcription factor orchestrates multiple features of cell wall development in *Medicago*  
943 *truncatula*. *Plant J.* 63, 100-114. doi: <https://doi.org/10.1111/j.1365-313X.2010.04223.x>.
- 944 Zhong, R., Kandasamy, M.K., and Ye, Z.-H. (2020). XND1 regulates secondary wall deposition  
945 in xylem vessels through the inhibition of VND functions. *Plant Cell Physiol.* 62, 53-65.  
946 doi: 10.1093/pcp/pcaa140.
- 947 Zhong, R., Lee, C., McCarthy, R.L., Reeves, C.K., Jones, E.G., and Ye, Z.-H. (2011).  
948 Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and  
949 MYB transcription factors. *Plant Cell Physiol.* 52, 1856-1871. doi: 10.1093/pcp/pcr123.
- 950 Zhong, R., Lee, C., and Ye, Z.-H. (2010). Evolutionary conservation of the transcriptional network  
951 regulating secondary cell wall biosynthesis. *Trends Plant Sci.* 15, 625-632. doi:  
952 <https://doi.org/10.1016/j.tplants.2010.08.007>.
- 953 Zhong, R., Lee, C., Zhou, J., McCarthy, R.L., and Ye, Z.-H. (2008). A battery of transcription  
954 factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant*  
955 *Cell* 20, 2763-2782. doi: 10.1105/tpc.108.061325.
- 956 Zhong, R., McCarthy, R.L., Haghghat, M., and Ye, Z.-H. (2013). The poplar MYB master  
957 switches bind to the SMRE site and activate the secondary wall biosynthetic program  
958 during wood formation. *PLoS One* 8, e69219. doi: 10.1371/journal.pone.0069219.
- 959 Zhong, R., Morrison, W.H., 3rd, Himmelsbach, D.S., Poole, F.L., 2nd, and Ye, Z.H. (2000).  
960 Essential role of caffeoyl coenzyme A O-methyltransferase in lignin biosynthesis in woody  
961 poplar plants. *Plant Physiol.* 124, 563-578. doi: 10.1104/pp.124.2.563.
- 962 Zhong, R., Richardson, E.A., and Ye, Z.-H. (2007). The MYB46 transcription factor is a direct  
963 target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. *Plant Cell* 19,  
964 2776-2792. doi: 10.1105/tpc.107.053678.
- 965 Zhong, R., and Ye, Z.H. (2009). Transcriptional regulation of lignin biosynthesis. *Plant Signal*  
966 *Behav.* 4, 1028-1034. doi: 10.4161/psb.4.11.9875.
- 967 Zhong, R., and Ye, Z.-H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate  
968 a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.*  
969 53, 368-380. doi: 10.1093/pcp/pcr185.
- 970 Zhou, J., Zhong, R., and Ye, Z.-H. (2014). Arabidopsis NAC domain proteins, VND1 to VND5,  
971 are transcriptional regulators of secondary wall biosynthesis in vessels. *PLoS One* 9,  
972 e105726. doi: 10.1371/journal.pone.0105726.
- 973 Zhou, X., Yang, S., Lu, M., Zhao, S., Cai, L., Zhang, Y., et al. (2020). Structure and monomer  
974 ratio of lignin in *C3H* and *HCT* RNAi transgenic poplar saplings. *ChemistrySelect* 5, 7164-  
975 7169. doi: 10.1002/slct.202000365.

976

Figure 1.TIF

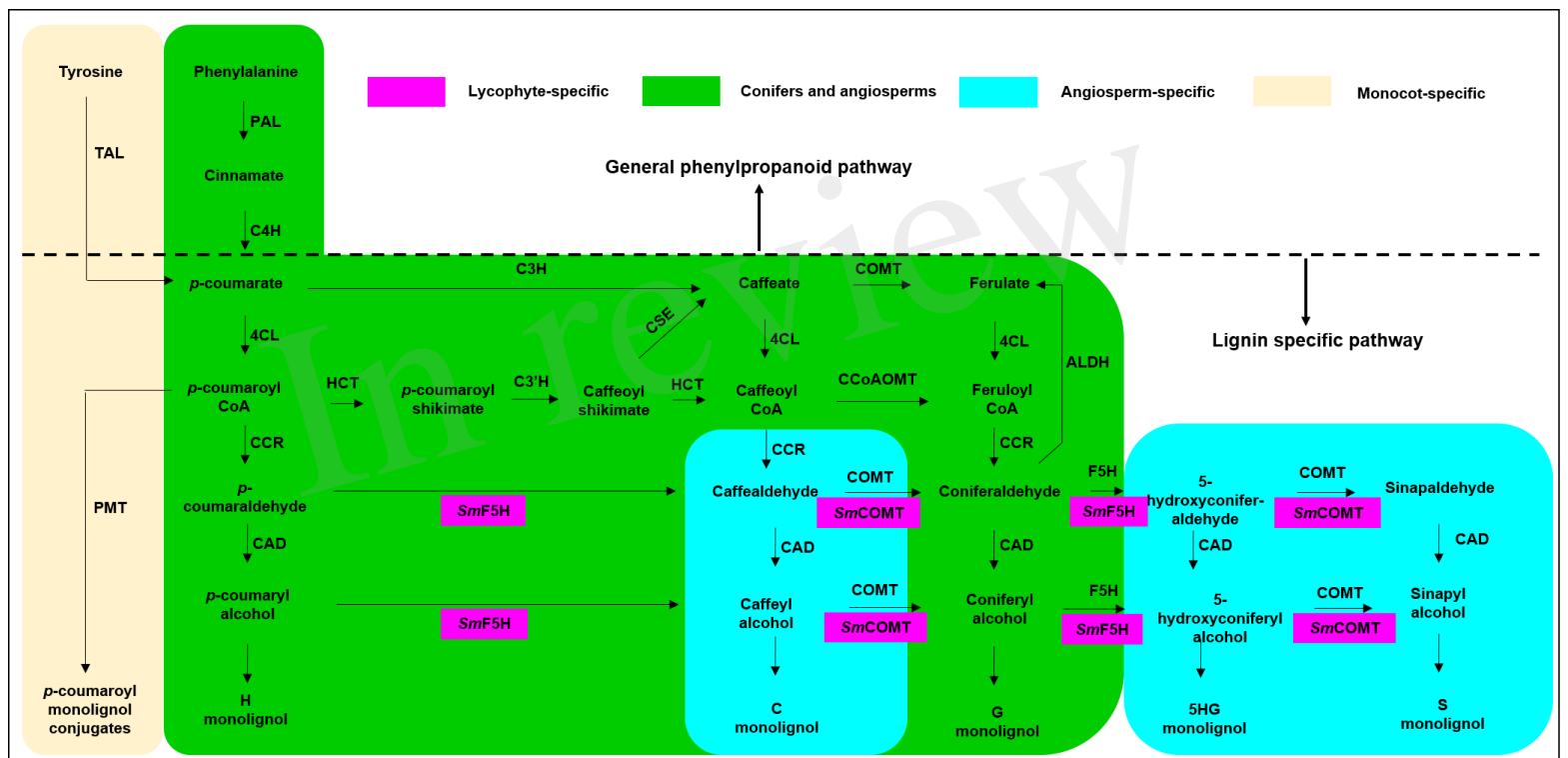


Figure 2.TIF

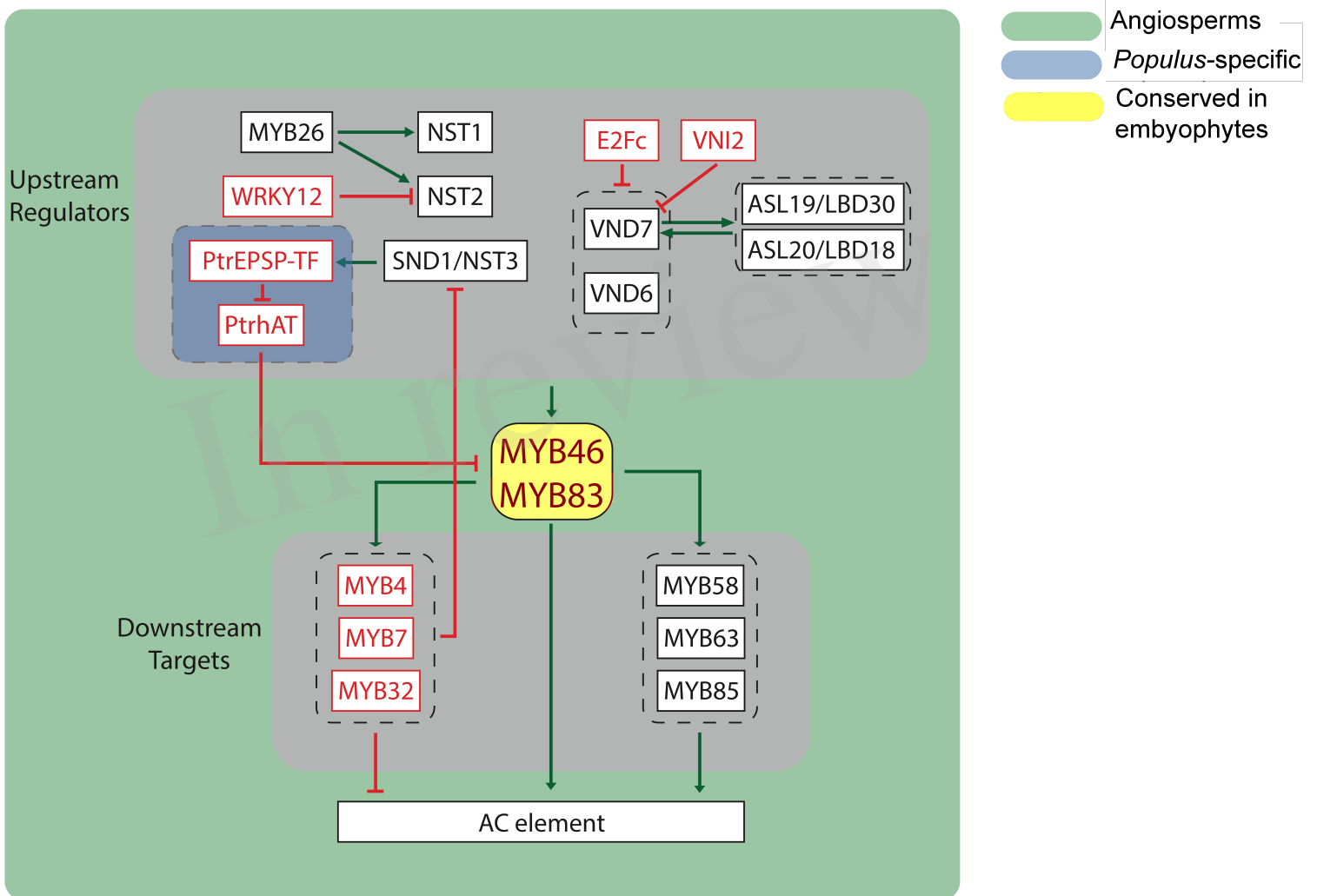
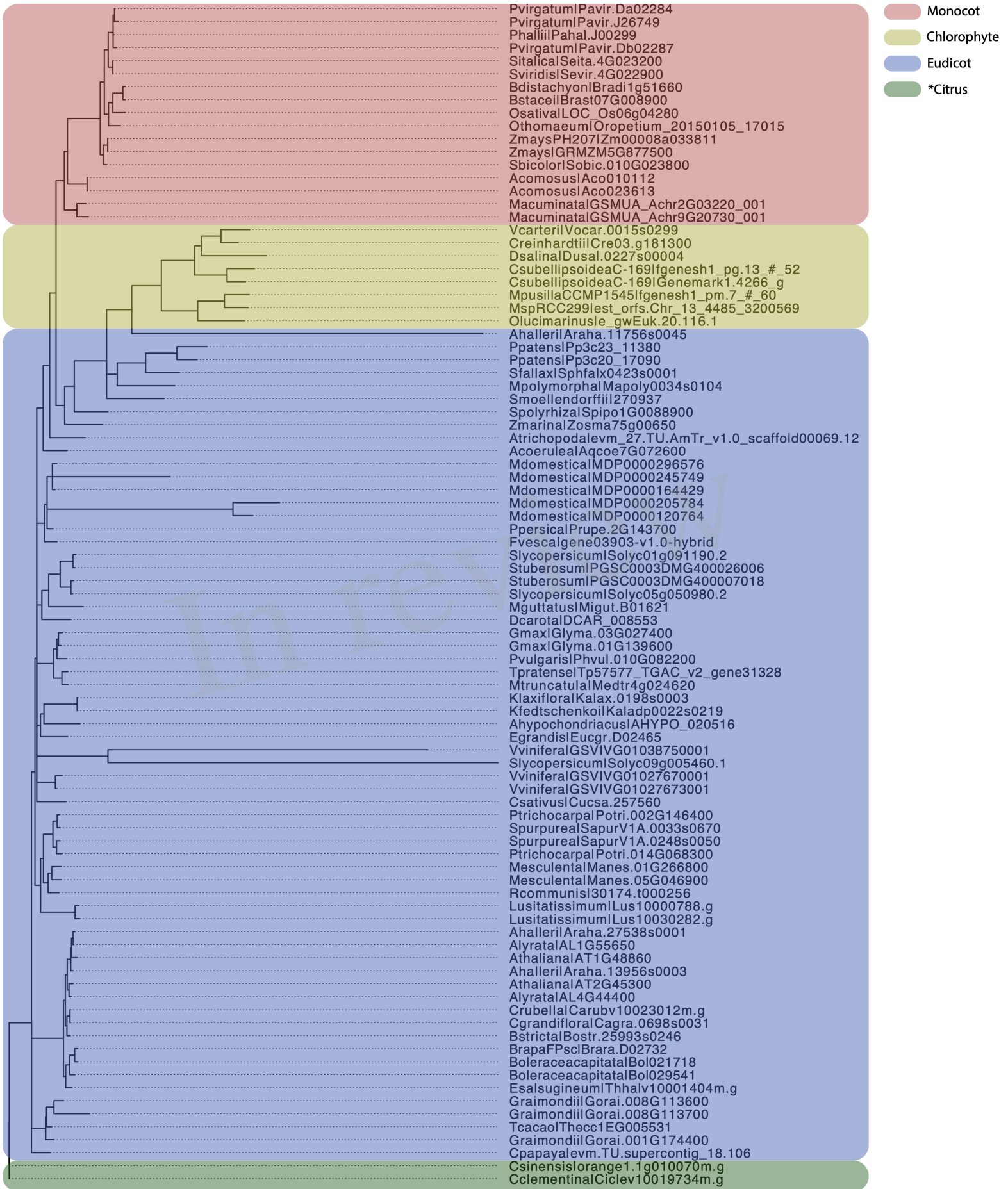


Figure 3.TIFF



0.4