Genome-wide analysis of the lignin toolbox for *morus* and the roles of lignin related genes in response to zinc stress

Nan Chao [1, 2], **Ting Yu** [1], **Chong Hou** [1], **Li Liu** [1, 2], **Lin Zhang** [Corresp. 1, 2]

[1] Jiangsu Key Laboratory of Sericultural Biology and Biotechnology, School of Biotechnology, Jiangsu University of Science & Technology, Zhenjiang, Jiangsu Province, China

[2] Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Zhenjiang, Jiangsu Province, China

Corresponding Author: Lin Zhang
Email address: moruszhanglin@126.com

Mulberry (*Morus,* Moraceae) is an important economic plant with nutritional, medicinal, and ecological values. Lignin in mulberry can affect the quality of forage and the saccharification efficiency of mulberry twigs. The availability of the *Morus notabilis* genome makes it possible to perform a systematic analysis of the genes encoding the 11 protein families specific to the lignin branch of the phenylpropanoid pathway, providing the core genes for the *lignin toolbox* in mulberry. We performed genome-wide screening, which was combined with *de novo* transcriptome data for *Morus notabilis* and *Morus alba* variety *Fengchi,* to identify putative members of the lignin gene families followed by phylogenetic and expression profile analyses. We focused on *bona fide* clade genes and their response to zinc stress were further distinguished based on expression profiles using RNA-seq and RT-qPCR. We finally identified 31 *bona fide* genes in *Morus notabilis* and 25 *bona fide* genes in *Fengchi.* The putative function of these *bona fide* genes was proposed, and a lignin toolbox that comprised 19 genes in *mulberry* was provided, which will be convenient for researchers to explore and modify the monolignol biosynthesis pathway in mulberry. We also observed changes in the expression of some of these lignin biosynthetic genes in response to stress caused by excess zinc in *Fengchi* and proposed that the enhanced lignin biosynthesis in lignified organs and inhibition of lignin biosynthesis in leaf is an important response to zinc stress in mulberry.

Genome-wide analysis of the lignin toolbox for *morus* and the roles of lignin related genes in response to zinc stress

Nan Chao [1,2], Ting Yu [1], Chong Hou[1], Li Liu [1,2], Lin Zhang [1,2†],

1. Jiangsu Key Laboratory of Sericultural Biology and Biotechnology, School of Biotechnology, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212018, China
2. Key Laboratory of Silkworm and Mulberry Genetic Improvement, Ministry of Agriculture and Rural Affairs, Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Zhenjiang, Jiangsu 212018, China

† To whom correspondence should be addressed: Tel (Fax): +86 511 8561 6638

E-mail: Zhanglinsri@126.com, Jiangsu University of Science and Technology.

The English in this document has been checked by at least two professional editors, both native speakers of English. Letpub has provided the Linguistical service for this MS.

Abstract: Mulberry (*Morus,* Moraceae) is an important economic plant with nutritional, medicinal, and ecological values. Lignin in mulberry can affect the quality of forage and the saccharification efficiency of mulberry twigs. The availability of the *Morus notabilis* genome makes it possible to perform a systematic analysis of the genes encoding the 11 protein families specific to the lignin branch of the phenylpropanoid pathway, providing the core genes for the *lignin toolbox* in mulberry. We performed genome-wide screening, which was combined with *de novo* transcriptome data for *Morus notabilis* and *Morus alba* variety *Fengchi,* to identify putative members of the lignin gene families followed by phylogenetic and expression profile analyses. We focused on *bona fide* clade genes and their response to zinc stress were further distinguished based on expression profiles using RNA-seq and RT-qPCR. We finally identified 31 *bona fide* genes in *Morus notabilis* and 25 *bona fide* genes in *Fengchi.* The putative function of these *bona fide* genes was proposed, and a lignin toolbox that comprised 19 genes in *mulberry* was provided, which will be convenient for researchers to explore and modify the monolignol biosynthesis pathway in mulberry. We also observed changes in the expression of some of these lignin biosynthetic genes in response to stress caused by excess zinc in *Fengchi* and proposed that the enhanced lignin biosynthesis in lignified organs and inhibition of lignin biosynthesis in leaf is an important response to zinc stress in mulberry.

Key words: gene family, genome-wide, lignin, mulberry, zinc stress

Introduction

Lignin is an important component of plant cell walls and has important functions in plant growth and stress resistance (Chun, et al. 2019). In turn, owing to its recalcitrant nature and complexity, lignin limits the efficient conversion of lignocellulosic biomass to ethanol (Ragauskas, et al. 2014). The modification of trees with less lignin or with more-degradable lignin along with normal growth, which can improve the quality of forage and saccharification efficiency, has become a hot topic (Dixon, et al. 2014; Umezawa 2018).

The lignin biosynthesis pathway has been deciphered and revised since its discovery decades ago (Whetten and Sederoff 1995). As of now, a total of 11 enzymes have been identified to play a role in monolignol biosynthesis (Zhao 2016). The monolignol biosynthesis pathway generally refers to the branch of phenylpropanoid pathway starting with the deamination of phenylalanine and leading to the production of hydroxycinnamyl alcohols. The general phenylpropanoid pathway contains phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H) and 4-coumarate: CoA ligase (4CL) and provides hydroxycinnamoyl-CoA esters as precursors for a wide range of end products, including lignin, flavonoids, anthocyanins and condensed tannins. In the monolignol-specific biosynthesis pathway, hydroxycinnamoyl-CoA esters undergo successive hydroxylation and O-methylation of their aromatic rings, as well as redox reactions, to produce the monolignols (Zhao 2016). Coumaroyl shikimate 3'-hydroxylase (C3'H) and ferulate 5-hydroxylase (F5H) are responsible for the hydroxylation process. Shikimate *O*-hydroxycinnamoyl -transferase (HCT), caffeoyl CoA 3-*O*-methyltransferase (CCoAOMT) and caffeate/5-

hydroxyferulate *O*-methyltransferase (COMT) are involved in the O-methylation process. Caffeoyl shikimate esterase (CSE) was recently discovered to convert caffeoyl shikimate into caffeate and consists of a bypass with 4CL (Saleme, et al. 2017; Vanholme, et al. 2013). Redox reactions are catalyzed successively by cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) to achieve the conversion of the side-chain carboxyl to an alcohol group. CCR and CAD constitute the primary pathway for monolignol biosynthesis (Zhao 2016).

Mulberry (*Morus,* Moraceae) is an important economic plant in Asia with considerable nutritional and medicinal values (Yuan and Zhao 2017). *Moraceae* is one of the closest relatives of *Rosaceae* and mulberry diverged from *Cannabis sativa* (Cannabaceae) 63.5 Mya, from apple/strawberry (Rosaceae) 88.2 Mya and from *Medicago truncatula (*Fabales) 101.6 Mya (He, et al. 2013; Jiao, et al. 2020). Many studies have shown the great potential of this plant in the energy, food and pharmaceutical industries. Mulberry has long been cultivated for sericulture, which shaped the world's history through the Silk-Road. Furthermore, a large number of by-products of branches twigs have been produced from the large-scale cultivation of mulberry trees in traditional sericulture, and mulberry has been gradually considered a potentially new energy plant providing biomass for the production of biofuels (Łochyńska 2015; Tang, et al. 2012). Studies of lignin biosynthesis have been widely reported for energy plants and forage plants, such as poplar, *Medicago sativa* L. and *Eucalyptus grandis* (Carocha, et al. 2015; Hamberger, et al. 2007; Lee, et al. 2011; Shi, et al. 2010). Recently, Wang et al. characterized four *Ma4CL* genes

from *M. atropurpurea cv. Jialing No. 40*. and revealed the functional divergence of Ma4CL (Wang, et al. 2016).

The availability of the *Morus notabilis* genome and an increasing number of transcriptomic data for mulberry allows comprehensive genome-wide analyses of lignin biosynthesis genes in this species (Li, et al. 2014). In addition, a recent study has been reported to reveal the chromosome-level genome of *Morus alba* (Jiao, et al. 2020). Genome-wide screening, combined with *de novo* transcriptome data, was performed in this study on *Morus notabilis and Morus alba* variety *Fengchi* to obtain the genes putatively included in the 11 monolignol gene families. *M. alba* is one of the most widely cultivated mulberry in China. *M. alba variety Fengchi* is a new variety created by Sericultural Research Institute, Chinese Academy of Agricultural Sciences, expected to spread and grow in extreme environment conditions and used as heavy metal hyperaccumulators and forage. A phylogenetic tree and expression profile were used to further identify the *bona fide* genes involved in lignin biosynthesis, and finally, we provided a lignin toolbox consisting of 19 genes in *Morus notabilis* and 17 genes in *Morus alba* variety *Fengchi,* which will be convenient for researchers to explore and modify the monolignol biosynthesis pathway in this genus. We also assessed the potential roles of lignin biosynthetic genes in response to stress caused by the excess of zinc in *Fengchi* and proposed that the promotion of lignification in lignifying organs, associated with the inhibition of lignin deposition in leaves, is an important response to zinc stress in mulberry.

Materials and Methods

94 **Plant Materials**

95 The materials used in this study were obtained from the National Germplasm Resource Nursery
96 of the Institute of Sericulture, Chinese Academy of Agricultural Sciences. Annual seedlings,
97 *Morus alba* L. variety *Fengchi,* were transplanted into plastic flowerpots, and the potted plants
98 were irrigated with 400 ml/kg of Murashige and Skoog (MS) medium to provide nutrients
99 (Susheelamma, et al. 1996). Zinc sulfate powder was applied near the roots of the mulberry trees
100 as excess zinc stress treatment (450 mg/kg). Changes of proline and superoxide dismutase (SOD)
101 concentration were determined on the 15th day (Figure S1). The root, stem and leaf tissues were
102 quickly frozen in liquid nitrogen and stored at -80°C. This experiment was performed using three
103 biological replicates. These collected samples were used for both RNA-seq and RT-qPCR
104 (quantitative real-time PCR) analysis.

105 **Genome-wide screening of candidate genes for the lignin toolbox in mulberry**

106 *Bona fide* genes involved in lignin biosynthesis with functional characterization from different
107 plants were collected as a query sequence for an HMMer search using MorusDB online
108 ((https://morus.swu.edu.cn /morusdb) (Li, et al. 2014). The sequence identity (>45%), e-value (<e-
109 100) and full score (>400) were used to screen for candidate gene family members. For some gene
110 family members such as CSE and CCoAOMT, the thresholds were flexible to obtain as many as
111 possible candidate genes. A local blastp search was also performed to identify the *Selaginella*
112 *moellendorffii* SmF5H and *Fengchi* homologs (Christiam, et al. 2009). All of the query sequences
113 used and candidate genes obtained are available in **Table S1.**

114 **De novo transcriptome assembly of *Morus alba* variety *Fengchi***

Transcriptome *de novo* assembly was carried out with the short reads assembling program Trinity (Grabherr, et al. 2011). Unigenes were aligned by BLASTx ($e < 0.00001$) to protein databases in nr, Swiss-Prot, KEGG and COG/KOG. The best alignment results were chosen to determine the sequence direction of unigenes. When a unigene could not be aligned to any of these protein databases, the protein coding sequence and sequence direction were confirmed using ESTscan (Iseli, et al. 1999). The data set is available with accession number PRJNA660559 in National Center for Biotechnology Information (NCBI)

Sequence alignment and phylogenetic analysis

Putative protein sequences of different plant species were used for alignment and phylogenetic analysis. Sequences used for phylogenetic analysis were screened from various sources based on the platform PLAZA 3.0 (http://bioinformatics.psb.ugent.be /plaza/). Sequences from the gymnosperm *Picea sitchensis* and the fern *Selaginella moellendorffii* were obtained using BlastP in the NCBI database. Bona fide lignin related genes in different plants were obtained based on published studies (Carocha, et al. 2015; Raes, et al. 2003). Alignment was performed using DNAman 8.0 (Lynnon BioSoft) with default parameters. Phylogenetic trees were constructed using Mega 7.0 with the maximum-likelihood method (Kumar, et al. 2016). The phylogenetic tree was assessed by bootstrapping using 1000 bootstrap replicates and marked above nodes only if greater than 50. The JTT substitution model and G+I rates among sites model were selected as parameters for building the tree. The putative protein sequences used are listed in **Table S2.**

Expression profile analysis

The gene expression based on the large-scale transcriptome data was calculated and normalized to

RPKM (reads per kb per million reads). Transcriptome data of different tissues and organs (root, bark, leaves, winter bud, male-flower) in *Morus notabilis* was obtained from Mrousdb (https://morus.swu.edu.cn /morusdb) (Li, et al. 2014). RNA-seq data for *Fengchi* different organs (root, stem and leaf) was aligned to de-novo transcriptome assembly of *Morus alba variety Fengchi* using bowtie2 and RPKM values for unigenes were calculated using deptools v2.0 based on the bam files (Langdon 2015; Ramirez, et al. 2014). RT-qPCR (quantitative real-time PCR) was also performed to validate gene expressions of 23 *bona fide* clade genes in different organs and the change of their expression levels after zinc treatment using ABI StepOnePlus™ Real-Time PCR System (USA). Genes that showed preferential expression in lignifying tissue or organs (bark, root and stem) were considered as candidate lignin-related genes. Primers based on the coding sequences of these genes were designed using Primer-Blast. The primers are available in **Table S3** and the melt curve of each gene is provided in Figure S2. *Actin* was used as reference gene according to previous studies (Shukla, et al. 2019). Tbtools was used to visualize the expression profile (Chen, et al. 2020), and Graphpad Prism8.0 was used to visualize the RT-qPCR results. SPSS19.0 was used to perform T-test and ANOVA, $p<0.05$ was marked as significant. Three biological replicates were considered for transcriptome data and two biological replicates with three technical replicates respectively were performed for RT-qPCR.

Results

Genome-wide screening of monolignol biosynthesis pathway-related genes

Morusdb (https://morus.swu.edu.cn/morusdb), which provides the genome and transcriptome information for *Morus notabilis,* was used to perform genome-wide screening of candidate genes

involved in monolignol biosynthesis. Finally, we obtained 56 candidate genes based on the HMMer search and blastp results (Table S1). In addition, we identified their homologs in *Fengchi*, a *Morus alba* variety bred by our institute, based on our *de novo* transcriptome data. Most (49/56) of the corresponding homologs were identified in *Fengchi* using candidate genes in *Morus notabilis* as a reference sequence.

Phenylalanine Ammonia-Lyase (PAL)

PAL (EC: 4.3.1.5) catalyzes the deamination of phenylalanine to produce cinnamic acid and is the initial step in the general phenylpropanoid pathway. We constructed a phylogenetic tree (Figure 1A) using both (*Morus notabilis* Mn) *MnPAL* and (*Fengchi* Fc) *FcPAL* and *bona fide* PAL data reported in other species. Seven *MnPAL*s were identified and clustered as *bona fide PAL*s. However, only three homologs, *FcPAL3*, *FcPAL6* and *FcPAL7,* were found in *Fengchi* based on *de novo* transcriptome data. MnPAL7 and FcPAL7 were quite divergent compared with other PALs in angiosperms and are closer to PALs from gymnosperms, which is similar to EgrPAL2 in *Eucalyptus grandis* (Carocha, et al. 2015). Both *MnPAL7* and *FcPAL7* showed a low expression level in various tissues and organs and exhibited no obvious preference in the lignified tissues and organs (Figure 1B). MnPAL1, 2, 4, and 5 are phylogenetically close to AtPAL1 and AtPAL2, which have been reported to be mainly involved in anthocyanin production (Cochrane, et al. 2004; Huang, et al. 2010). *MnPAL1* and *5* were preferentially expressed in lignifying organs and tissues (root and bark), which differed from the expression patterns of *MnPAL2* and *4* (Figure 1B). *MnPAL1, 2, 4, and 5* (L484_024371, L484_024373, L484_024372, L484_024369) have high sequence identity (Aligned protein sequence identity> 99%) and are located close to each other in

the genome, forming a gene cluster. Although *MnPAL1, 2, 4,* and *5* showed high expression levels in the studied organs and tissues in *Morus notabilis*, we could not find or distinguish homologs of *MnPAL1, 2, 4,* and *5* in *Fengchi*. Mn PAL3, 6 and FcPAL3, 6 are phylogenetically close to AtPAL4, 5 and PtrPAL4, 5, which are reported to express more specifically in xylem tissues (Raes, et al. 2003). *MnPAL3* and *FcPAL3* also showed an expression preference in the root, stem or bark, with a high overall expression level, while *MnPAL6* and *FcPAL6* showed low overall expression levels in all of the examined organs and tissues (Figure 1B). RT-qPCR results also validated the expression preference of *FcPAL3* in stems (Figure S3). Based on the above facts, *MnPAL1, 3,5* and *FcPAL3* are the PAL genes most likely to be involved in lignification.

4-Coumaric acid coenzyme A ligase (4CL)

4CL (EC: 6.2.1.12) belongs to the ANL (AMP-producing adenylating superfamily of enzymes) superfamily and catalyzes the formation of CoA thiol esters of 4-coumarate and other 4-hydroxycinnamates, which are important input metabolites, especially for lignin biosynthesis and flavonoid biosynthesis (Ehlting, et al. 1999). The *bona fide* 4CL clade in angiosperm comprises three classes. Clade I contains 4CLs, which are mainly involved in lignin biosynthesis, including At4CL1, 3, 4, Pto4CL1, 3, 4, 5, Mn4CL1, 2, 4 and Fc4CL1, 2, 4 (Figure 1C). The expression profiles of *Mn4CL1*, *2*, *4* and *Fc4CL1*, *2*, *4* also indicated preferential expression in the root, stem or bark (Figure 1D and Figure S3). Mn4CL3 and Fc4CL3 were clustered in Clade II together with At4CL3, Pto4CL2 and Os4CL2, which have been reported to be associated with flavonoid and soluble phenolic biosynthesis (Gui, et al. 2011; Li, et al. 2015; Rao, et al. 2015). Mn4CL3 showed a high expression level in male flowers but a low expression level in winter buds, consistent with

its possible function in flavonoid biosynthesis. The third clade only contained Os4CLs (Os4CL1/3/4/5), which are thought to be distinct from the lignin-associated clade I 4CLs found in dicots. We also found that Mn4CL5, 6 and Fc4CL5, 6 were in a separate cluster and were phylogenetically close to (*Plagiochasma appendiculatum*) Pa4CL, the liverwort *Plagiochasma appendiculatum*. *Mn4CL5* showed a similar expression pattern to *Mn4CL3* and a high expression in male flowers and bark. *Mn4CL6* had an expression specific to male flowers. These facts indicate that *Mn4CL5*, *6* may also be involved in flavonoid and soluble phenolic biosynthesis, given the high flavonoid content in mulberry. Mn4CL5, 6 and Fc4CL5, 6 are divergent from 4CLs in angiosperms and still need to be further studied to identify their roles in mulberry. Therefore, *Mn4CL1, 2, 3, 4* and *Fc4CL1, 2, 3*, 4 are the 4CL genes most likely to be involved in lignification.

Hydroxylation steps in the general phenylpropanoid pathway

C4H (EC: 1.14.13.11) and C3'H (EC: 1.14.14.1) catalyze the hydroxylation steps. C4H and C3'H belong to the CYP73 and CYP98 families, respectively which are members of the cytochrome P450 monooxygenase superfamily. C4H is generally encoded by small gene family, except in *Arabidopsis*, which has only one *AtC4H*. Studies in *Populus* have shown a C4H-C3'H complex that more efficiently catalyzes hydroxylation steps (Chen, et al. 2011). Here, we identified three candidate *C4H*s in mulberry. MnC4H1, 2 and FcC4H1, 2 clustered with AtC4H and PoptrC4H1, 2 as Clade I, which is responsible for lignin biosynthesis (Figure 2A). *MnC4H1, 2* showed a high expression in all organs and tissues. *FcC4H1, 2* was preferentially expressed in lignified organs (Figure 2C, Figure S3). MnC4H3 and FcC4H3 were grouped with PoptrC4H3 and are distinct from MnC4H1, 2 and FcC4H1, 2. Similar to PoptrC4H3, *MnC4H3* and *FcC4H3* had a very low

220 expression level in all of the studied organs. Therefore, *MnC4H1, 2* and *FcC4H1, 2* are the *C4H*
221 genes most likely to be involved in lignification.

222 Although C3'H was shown to catalyze the conversion of *p*-coumaric acid into caffeic acid *in*
223 *vitro*, further studies demonstrated that its activity *in vitro* is the conversion of *p*-coumaroyl
224 shikimate to caffeoyl shikimate (Abdulrazzak, et al. 2006; Franke, et al. 2002, Schoch et al. 2001).
225 Based on our phylogenetic analysis, only MnC3'H1 and FcC3'H1 clustered with StC3'Hs as *bona*
226 *fide* clade II (Figure 2B). It is interesting to note that MnC3'H1 and FcC3H1 are phylogenetically
227 closer to C3'Hs in monocots, other than C3'Hs such as AtC3'H and PoptrC3'H in dicots (Figure
228 2B). Other candidates, including MnC3'H2-5 or FcC3'H2-5, is in a separate cluster without any
229 C3'H orthologs in other plants. It seems that C3'H in mulberry is similar to that in *A. thaliana*,
230 which also has only one C3'H (Raes et al., 2003). *MnC3'H1* showed a high expression in all of the
231 studied organs, and *FcC3'H1* was preferentially expressed in the stem, according to both
232 transcriptome data and RT-qPCR, which likely involves lignification (Figure 2C and Figure S3).
233 Therefore, *MnC3'H1* and *FcC'3H1* are the *C3H* genes most likely to be involved in lignification.

234 **Hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase (HCT) and caffeoyl**
235 **shikimate esterase (CSE)**

236 HCT (EC: 2.3.1.133) combined with C3'H (*p*-coumarate 3-hydroxylase) catalyzes two steps
237 to change the carbon flux from H to G and S lignin units. HCT belongs to the BAHD
238 acyltransferase family and is able to utilize a variety of non-native substrates (Chiang, et al. 2018;
239 D'Auria 2006). *P*-coumaroyl-CoA and caffeoyl-CoA are preferential substrates for HCTs, and
240 HCTs catalyze the acylation of CoA esters with shikimate, producing shikimate esters containing

241 coumaric acid or caffeic acid. The reverse reaction for the formation of caffeoyl-CoA from
242 caffeoyl shikimate is also catalyzed by HCT. Hydroxy-cinnamoyl CoA: quinate
243 hydroxycinnamoyl transferase (HQT) is another acyl transferase that uses quinic acid instead of
244 shikimic acid as the acceptor compound and is involved in chlorogenic acid biosynthesis, not lignin
245 biosynthesis (Niggeweg, et al. 2004). We constructed a phylogenetic tree using both HCTs and
246 HQTs to distinguish *bona fide* HCT clades (Figure 3A). Six candidate MnHCTs were grouped as
247 *bona fide* HCTs with HCTs in angiosperms. MnHCT2, 3 (L484_000457, L484_018078) and
248 MnHCT5, 6 (L484_017530, L484_017529) had high sequence similarity (aligned protein sequence
249 identity>95%). We could not distinguish *FcHCT2*, *3* and *FcHCT5*, *6* based only on transcripts;
250 therefore, we named *FcHCT2* and 5 based on their similar expression pattern to *MnHCT2* and
251 *MnHCT5*. Among all *MnHCTs*, *MnHCT1*, *2* and *FcHCT1, 2* were preferentially expressed in
252 lignified organs and tissues (stems, roots and bark) and are likely involved in monolignol
253 biosynthesis (Figure 3C. Figure S3). Other *MnHCT*s and *FcHCT*s showed relatively low
254 expression in all organs and tissues. MnHCT5, 6 and FcHCT5 are phylogenetically divergent from
255 other *MnHCTs* and *MnHCTs* and showed preferential expression in the leaf (*FcHCT5*) and in the
256 winter bud and leaf (*MnHCT5*), which indicates their possible different roles as opposed to
257 lignification in mulberry. Therefore, *MnHCT1, 2* and *FcHCT1, 2* are the *HCT* genes most likely
258 to be involved in lignification.

259 AtCSE was first characterized as lysoPL2, a member of the monoacylglycerol lipase (MAGL)
260 gene family in *Arabidopsis* (Kim et al. 2016; Gao et al. 2010). AtCSE was first reported as caffeoyl
261 shikimate esterase by Vanholme et al. (2013) in *Arabidopsis* because of its ability to convert

caffeoyl shikimate into caffeate. Further analysis of an *A. thaliana cse-2* (caffeoyl shikimate esterase 2) knockout mutant that presented a reduced lignin content enriched in H units and depleted in S units indicated the involvement of CSE (EC: 3.1.1.) in lignin biosynthesis (Vanholme, et al. 2013). CSE competes with HCT for the substrate caffeoyl shikimate. MnCSE1 and FcCSE1 are phylogenetically close to AtCSE, PoptrCSE1,2 and MtCSE, which have been reported to be involved in lignin biosynthesis (Figure 3B) (Ha, et al. 2016; Saleme, et al. 2017). In addition, *FcCSE1* showed preferential expression in lignified organs and tissues (Figure 3C, Figure S3). MnCSE2 and FcCSE2 showed close relationship with AtGAML1 which was reported to harbor MAG lipase activities and lysophosphatidylcholine (LPC) and/or lysophosphatidy - lethanolamine (LPE) hydrolase activities. MnCSE3, without a homolog in *Fengchi*, was far from the *bona fide* CSEs and cluster with AtMAGL9 and 12. MnCSE3 had an expression preference in winter buds and male flowers. In general, *MnCSE1* and *FcCSE1* are lignin-related CSEs, but *MnCSE2* and *FcCSE2* are monoacylglycerol lipase.

The methylation steps

COMT (EC: 2.1.1.68) and CCoAOMT (EC: 2.1.1.104) are both involved in the methylation steps of the monolignol pathway (Zhong, et al. 1998). CCoAOMT catalyzes the methylation of caffeoyl CoA to produce feruloyl CoA and is reported to be responsible for G and S-type lignin. Only one CCoAOMT in mulberry, MnCCoAOMT1 or FcCCoAOMT1, clusters in the *bona fide* clade with AtCCoAOMT and PtoCCoAOMT1 and 2 (Figure 4A). Both *MnCCoAOMT1* and *FcCCoAOMT1* showed a high expression level in the lignified organs (Figure 4C, Figure S3). *FcCCoAOMT1* had the highest expression level in the stems, about 50-fold higher than that in the leaves (Figure 4C).

283 *MnCCoAOMT1* had high expression in the root, bark and male flowers, with the highest expression
284 in the root (two-fold higher than the expression in the bark or male flower, five-fold higher than
285 the expression in the leaf). Two other candidates, *MnCCoAOMT2*, *3* and *FcCCoAOMT2*, *3* belong
286 to the CCoAOMT-like clade and had a low expression level in all organs and tissues. Therefore,
287 *MnCCoAOMT1* and *FcCCoAOMT1* are the *CCoAOMT1* genes most likely to be involved in
288 lignification.

289 In angiosperms, COMT (EC: 2.1.1.68) is involved in the synthesis S precursors and is now
290 considered to be primarily involved in the synthesis of S units through the preferential methylation
291 of 5-hydroxyconiferyl aldehyde into sinapaldehyde based on functional analysis in several species
292 (Davin, et al. 2008) In mulberry, only one COMT, MnCOMT4 or FcCOMT4 was identified as a
293 *bona fide* COMT together with AtCOMT and PoptrCOMT1,2, based on our phylogenetic analysis
294 (Figure 4B). *MnCOMT4* and *FcCOMT4* showed obvious expression preference in the lignified
295 organs and tissues (Figure 4C, Figure S3) and should be responsible for lignin biosynthesis in
296 mulberry. Other candidate MnCOMTs and FcCOMTs were in a separate cluster and
297 phylogenetically far from the *bona fide* clade. *MnCOMT2*, *5* showed a relatively high expression
298 in male flowers compared with that in the leaf, winter bud and bark. *MnCOMT1* and *6* showed a
299 very low expression in all of the detected organs, and the RPKM of *MnCOMT6* based on the
300 published transcriptome data is not available. *FcCOMT1*, *2*, *3*, *5,* and *6* had similar expression
301 patterns, with overall low expression in all organs and a relatively high expression in roots
302 compared with the stem and leaf. These COMT-like genes need more evidence and functional
303 analysis for elucidating their roles in mulberry. Therefore, *MnCOMT4* and *FcCOMT4* are the

COMT genes most likely to be involved in lignification.

Hydroxylation step specific for S lignin production

F5H (EC:1.14.13) belongs to the CYP84 family and is similar to C3H and C4H as a member of the cytochrome P450 monooxygenases. F5H (EC: 1.14.13), also called CAld5H because of its substrate preference for coniferaldehyde/coniferyl alcohol (Humphreys, et al. 1999), catalyzes the hydroxylation step specific for the production of sinapyl alcohol and, ultimately, S lignin. The discovery of SmF5H in the lycophyte *Selaginella moellendorffii* revealed a novel P450 (CYP788A1) (Weng, et al. 2008). SmF5H shares only 37% amino acid sequence identity with its angiosperm counterparts and can also use *p*-coumaraldehyde and *p*-coumaryl alcohol as substrates to efficiently produce caffeoyl aldehyde and caffeoyl alcohol. Therefore, SmF5Hs can divert G-substituted intermediates toward S lignin synthesis through related but distinct pathways compared with angiosperms (Weng and Chapple 2010). In addition to the genome-wide screening using F5Hs from angiosperms, we carried out blastp using SmF5H as a query to find more F5H-like sequences MnF5H1, 2 and FcF5H1, 2 were identified as candidate F5H based on a Hmmer search using F5Hs from angiosperms. MnF5H1 and FcF5H1 clustered with AtF5H and PoptrF5H1, 2 and belong to the *bona fide* clade in angiosperms (Figure 5A). *MnF5H1* and *FcF5H1* showed obvious expression preference in lignified organs and tissues and are likely to be involved in lignin biosynthesis (Figure 5B, Figure S3). In contrast, *MnF5H2* and *FcF5H2* were far from the *bona fide* clade and had very low expressions in all organs and tissues. Other candidate F5Hs named MnF5H3(Sm), MnF5H4(Sm) or FcF5H3(Sm), FcF5H4(Sm) were identified, sharing about 45% protein sequence identity with SmF5H. *MnF5H3(Sm*), *MnF5H4(Sm)* and *FcF5H3(Sm),*

325 *FcF5H4(Sm)* showed relatively high expression in the root, and *MnF5H3(Sm)* also had a high
326 overall expression in the bark, male flowers and leaf. These SmF5H-like proteins in mulberry may
327 be involved in the response to zinc stress since *FcF5H3(Sm)* and *FcF5H4(Sm)* both showed a
328 decreased expression in the leaf after zinc treatment (Figure S4). Therefore, *MnF5H1* and *FcF5H1*
329 are the *F5H* genes most likely to be involved in lignification.

330 **The last two reductive steps**

331 CCR (EC: 1.2.1.44) is the first committed enzyme for a specific branch of monolignol
332 biosynthesis and converts various cinnamoyl-CoA esters (*p*-coumaroyl-CoA, caffeoyl-CoA,
333 feruloyl-CoA and sinapoyl-CoA) to produce their corresponding hydroxycinnamaldehydes, which
334 are further reduced into different monolignols by another reductase called cinnamyl-alcohol
335 dehydrogenase (CAD EC: 1.1.1.195). CCR and CAD are involved in the primary pathway of
336 monolignol biosynthesis. A recent study showed that PoptrCAD1 and PoptrCCR2 can form a
337 complex to regulate monolignol biosynthesis in *Populu*s (Yan, et al. 2019).

338 Six candidate CCRs from the mulberry genome were screened; phylogenetic analysis showed
339 that MnCCR1, 2 and FcCCR1, 2 belonged to the *bona fide* clade with AtCCR1,2, MtCCR1,2 and
340 PtoCCR1,7 (Figure 6A). Further motif-aware analysis based on our previously reported workflow
341 further validated that MnCCR1,2 and FcCCR1,2 belonged to *bona fide* CCRs and other candidate
342 CCRs should be CCR-like (Figure S5) (Chao, et al. 2019). Similar to AtCCR1, 2 and PtoCCR1,
343 7, different *bona fide* CCRs in mulberry also had different expression patterns. *MnCCR1* and
344 *FcCCR1* had a high overall expression, with the highest expression level in the bark or stem, while
345 *MnCCR2* and *FcCCR2* had quite low expression in all organs and tissues (Figure 6C, Figure S3).

Therefore, *MnCCR1* and *FcCCR1* are likely to play a predominant role in monolignol biosynthesis. As both phylogenetic analysis and motif-aware analysis showed, other MnCCRs and FcCCRs belong to the CCR-like cluster with unknown functions.

CAD is the last enzyme in monolignol biosynthesis and uses various phenylpropenyl aldehyde derivatives as substrates to ensure the diversity of lignin. We obtained six candidate MnCADs, and the corresponding homologs in *Fengchi* were found except MnCAD4. FcCAD4 was different from all six candidate MnCADs, with high (73.88%) protein sequence identity to AtCAD1 which was reported to have very low catalytic activity *in vitro* and play roles in lignification of elongating stems (Eudes, et al. 2006; Kim, et al. 2004). MnCAD1, 2, 3, 4, 5 and FcCAD1, 2, 3, 5 belong to *bona fide* clades (Figure 5B). MnCAD3, 4 and FcCAD3 clustered with AtCAD4, 5, PtoCAD1, and BdCAD5, which have been reported to be involved in lignin biosynthesis. Although the expression data for MnCAD3, 4 were not available in Morusdb, *FcCAD3* showed expression specific to lignified organs based on our transcriptome data and RT-qPCR (Figure 5C, Figure S3). In addition to the above *bona fide* CADs, there is another kind of *bona fide* CAD with the present PtrSAD in *Populus* (Li, et al. 2001). This PtrSAD has been reported to prefer sinapaldehyde as a substrate,however, our previous study on PtoCAD2 showed no obvious substrate preference (Chao, et al. 2014). MnCAD1,2 and FcCAD1, 2 are phylogenetically close to the so-called PtrSAD and PtoCAD2 and cluster as another *bona fide* clade. *MnCAD1* showed no obvious expression preference while *FcCAD1* exhibit a preference for lignified organs based on RT-qPCR results (Figure 5C, Figure S3). *MnCAD2* and *FcCAD2* showed an expression preference for winter-bud or leaf (Figure 5C, Figure S3).

Summary of the *lignin toolbox* for mulberry and the response to zinc ion stress

Genes considered the lignin toolbox in mulberry were annotated in TableS4. Hierarchical clustering depicted a similar expression pattern for bona fide lignin biosynthetic genes we described above, which differed from that of genes identified as 'like' genes" (candidate genes excluded from bona fide clade) (Figure S6 A). It was obvious that the *bona fide* genes had a higher overall expression in the studied organs than the 'like' genes. 21 of total 31 *bona fide* genes in mulberry can be classified as two main clusters based on their expression patterns. Cluster I (indicated as a blue star r) includes genes with obvious expression preferences in lignified organs such as stem and bark, and cluster II (indicated as a red star) includes genes with a high expression but no obvious preference in lignified organs and tissues. RT-qPCR also validated the expression preference in lignified organs for the lignin-related genes comprise the lignin toolbox for mulberry (Figure S6B, Figure S3 and Table S4).

Mulberry has been reported as heavy metal hyperaccumulators. Our results showed that lignin-related genes play important roles in responding to zinc stress. All we detected *bona fide* clade genes (22/23) including 16 core genes in lignin toolbox in *Fengchi* show significant expression change in at least one organ (Figure 7A). Only *FcCCoAOMT*1showed no significant change in any detected organs after zinc excess treatment (Figure 7A, Figure S3, Table S4). Monolignol biosynthesis pathway in *Fengchi* showed overall up-regulation in root and stem but down-regulation in leaf (Figure 7B). Most of these bona fide clade genes (13/22) were down-regulated in leaves. Five genes showed up-regulation in both leaf and lignified organs (Figure 7C). It is likely that the promotion of lignin biosynthesis in lignified organs in mulberry is an important

388 way to respond to zinc stress.

389 **Discussion**

390 **Lignification toolbox in mulberry**

391 Genes involving in secondary metabolism in mulberry were reported to have faster
392 evolutionary rate (Jiao, et al. 2020). Lignin biosynthesis is important pathway in land plants.
393 Genome-wide screening of candidate genes involving in monolignol biosynthesis was performed
394 here in mulberry. In total, 31 *bona fide* clade genes were obtained in *Morus notabilis* based on
395 phylogenetic analysis, and 25 *bona fide* homologs were found in *Fengchi,* similar to *Populus* (25)
396 and *E. grandis* (38) (Table S4) (Carocha, et al. 2015; Shi, et al. 2010). The loss of *MnHCT* and
397 *MnPAL* homologs in *Fengchi* resulted in the above-mentioned change in the total numbers. The
398 *bona fide* genes described above had a similar expression pattern and a higher overall expression
399 in the studied organs compared with genes identified as 'like' genes. Combined with the expression
400 profile in different organs of *Morus notabilis* and *Fengchi*, a total of 19 genes were identified as
401 *bona fide* lignification-related genes, which is similar to that in *E. grandi*s (17). These 19 genes
402 were preferentially expressed in the lignified organs and tissues and probably represent the core
403 lignification toolbox in mulberry (Figure S6B, Figure S3, Table S4).

404 **The lignin biosynthesis pathway plays an important role in the response to stress caused by**
405 **excess zinc in mulberry**

406 Zinc is a trace element that is necessary for a healthy immune system and is important for
407 people to maintain their fitness level. Studies have shown that dietary zinc can act as sleep
408 modulator and is necessary for brain development and function (Cherasse and Urade 2017;

Hambidge 2000). Mulberry is a woody plant with resistance to zinc ions, and both leaves and fruits of mulberry are known as sites rich in zinc (Jiang, et al. 2017; Srivastava, et al. 2006). Black mulberry (*Morus nigra*) juice has high amounts of zinc and iron, which could help to improve the micronutrient status of pregnant women and children (Khalid, et al. 2011) A deficiency or excess zinc leads to oxidative stress. Moreover, the Zn-deficiency leads to abnormal development of leaves in mulberry (Kumar Tewari, et al. 2008).

Mulberry is able to uptake the heavy metal and was reported to immigrate 254532.8mg Zn every square meter plough layer soil (Jiang, et al. 2017). The contents of zinc in different mulberry organs (leaf, root, bark and stem) are greatly different (Jiang, et al. 2017). After excess zinc treatment, 23 core genes involved in lignin biosynthesis except *CCoAOMT1* showed obvious expression changes in different organs (Figure 7). Monolignol biosynthesis pathway in *Fengchi* showed overall up-regulation in root and stem but down-regulation in leaf (Figure 7B). Relatively high expression of lignin related genes (total 24 genes) was also reported in response to zinc exposure in roots of *Thlaspi caerulescens*, one of the natural zinc hyperaccumulator species (van de Mortel, et al. 2006). Lignin has been reported to act as a metal-absorbing matrix in response to metal stress (Bhardwaj, et al. 2014). It is likely that the promotion of lignin biosynthesis in lignified organs association with inhibition of lignin biosynthesis in leaf in mulberry is an important response to zinc stress. A similar situation was reported for *Lens Culinaris* and *Phaseolus Mungo* subjected to lead stress (Haider and Azmat 2012).

Supplementary Data

Figure S1 Pro concentration and SOD concentration in mulberry leaves, roots and stems at 15 days after treatment

Figure S2 The melt curve of each gene for RT-qPCR experiments

Figure S3 RT-qPCR results for 23 *bona fide* clade genes in *Fengchi*

Figure S4 RT-qPCR for FcF5H3(Sm) and FcF5H4(Sm) in *Fengchi*

Figure S5 Alignment and motif analysis of CCRs. The motifs were marked in red box.

Figure S6 Expression profile of all candidate genes in mulberry.

Table S1 Genome-wide screening of monolignon biosynthesis pathway-related genes

Table S2 Sequence used for lignin-related gene analysis in mulberry

Table S3 Primers used for RT-qPCR

Table S4 Putative functional classification of 56 candidate lignification-related genes

Table S5 Raw data for RT-qPCR

Table S6 Basic annotation of all unigenes

Acknowledgments

This work was jointly supported by the Crop Germplasm Resources Protection Project of the Ministry of Agriculture and Rural Affairs of the People's Republic of China (19190172), National Infrastructure for Crop Germplasm Resources (NICGR-2019-43), China Agriculture Research System of MOF and MARA and Jiangsu Province Postgraduate Research and Practice Innovation Project (KYCX20_3166).

Conflict of Interests

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

Lin Zhang, Li Liu and Nan Chao guided the work and provided advices; Nan Chao and Ting Yu performed the experiments; Nan Chao analyzed the data, organized the figures and wrote the manuscript. Chong Hou performed proof-reading.

Availability of data and material

The raw sequencing data has been deposited in NCBI: https://www.ncbi.nlm.nih.gov/sra/PRJNA660559. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/ENA/GenBank under the accession GJFU00000000. The version described in this paper is the first version, GJFU01000000.

References

Abdulrazzak N, Pollet B, Ehlting J, Larsen K, Asnaghi C, Ronseau S, Proux C, Erhardt M, Seltzer V, Renou JP, Ullmann P, Pauly M, Lapierre C, Werck-Reichhart D (2006) A coumaroyl-ester-3-hydroxylase insertion mutant reveals the existence of nonredundant meta-hydroxylation pathways and essential roles for phenolic precursors in cell expansion and plant growth. Plant Physiol 140:30-48

Bhardwaj R, Handa N, Sharma R, Kaur H, Kohli S, Kumar V, Kaur P (2014) Lignins and Abiotic Stress: An Overview. In: Physiological Mechanisms and Adaptation Strategies in Plants Under Changing Environment. pp 267-296

Carocha V, Soler M, Hefer C, Cassan-Wang H, Fevereiro P, Myburg AA, Paiva JA, Grima-Pettenati J (2015) Genome-wide analysis of the lignin toolbox of Eucalyptus grandis. New Phytol 206:1297-1313

Chao N, Jiang WT, Wang XC, Jiang XN, Gai Y (2019) Novel motif is capable of determining CCR and CCR-like proteins based on the divergence of CCRs in plants. Tree Physiol 39:2019-2026

Chao N, Liu SX, Liu BM, Li N, Jiang XN, Gai Y (2014) Molecular cloning and functional analysis of nine cinnamyl alcohol dehydrogenase family members in Populus tomentosa. Planta 240:1097-1112

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020) TBtools - an integrative toolkit developed for interactive analyses of big biological data. Mol Plant

Chen HC, Li Q, Shuford CM, Liu J, Muddiman DC, Sederoff RR, Chiang VL (2011) Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. Proc Natl Acad Sci U S A 108:21253-21258

Chiang YC, Levsh O, Lam CK, Weng JK, Wang Y (2018) Structural and dynamic basis of substrate permissiveness in hydroxycinnamoyltransferase (HCT). PLoS Comput Biol 14:e1006511

Christiam, Camacho, George, Coulouris, Vahram, Avagyan, Ning, Ma, Jason, Papadopoulos (2009) BLAST+: architecture and applications. BMC Bioinformatics

Chun HJ, Baek D, Cho HM, Lee SH, Jin BJ, Yun DJ, Hong YS, Kim MC (2019) Lignin biosynthesis genes play critical roles

in the adaptation of Arabidopsis plants to high-salt stress. Plant Signal Behav 14:1625697

Cochrane FC, Davin LB, Lewis NG (2004) The Arabidopsis phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. Phytochemistry 65:1557-1564

Davin LB, Jourdes M, Patten AM, Kim KW, Vassao DG, Lewis NG (2008) Dissection of lignin macromolecular configuration and assembly: comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis. Nat Prod Rep 25:1015-1090

D'Auria JC (2006) Acyltransferases in plants: a good time to be BAHD. Curr Opin Plant Biol 9:331-340

Dixon RA, Reddy MS, Gallego‐Giraldo L (2014) Monolignol Biosynthesis and its Genetic Manipulation: The Good, the Bad, and the Ugly. Recent Advances in Polyphenol Research, Volume 4 1-38

Ehlting J, Büttner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E (1999) Three 4‐coumarate: coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. The Plant Journal 19:9-20

Eudes A, Pollet B, Sibout R, Do CT, Seguin A, Lapierre C, Jouanin L (2006) Evidence for a role of AtCAD 1 in lignification of elongating stems of Arabidopsis thaliana. Planta 225:23-39

Franke R, Hemm MR, Denault JW, Ruegger MO, Humphreys JM, Chapple C (2002) Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of Arabidopsis. The Plant Journal 30:47-59

Gao W, Li HY, Xiao S, Chye ML (2010) Acyl‐CoA‐binding protein 2 binds lysophospholipase 2 and lysoPC to promote tolerance to cadmium‐induced oxidative stress in transgenic Arabidopsis. The Plant Journal 62:989-100

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644-U130

Gui J, Shen J, Li L (2011) Functional characterization of evolutionarily divergent 4-coumarate:coenzyme a ligases in rice. Plant Physiol 157:574-586

Ha CM, Escamilla-Trevino L, Yarce JC, Kim H, Ralph J, Chen F, Dixon RA (2016) An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in Medicago truncatula. Plant J 86:363-375

Haider S, Azmat R (2012) Failure of Survival Strategies in Adaption of Heavy Metal Environment in Lens Culinaris and Phaseolus Mungo. Pak J Bot 44:1959-1964

Hamberger B, Ellis M, Friedmann M, Clarice D, Barbazuk B, Douglas CJ (2007) Genome-wide analyses of phenylpropanoid-related genes in Populus trichocarpa, Arabidopsis thaliana, and Oryza sativa: the Populus lignin toolbox and conservation and diversification of angiosperm gene families. Canadian Journal of Botany 85:1182-1201

Hambidge M (2000) Human Zinc Deficiency. J Nutr 130:1344S

He N, Zhang C, Qi X, Zhao S, Tao Y, Yang G, Lee T, Wang X, Cai Q, Li D (2013) Draft genome sequence of the mulberry tree Morus notabilis. Nature Communications 4:2445-2445

Huang J, Gu M, Lai Z, Fan B, Shi K, Zhou Y-H, Yu J-Q, Chen Z (2010) Functional analysis of the Arabidopsis PAL gene family in plant growth, development, and response to environmental stress. Plant Physiol 153:1526-1538

Humphreys JM, Hemm MR, Chapple C (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci U S A 96:10045-10050

Iseli C, Jongeneel CV, Bucher P (1999) ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. Proc Int Conf Intell Syst Mol Biol 138-148

Jiang Y, Huang R, Yan X, Jia C, Long T (2017) Mulberry for environmental protection.

524 Jiao F, Luo R, Dai X, Liu H, Yu G, Han S, Lu X, Su C, Chen Q, Song Q, Meng C, Li F, Sun H, Zhang R, Hui T, Qian Y, Zhao
525 A, Jiang Y (2020) Chromosome-level reference genome and population genomic analysis provide insight into the
526 evolution and improvement of domesticated mulberry (Morus alba L). Mol Plant
527 Khalid N, Fawad SA, Ahmed I (2011) Antimicrobial activity, phytochemical profile and trace minerals of black
528 mulberry (Morus Nigra L.) fresh juice. Pak J Bot 43:91-96
529 Kim RJ, Kim HJ, Shim D, Suh MC (2016) Molecular and biochemical characterizations of the monoacylglycerol lipase
530 gene family of Arabidopsis thaliana. Plant J 85:758-771
531 Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger
532 Datasets. Molecular Biology & Evolution 33:1870
533 Kumar Tewari R, Kumar P, Nand Sharma P (2008) Morphology and physiology of zinc-stressed mulberry plants. J
534 Plant Nutr Soil Sci 171:286-294
535 Langdon WB (2015) Performance of genetic programming optimised Bowtie2 on genome comparison and analytic
536 testing (GCAT) benchmarks. BioData Min 8:1
537 Lee Y, Chen F, Gallego-Giraldo L, Dixon RA, Voit EO (2011) Integrative analysis of transgenic alfalfa (Medicago sativa
538 L.) suggests new metabolic control mechanisms for monolignol biosynthesis. PLoS Comp Biol 7:
539 Li L, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL (2001) The last step of syringyl monolignol
540 biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. Plant Cell
541 13:1567-1586Li T, Qi X, Zeng Q, Xiang Z, He N (2014) MorusDB: a resource for mulberry genomics and genome
542 biology. Database 2014:
543 Li X, Weng J, Chapple C (2008) Improvement of biomass through lignin modification. Plant J 54:569-581
544 Li Y, Im Kim J, Pysh L, Chapple C (2015) Four Isoforms of Arabidopsis 4-Coumarate: CoA Ligase Have Overlapping yet
545 Distinct Roles in Phenylpropanoid Metabolism. Plant Physiol 169:2409-2421
546 Łochyńska M (2015) Energy and nutritional properties of the white mulberry (Morus alba l.). Journal of Agricultural
547 Science and Technology 5:709-716
548 Niggeweg R, Michael AJ, Martin C (2004) Engineering plants with increased levels of the antioxidant chlorogenic acid.
549 Nat Biotechnol 22:746-754
550 Raes J, Rohde A, Christensen JH, De Peer YV, Boerjan W (2003) Genome-Wide Characterization of the Lignification
551 Toolbox in Arabidopsis. Plant Physiol 133:1051-1071
552 Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M (2014)
553 Lignin Valorization: Improving Lignin Processing in the Biorefinery. Science 344:1246843
554 Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T (2014) deepTools: a flexible platform for exploring deep-
555 sequencing data. Nucleic Acids Res 42:W187-191
556 Rao G, Pan X, Xu F, Zhang Y, Cao S, Jiang X, Lu H (2015) Divergent and Overlapping Function of Five 4-
557 Coumarate/Coenzyme A Ligases from Populus tomentosa. Plant Molecular Biology Reporter 33:841-854
558 Saleme MLS, Cesarino I, Vargas L, Kim H, Vanholme R, Goeminne G, Van Acker R, Fonseca FCA, Pallidis A, Voorend
559 W, Junior JN, Padmakshan D, Van Doorsselaere J, Ralph J, Boerjan W (2017) Silencing CAFFEOYL SHIKIMATE
560 ESTERASE Affects Lignification and Improves Saccharification in Poplar. Plant Physiol 175:1040-1057
561 Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, Ullmann P, and Werck-Reichhart D. 2001. CYP98A3 from
562 Arabidopsis thaliana is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J Biol*
563 *Chem* 276:36566-36574
564 Shi R, Sun YH, Li Q, Heber S, Sederoff R, Chiang VL (2010) Towards a systems approach for lignin biosynthesis in

Populus trichocarpa: transcript abundance and specificity of the monolignol biosynthetic genes. Plant Cell Physiol 51:144-163

Shukla P, Reddy RA, Ponnuvel KM, Rohela GK, Shabnam AA, Ghosh MK, Mishra RK (2019) Selection of suitable reference genes for quantitative real-time PCR gene expression analysis in Mulberry (Morus alba L.) under different abiotic stresses. Mol Biol Rep 46:1809-1817

Susheelamma BN, Shekar KR, Sarkar A, Rao MR, Datta RK (1996) Genotype and hormonal effects on callus formation and regeneration in mulberry. Euphytica 90:25-29

Tang Y, Liu Q, Chen F (2012) Preparation and characterization of activated carbon from waste ramulus mori. Chem Eng J 203:19-24

Umezawa T (2018) Lignin modification in planta for valorization. Phytochem Rev 17:1305-1327

van de Mortel JE, Almar Villanueva L, Schat H, Kwekkeboom J, Coughlan S, Moerland PD, Ver Loren van Themaat E, Koornneef M, Aarts MG (2006) Large expression differences in genes for iron and zinc homeostasis, stress response, and lignin biosynthesis distinguish roots of Arabidopsis thaliana and the related metal hyperaccumulator Thlaspi caerulescens. Plant Physiol 142:1127-1147

Vanholme R, Cesarino I, Rataj K, Xiao Y, Sundin L, Goeminne G, Kim H, Cross J, Morreel K, Araujo P (2013) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis. Science 341:1103-1106

Wang CH, Yu J, Cai YX, Zhu PP, Liu CY, Zhao AC, Lu RH, Li MJ, Xu FX, Yu MD (2016) Characterization and Functional Analysis of 4-Coumarate:CoA Ligase Genes in Mul-berry. PLoS One 11:e0155814

Weng J-K, Li X, Stout J, Chapple C (2008) Independent origins of syringyl lignin in vascular plants. Proceedings of the National Academy of Sciences 105:7887-7892

Weng JK, Chapple C (2010) The origin and evolution of lignin biosynthesis. New Phytol 187:273-285

Yan X, Liu J, Kim H, Liu B, Huang X, Yang Z, Lin YJ, Chen H, Yang C, Wang JP, Muddiman DC, Ralph J, Sederoff RR, Li Q, Chiang VL (2019) CAD1 and CCR2 protein complex formation in monolignol biosynthesis in Populus trichocarpa. New Phytol 222:244-260

Whetten R, Sederoff R (1995) Lignin Biosynthesis. Plant Cell 7:1001-1013Yuan Q, Zhao L (2017) The Mulberry (Morus alba L.) Fruit—A Review of Characteristic Components and Health Benefits. J Agric Food Chem 65:10383-10394

Zabed H, Sahu JN, Boyce AN, Faruq G (2016) Fuel ethanol production from lignocellulosic biomass: An overview on feedstocks and technological approaches. Renewable & Sustainable Energy Reviews 66:751-774

Zhao Q (2016) Lignification: flexibility, biosynthesis and regulation. Trends Plant Sci 21:713-721

Zhong R, Iii WH, Negrel J, Ye ZH (1998) Dual methylation pathways in lignin biosynthesis. Plant Cell 10:2033-2046

Figure Legends

Figure 1. Phylogenetic analysis and expression profile of PAL and 4CL gene family in mulberry.

(A). Phylogenetic analysis of PALs; (B). Expression profiles of PAL gene family in different tissues or organs in *Morus notabilis* and *Fengchi*; (C). Phylogenetic analysis of 4CLs; (D). Expression profiles of 4CL gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating PALs or 4CLs from dicots, blue full circles indicating PALs or 4CLs from monocots, green full circles indicating PALs or 4CLs from gymnosperms and yellow full circles indicating PALs from ferns or moss. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without Zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment. *Bona fide* clades were marked using different color shading

Figure 2. Phylogenetic analysis and expression profile of C3'H and C4H gene families in mulberry.

(A). Phylogenetic analysis of C4Hs; (B). Phylogenetic analysis of C3'Hs. c. Expression profiles of C3'H and C4H gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating proteins from dicots, blue full circles indicating proteins from monocots, green full circles indicating proteins from gymnosperms and yellow full circles indicating proteins from ferns or moss. *Bona fide* clades were marked using different color shadings. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

621

622 **Figure 3. Phylogenetic analysis and expression profile of HCT and CSE gene families in**
623 **mulberry**.

624 (A). Phylogenetic analysis of HCTs; (B). Phylogenetic analysis of CSEs; (C). Expression profiles
625 of HCT and CSE gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red
626 full circles indicating proteins from dicots, red empty circles indicating HQTs, blue full circles
627 indicating proteins from monocots, green full circles indicating proteins from gymnosperms and
628 yellow full circles indicating proteins from ferns or moss. *Bona fide* clades were marked using
629 different color shadings. Putative protein sequences were used for phylogenetic analysis and the
630 sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating
631 *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc
632 treatment.

633

634 **Figure 4. Phylogenetic analysis and expression profile of CCoAOMT and COMT gene**
635 **families in mulberry**.

636 (A). Phylogenetic analysis of CCoAOMTs; (B). Phylogenetic analysis of COMTs; (C). Expression
637 profiles of CCoAOMT and COMT gene family in different tissues or organs in *Morus notabilis*
638 and *Fengchi.* Red full circles indicating proteins from dicots, blue full circles indicating proteins
639 from monocots, green full circles indicating proteins from gymnosperms and yellow full circles
640 indicating proteins from ferns or moss. *Bona fide* clades were marked using different color
641 shadings. Putative protein sequences were used for phylogenetic analysis and the sequences
642 information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.*
643 L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

644

Figure 5. Phylogenetic analysis and expression profile of F5H gene family in mulberry.

(A). Phylogenetic analysis of F5Hs; (B). Expression profiles of F5H gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating F5Hs from dicots, blue full circles indicating F5Hs from monocots, green full circles indicating F5Hs from gymnosperms and yellow full circles indicating F5Hs from ferns or moss. *Bona fide* clades were marked using different color shadings. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

Figure 6. Phylogenetic analysis and expression profile of CCR and CAD gene families in mulberry.

(A). Phylogenetic analysis of CCRs; (B). Phylogenetic analysis of CADs; (C). Expression profiles of CCR and CAD gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating proteins from dicots, blue full circles indicating proteins from monocots, green full circles indicating proteins from gymnosperms and yellow full circles indicating proteins from ferns or moss. *Bona fide* clades were marked using different color shadings. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

667

668 **Figure 7. Expression change of *bona fide* clade genes in response to excess zinc stress in**
669 **mulberry.**

670 (A). Fold change of expression levels of 23 *bona fide* genes in *Fengchi* after excess zinc treatment;
671 (B). Overall change of monolignol pathway in different organs after excess zinc treatment in
672 *Fengchi*; (C). Clustering of 23 *bona fide* clade genes expression pattern in response to Zinc stress.
673 Two biological replicates with three technical replicates respectively were performed for RT-
674 qPCR. P-value was calculate using SPSS 19.0. "*" indicates $0.01<p<0.05$; "**" indicates
675 $0.001<p<0.01$ and "***" indicates $p<0.001$.

676 **Figure S1 Pro concentration and SOD concentration in mulberry leaves, roots and stems at**
677 **15 days after treatment**

678 **Figure S2 The melt curve of each gene for RT-qPCR experiments**

679 **Figure S3 RT-qPCR results for 23 *bona fide* clade genes in *Fengchi***

680 **Figure S4 RT-qPCR for FcF5H3(Sm) and FcF5H4(Sm) in *Fengchi***

681 **Figure S5 Alignment and motif analysis of CCRs. The motifs were marked in red box.**

682 **Figure S6 Expression profile of all candidate genes in mulberry**.
683 (A). Hierarchical clustering of expression profiles of 56 candidate genes based on transcriptome
684 data in mulberry; Blue star indicated the cluster I and red star indicate the cluster II. Red full circles
685 indicated the *bona fide* clade genes (B). Expression profiles of 23 *bona fide* clade genes in *Fengchi.*
686

Figure 1

Figure 1. Phylogenetic analysis and expression profile of PAL and 4CL gene family in mulberry.

(A). Phylogenetic analysis of PALs; (B). Expression profiles of PAL gene family in different tissues or organs in *Morus notabilis* and *Fengchi*; (C). Phylogenetic analysis of 4CLs; (D). Expression profiles of *4CL* gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating PALs or 4CLs from dicots, blue full circles indicating PALs or 4CLs from monocots, green full circles indicating PALs or 4CLs from gymnosperms and yellow full circles indicating PALs from ferns or moss. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2 . Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without Zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment. *Bona fide* clades were marked using different color shading

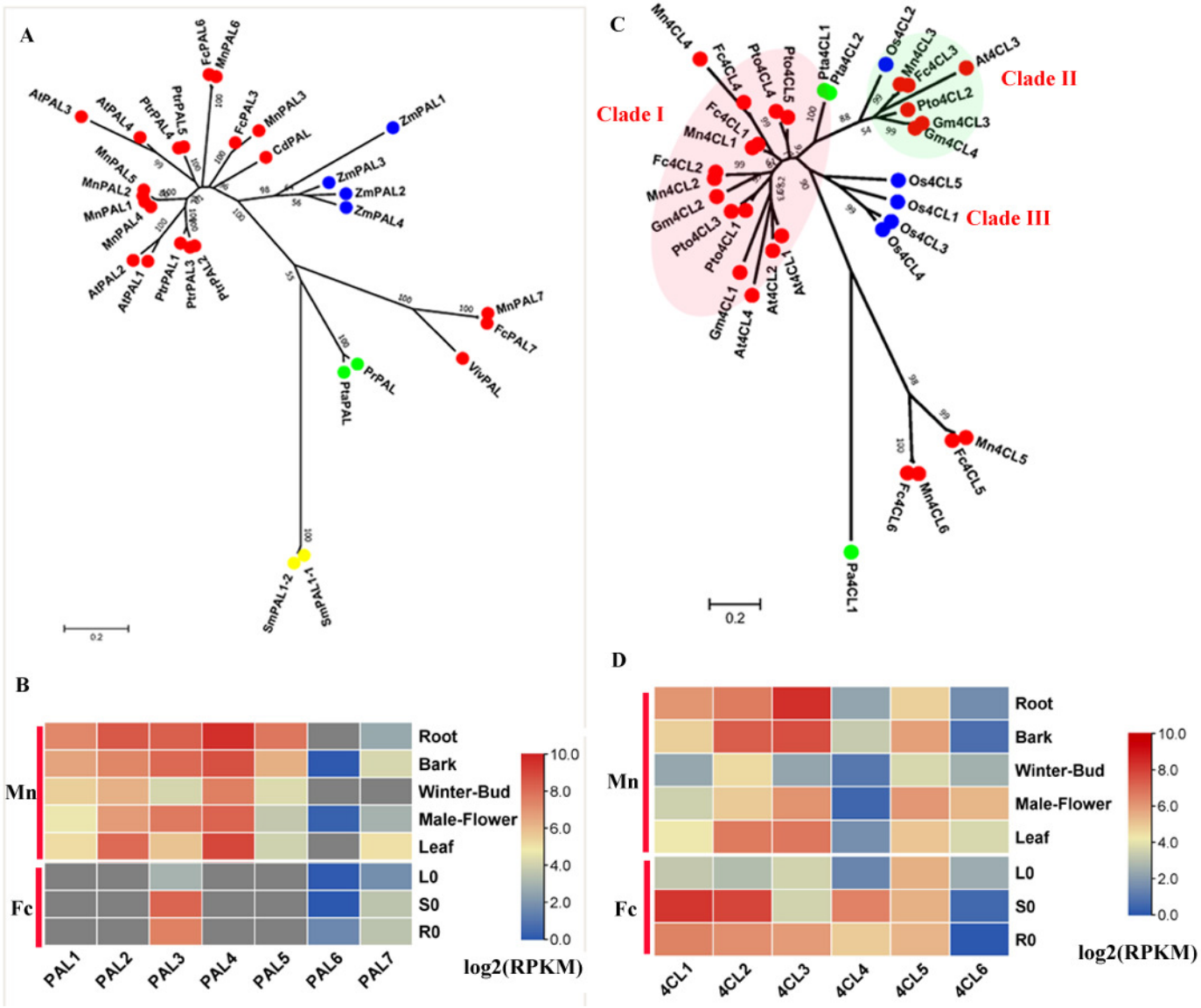

A
FcPAL6
MnPAL6
AtPAL3
AtPAL4
PtrPAL4
PtrPAL5
FcPAL3
MnPAL3
CdPAL
ZmPAL1
MnPAL5
MnPAL2
MnPAL1
MnPAL4
ZmPAL3
ZmPAL2
ZmPAL4
AtPAL2
AtPAL1
PtrPAL1
PtrPAL3
PtrPAL2
MnPAL7
FcPAL7
VivPAL
PtrPAL
PtaPAL
SmPAL1-2
SmPAL1-1
0.2
B
Mn
Fc
Root
Bark
Winter-Bud
Male-Flower
Leaf
L0
S0
R0
PAL1
PAL2
PAL3
PAL4
PAL5
PAL6
PAL7
10.0
8.0
6.0
4.0
2.0
0.0
log2(RPKM)
C
Clade I
Clade II
Clade III
Mn4CL4
Fc4CL4
Pto4CL4
Pto4CL5
Pta4CL1
Pta4CL2
Os4CL2
Mn4CL3
Fc4CL3
At4CL3
Pto4CL2
Gm4CL3
Gm4CL4
Fc4CL1
Mn4CL1
Fc4CL2
Mn4CL2
Gm4CL2
Pto4CL3
Pto4CL1
Os4CL5
Os4CL1
Os4CL3
Os4CL4
Gm4CL1
At4CL4
At4CL2
At4CL1
Mn4CL5
Fc4CL5
Fc4CL6
Mn4CL6
Pa4CL1
0.2
D
Mn
Fc
Root
Bark
Winter-Bud
Male-Flower
Leaf
L0
S0
R0
4CL1
4CL2
4CL3
4CL4
4CL5
4CL6
10.0
8.0
6.0
4.0
2.0
0.0
log2(RPKM)

Figure 2

Figure 2. Phylogenetic analysis and expression profile of C3’H and C4H gene families in mulberry.

(A). Phylogenetic analysis of C4Hs; (B). Phylogenetic analysis of C3’Hs. (C). Expression profiles of *C3’H* and *C4H* gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating proteins from dicots, blue full circles indicating proteins from monocots, green full circles indicating proteins from gymnosperms and yellow full circles indicating proteins from ferns or moss. *Bona fide* clades were marked using different color shadings. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

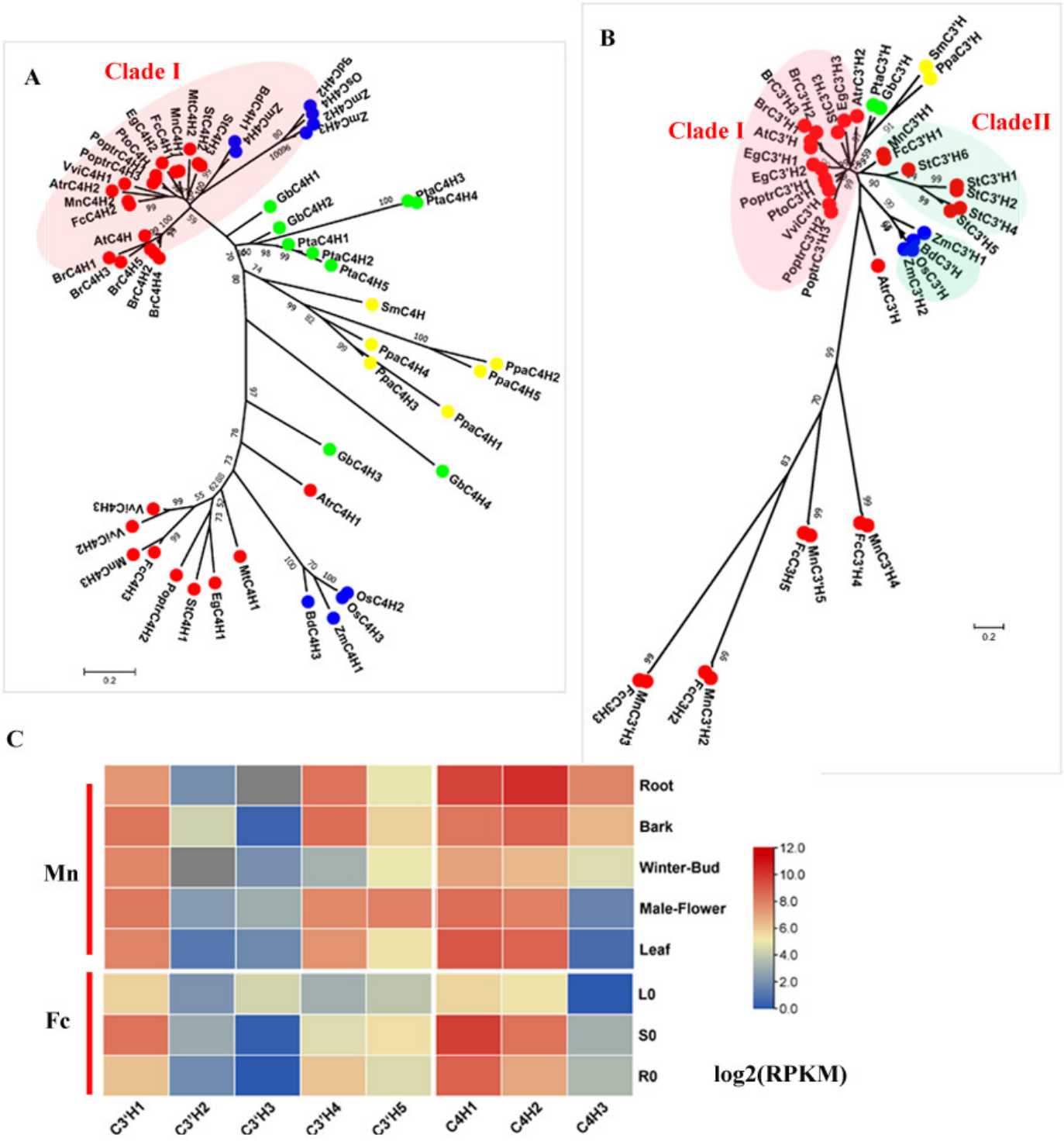

A
Clade I
MtC4H2
StC4H2
MnC4H2
FcC4H2
EgC4H2
PtoC4H2
PoptrC4H1
PoptrC4H3
VviC4H1
AtrC4H2
MnC4H1
FcC4H1
StC4H3
SlC4H1
BdC4H1
ZmC4H4
ZmC4H2
ZmC4H3
OsC4H4
OsC4H2
AtC4H
BrC4H1
BrC4H3
BrC4H5
BrC4H2
BrC4H4
GbC4H1
GbC4H2
PtaC4H1
PtaC4H2
PtaC4H5
PtaC4H3
PtaC4H4
SmC4H
PpaC4H4
PpaC4H3
PpaC4H2
PpaC4H5
PpaC4H1
GbC4H3
GbC4H4
AtrC4H1
VviC4H3
VviC4H2
MnC4H3
FcC4H3
PoptrC4H2
StC4H1
EgC4H1
MtC4H1
BdC4H3
ZmC4H1
OsC4H2
OsC4H3
0.2
B
Clade I
CladeII
BrC3'H3
BrC3'H2
BrC3'H1
AtC3'H
EgC3'H1
EgC3'H2
PoptrC3'H1
PtoC3'H
VviC3'H2
PoptrC3'H2
PoptrC3'H3
SlC3'H3
EgC3'H3
AtrC3'H2
PtaC3'H
GbC3'H
SmC3'H
PpaC3'H
MnC3'H1
FcC3'H1
StC3'H6
StC3'H1
StC3'H2
StC3'H4
StC3'H5
ZmC3'H1
BdC3'H
OsC3'H
ZmC3'H2
AtrC3'H
MnC3'H3
FcC3H2
MnC3'H2
FcC3H5
MnC3'H5
FcC3'H4
MnC3'H4
0.2
C
Mn
Fc
Root
Bark
Winter-Bud
Male-Flower
Leaf
L0
S0
R0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
log2(RPKM)
C3'H1
C3'H2
C3'H3
C3'H4
C3'H5
C4H1
C4H2
C4H3

Figure 3

Figure 3. Phylogenetic analysis and expression profile of HCT and CSE gene families in mulberry.

(A). Phylogenetic analysis of HCTs; (B). Phylogenetic analysis of CSEs; (C). Expression profiles of *HCT* and *CSE* gene family in different tissues or organs in *Morus notabilis* and *Fengchi .* Red full circles indicating proteins from dicots, red empty circles indicating HQTs, blue full circles indicating proteins from monocots, green full circles indicating proteins from gymnosperms and yellow full circles indicating proteins from ferns or moss. *Bona fide* clades were marked using different color shadings. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

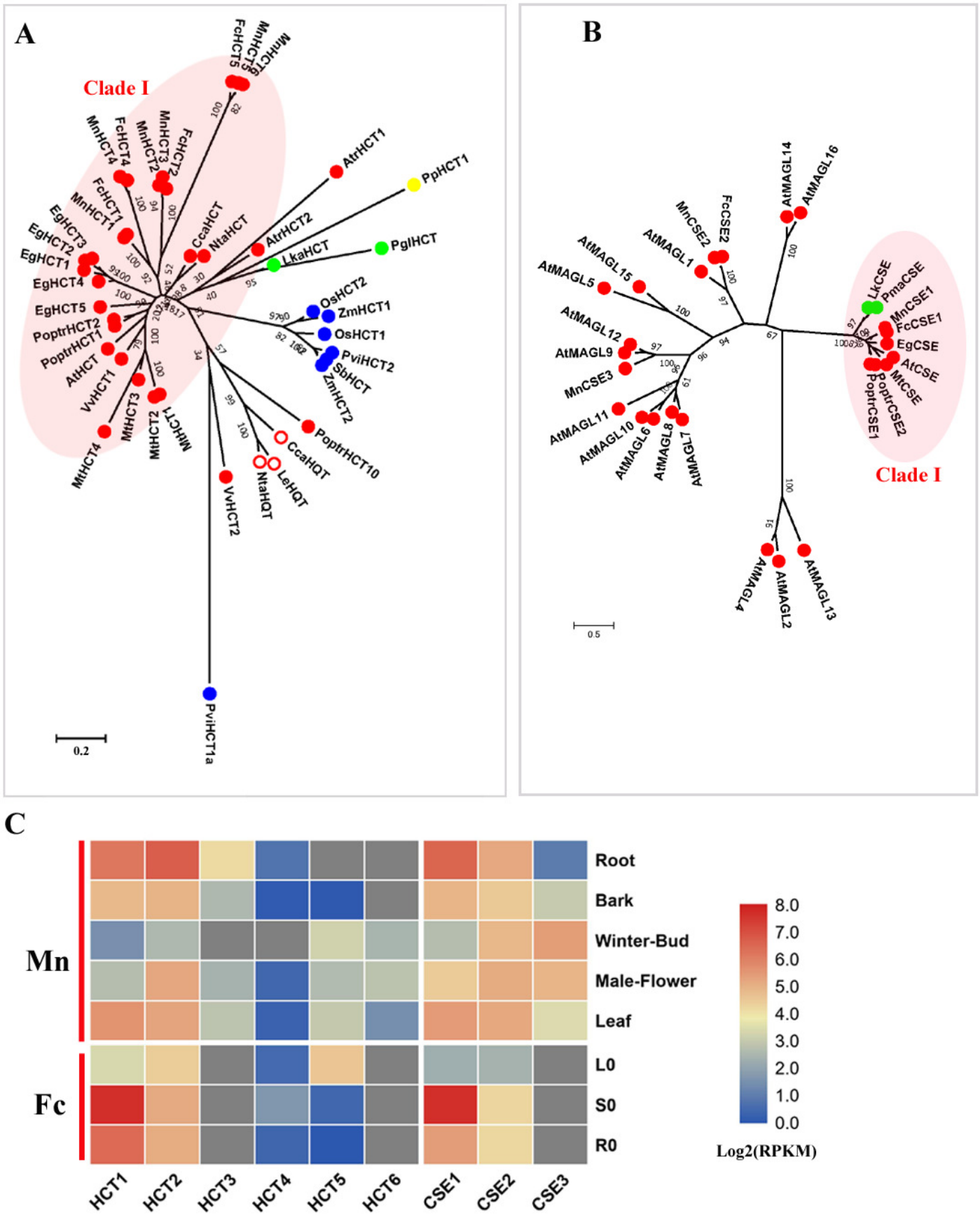

A
Clade I
FcHCT5
MnHCT5
MnHCT4
FcHCT4
MnHCT3
FcHCT3
FcHCT1
MnHCT1
EgHCT3
EgHCT2
EgHCT1
EgHCT4
EgHCT5
PoptrHCT2
PoptrHCT1
AtHCT
VvHCT1
MtHCT3
MtHCT2
MtHCT4
MnHCT2
CcaHCT
NtaHCT
AtrHCT2
LkaHCT
AtrHCT1
PpHCT1
PglHCT
OsHCT2
ZmHCT1
OsHCT1
PviHCT2
SbHCT
ZmHCT2
PoptrHCT10
CcaHQT
NtaHQT
LeHQT
VvHCT2
PviHCT1a
0.2
B
AtMAGL14
AtMAGL16
MnCSE2
FcCSE2
AtMAGL1
AtMAGL15
AtMAGL5
AtMAGL12
AtMAGL9
MnCSE3
AtMAGL11
AtMAGL10
AtMAGL6
AtMAGL8
AtMAGL7
LkCSE
PmaCSE
MnCSE1
FcCSE1
EgCSE
AtCSE
MtCSE
PoptrCSE2
PoptrCSE1
Clade I
AtMAGL4
AtMAGL2
AtMAGL13
0.5
C
Mn
Fc
Root
Bark
Winter-Bud
Male-Flower
Leaf
L0
S0
R0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
Log2(RPKM)
HCT1
HCT2
HCT3
HCT4
HCT5
HCT6
CSE1
CSE2
CSE3

Figure 4

Figure 4. Phylogenetic analysis and expression profile of CCoAOMT and COMT gene families in mulberry.

(A). Phylogenetic analysis of CCoAOMTs; (B). Phylogenetic analysis of COMTs; (C). Expression profiles of *CCoAOMT* and *COMT* gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating proteins from dicots, blue full circles indicating proteins from monocots, green full circles indicating proteins from gymnosperms and yellow full circles indicating proteins from ferns or moss. *Bona fide* clades were marked using different color shadings. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

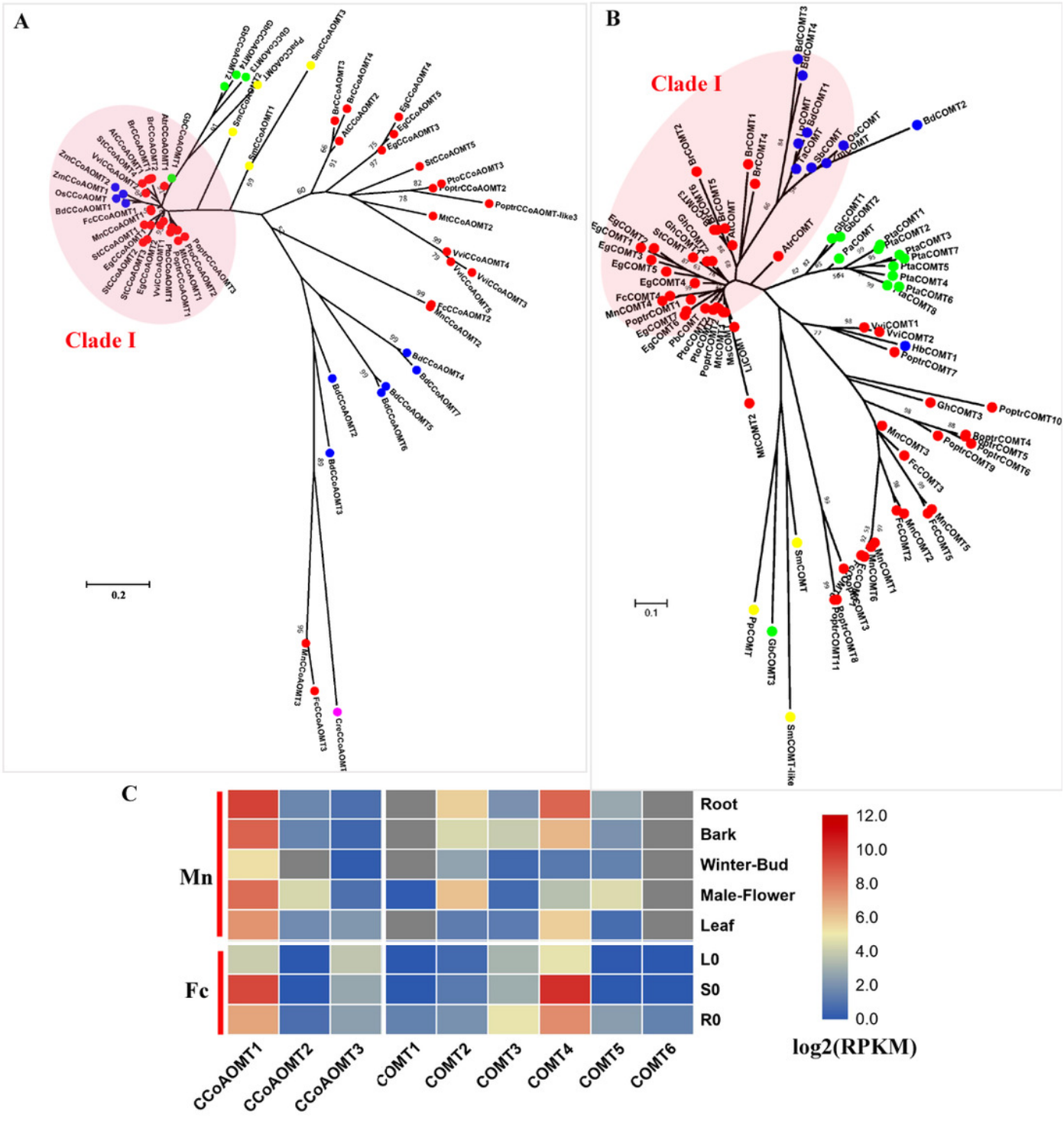

A
Clade I
0.2
B
Clade I
0.1
C
Mn
Fc
Root
Bark
Winter-Bud
Male-Flower
Leaf
L0
S0
R0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
log2(RPKM)
CCoAOMT1
CCoAOMT2
CCoAOMT3
COMT1
COMT2
COMT3
COMT4
COMT5
COMT6

Figure 5

Figure 5. Phylogenetic analysis and expression profile of F5H gene family in mulberry.

(A). Phylogenetic analysis of F5Hs; (B). Expression profiles of *F5H* gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating F5Hs from dicots, blue full circles indicating F5Hs from monocots, green full circles indicating F5Hs from gymnosperms and yellow full circles indicating F5Hs from ferns or moss. *Bona fide* clades were marked using different color shadings. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

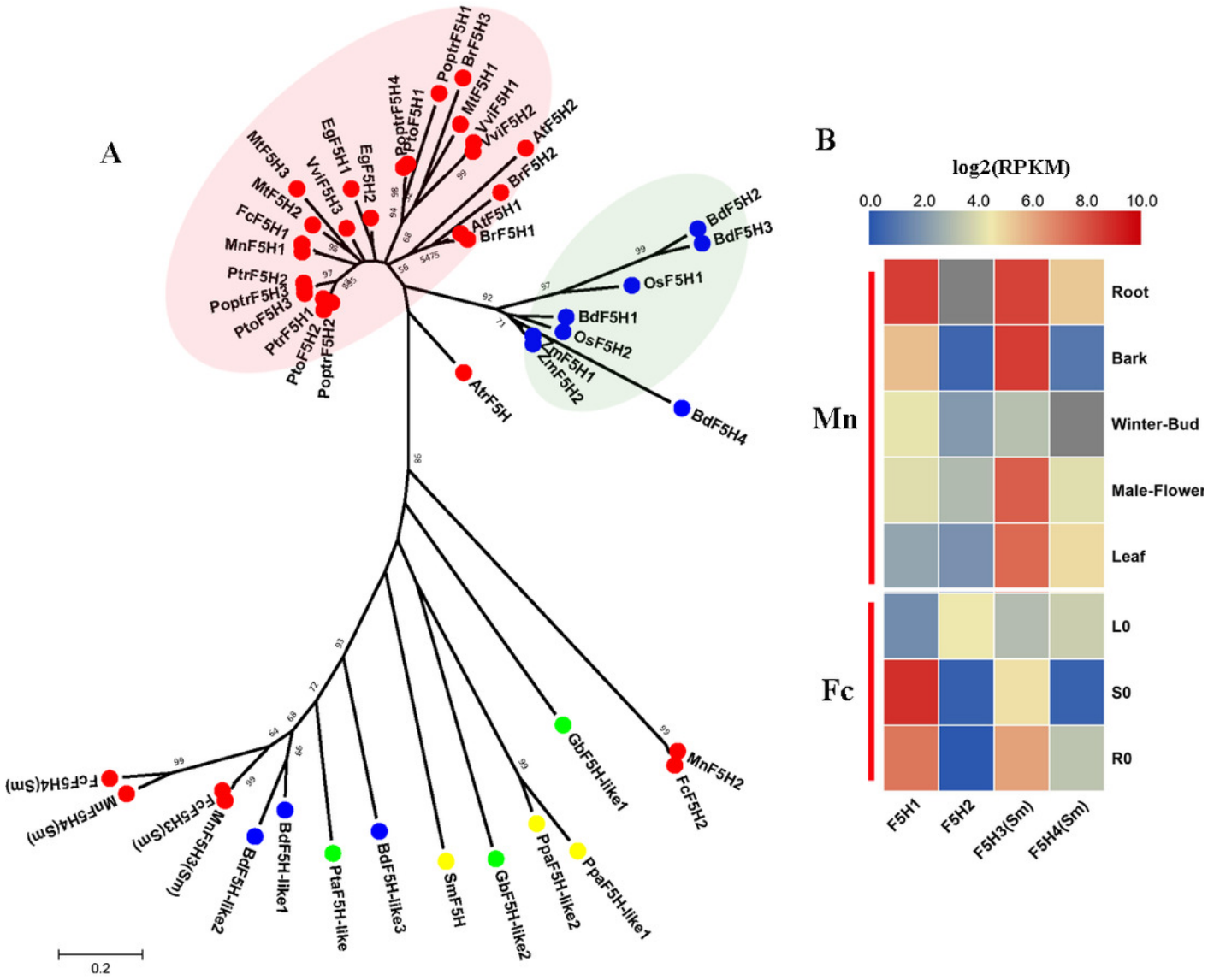

A
MtF5H3
MtF5H2
FcF5H1
MnF5H1
PtrF5H2
PoptrF5H3
PtoF5H3
PtrF5H1
PtoF5H2
PoptrF5H2
EgF5H1
VviF5H3
EgF5H2
PoptrF5H4
PtoF5H1
PoptrF5H1
BrF5H3
MtF5H1
VviF5H1
VviF5H2
AtF5H2
BrF5H2
AtF5H1
BrF5H1
AtrF5H
BdF5H2
BdF5H3
OsF5H1
BdF5H1
OsF5H2
ZmF5H1
ZmF5H2
BdF5H4
MnF5H2
FcF5H2
GbF5H-like1
PpaF5H-like2
PpaF5H-like1
GbF5H-like2
SmF5H
BdF5H-like3
PtaF5H-like
BdF5H-like1
BdF5H-like2
MnF5H3(Sm)
FcF5H3(Sm)
MnF5H4(Sm)
FcF5H4(Sm)
0.2
B
log2(RPKM)
0.0
2.0
4.0
6.0
8.0
10.0
Mn
Fc
Root
Bark
Winter-Bud
Male-Flower
Leaf
L0
S0
R0
F5H1
F5H2
F5H3(Sm)
F5H4(Sm)

Figure 6

Figure 6. Phylogenetic analysis and expression profile of CCR and CAD gene families in mulberry.

(A). Phylogenetic analysis of CCRs; (B). Phylogenetic analysis of CADs; (C). Expression profiles of *CCR* and *CAD* gene family in different tissues or organs in *Morus notabilis* and *Fengchi.* Red full circles indicating proteins from dicots, blue full circles indicating proteins from monocots, green full circles indicating proteins from gymnosperms and yellow full circles indicating proteins from ferns or moss. *Bona fide* clades were marked using different color shadings. Putative protein sequences were used for phylogenetic analysis and the sequences information is available in Table S2. Mn indicating *Morus notabilis* and Fc indicating *Fengchi.* L0, leaf without zinc treatment; S0, stem without zinc treatment; R0, root without zinc treatment.

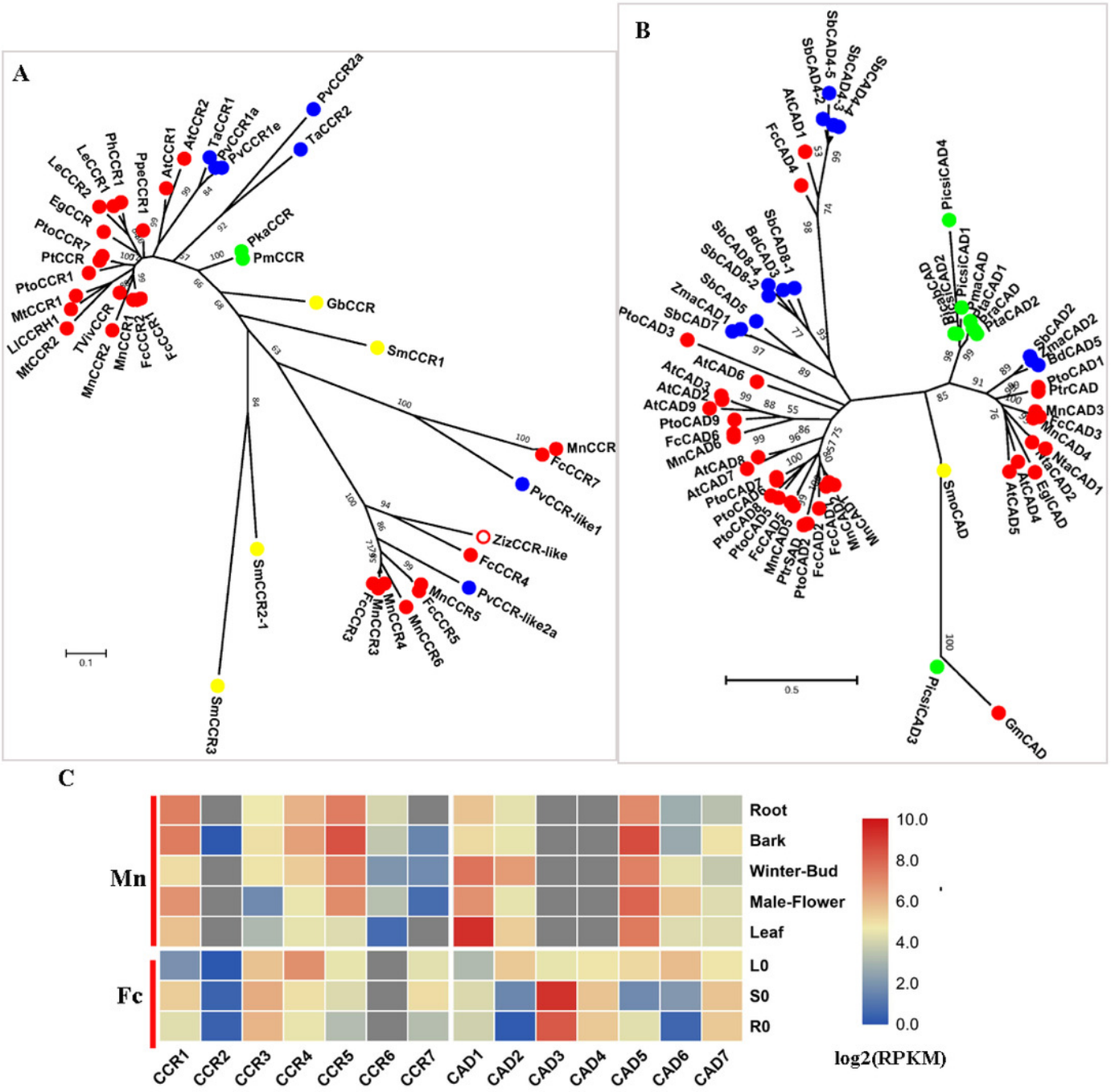
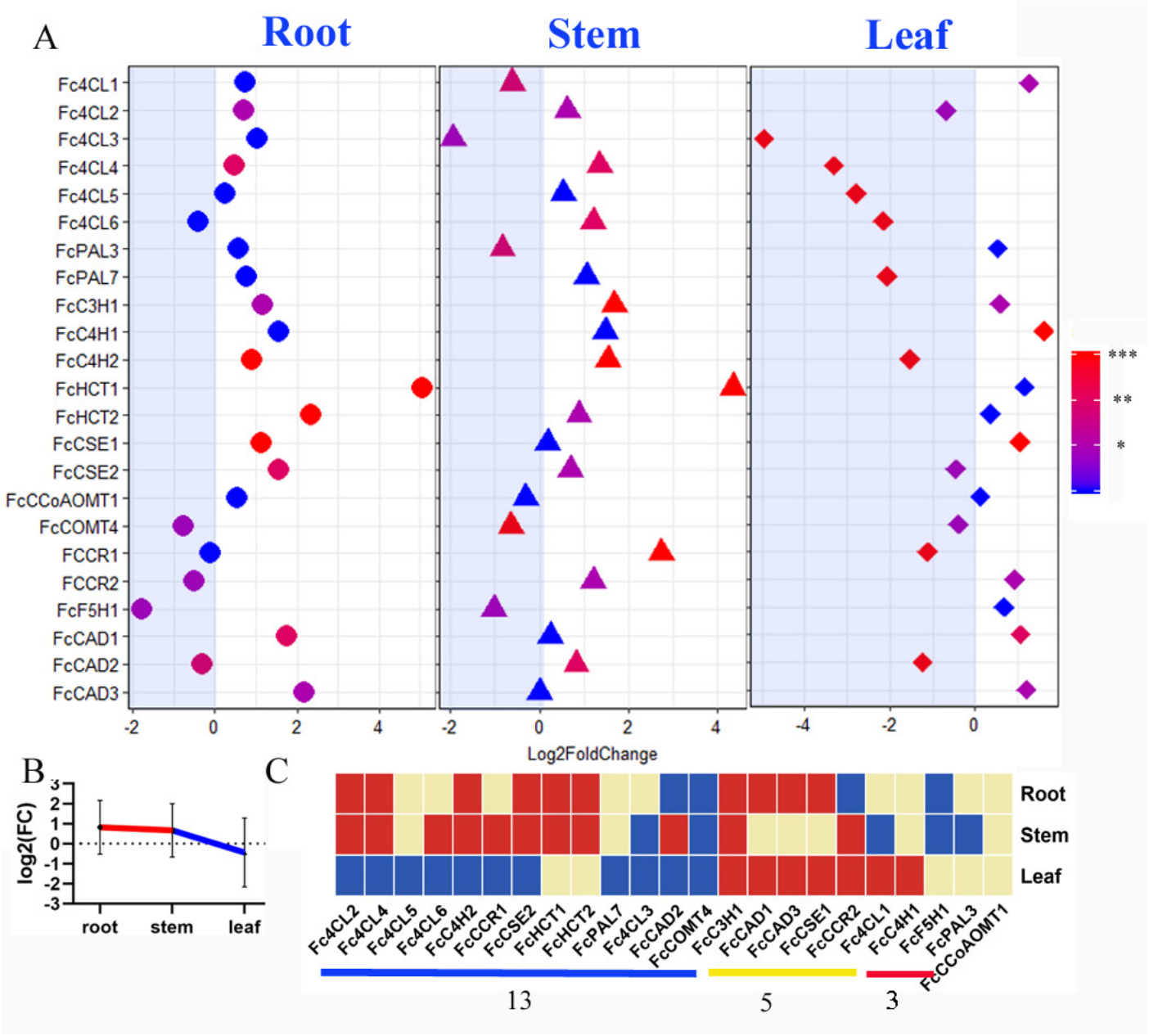

A
PvCCR2a
TaCCR2
PkaCCR
PmCCR
GbCCR
SmCCR1
MnCCR
FcCCR7
PvCCR-like1
ZizCCR-like
FcCCR4
PvCCR-like2a
MnCCR5
FcCCR5
MnCCR6
MnCCR4
FcCCR3
MnCCR3
SmCCR2-1
SmCCR3
FcCCR1
MnCCR1
MnCCR2
TVivCCR
LlCCRH1
MtCCR2
MtCCR1
PtoCCR1
PtCCR
PtoCCR7
EgCCR
LeCCR2
LeCCR1
PhCCR1
PpeCCR1
AtCCR1
AtCCR2
TaCCR1
PvCCR1a
PvCCR1e
0.1
B
SbCAD4-5
SbCAD4-3
SbCAD4-4
AtCAD1
FcCAD4
SbCAD8-1
BdCAD3
SbCAD8-4
SbCAD8-2
SbCAD5
ZmaCAD1
SbCAD7
PtoCAD3
AtCAD6
AtCAD3
AtCAD2
AtCAD9
PtoCAD9
FcCAD6
MnCAD6
AtCAD8
AtCAD7
PtoCAD7
PtoCAD6
PtoCAD8
PtoCAD5
FcCAD5
MnCAD5
PtrSAD
PtoCAD2
FcCAD2
MtCAD2
MtCAD1
FcCAD1
PicsiCAD4
PicsiCAD2
PicsiCAD
ZmaCAD
PtaCAD1
PtaCAD
PtaCAD2
SbCAD2
ZmaCAD2
BdCAD5
PtoCAD1
PtrCAD
MnCAD3
FcCAD3
MnCAD4
NtaCAD2
NtaCAD1
EglCAD
AtCAD4
AtCAD5
SmoCAD
PicsiCAD3
GmCAD
0.5
C
Mn
Fc
Root
Bark
Winter-Bud
Male-Flower
Leaf
L0
S0
R0
CCR1
CCR2
CCR3
CCR4
CCR5
CCR6
CCR7
CAD1
CAD2
CAD3
CAD4
CAD5
CAD6
CAD7
10.0
8.0
6.0
4.0
2.0
0.0
log2(RPKM)

Figure 7

Figure 7. Expression change of bona fide clade genes in response to excess zinc stress in mulberry.

(A). Fold change of expression levels of 23 *bona fide* genes in *Fengchi* after excess zinc treatment; (B). Overall change of monolignol pathway in different organs after excess zinc treatment in *Fengchi* ; (C). Clustering of 23 *bona fide* clade genes expression pattern in response to Zinc stress. Two biological replicates with three technical replicates respectively were performed for qRT-PCR. P-value was calculate using SPSS 19.0. “*” indicates $0.01<p<0.05$; “**” indicates $0.001<p<0.01$ and “***” indicates $p<0.001$.


A
Root
Stem
Leaf
Fc4CL1
Fc4CL2
Fc4CL3
Fc4CL4
Fc4CL5
Fc4CL6
FcPAL3
FcPAL7
FcC3H1
FcC4H1
FcC4H2
FcHCT1
FcHCT2
FcCSE1
FcCSE2
FcCCoAOMT1
FcCOMT4
FCCR1
FCCR2
FcF5H1
FcCAD1
FcCAD2
FcCAD3
-2
0
2
4
-4
Log2FoldChange
***
**
*
B
log2(FC)
3
2
1
0
-1
-2
-3
root
stem
leaf
C
Root
Stem
Leaf
Fc4CL2
Fc4CL4
Fc4CL5
Fc4CL6
FcC4H2
FcCCR1
FcCSE2
FcHCT1
FcHCT2
FcPAL7
Fc4CL3
FcCAD2
FcCOMT4
FcC3H1
FcCAD1
FcCAD3
FcCSE1
FcCCR2
Fc4CL1
FcC4H1
FcF5H1
FcPAL3
FcCCoAOMT1
13
5
3