Molecular and biochemical characterization of two 4-coumarate: CoA ligase genes in tea plant (Camellia sinensis)

Mingzhuo Li1,3 · Lili Guo2 · Yeru Wang2 · Yanzhi Li1 · Xiaolan Jiang1 · Yajun Liu1 · De-Yu Xie3 · Liping Gao2 · Tao Xia1

Received: 12 September 2021 / Accepted: 22 March 2022 / Published online: 12 May 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Key message Two 4-coumarate: CoA ligase genes in tea plant involved in phenylpropanoids biosynthesis and response to environmental stresses.

Abstract Tea plant is rich in flavonoids benefiting human health. Lignin is essential for tea plant growth. Both flavonoids and lignin defend plants from stresses. The biosynthesis of lignin and flavonoids shares a key intermediate, 4-coumaroyl-CoA, which is formed from 4-coumaric acid catalyzed by 4-coumaric acid: CoA ligase (4CL). Herein, we report two 4CL paralogs from tea plant, Cs4CL1 and Cs4CL2, which are a member of class I and II of this gene family, respectively. Cs4CL1 was mainly expressed in roots and stems, while Cs4CL2 was mainly expressed in leaves. The promoter of Cs4CL1 had AC, nine types of light sensitive (LSE), nine types of stress-inducible (SIE), and two types of meristem-specific elements (MSE). The promoter of Cs4CL2 also had AC and nine types of LSEs, but only had two types of SIEs and did not have MSEs. In addition, the LSEs varied in the two promoters. Based on the different features of regulatory elements, three stress treatments were tested to understand their expression responses to different conditions. The resulting data indicated that the expression of Cs4CL1 was sensitive to mechanical wounding, while the expression of Cs4CL2 was UV-B-inducible. Enzymatic assays showed that both recombinant Cs4CL1 and Cs4CL2 transformed 4-coumaric acid (CM), ferulic acid (FR), and caffeic acid (CF) to their corresponding CoA ethers. Kinetic analysis indicated that the recombinant Cs4CL1 preferred to catalyze CF, while the recombinant Cs4CL2 favored to catalyze CM. The overexpression of both Cs4CL1 and Cs4CL2 increased the levels of chlorogenic acid and total lignin in transgenic tobacco seedlings. In addition, the overexpression of Cs4CL2 consistently increased the levels of three flavonoid compounds. These findings indicate the differences of Cs4CL1 and Cs4CL2 in the phenylpropanoid metabolism.

Keywords 4-Coumaric acid: CoA ligase · Flavonoids · Lignin · Metabolic engineering of phenylpropanoid · Stress

Introduction

Phenylpropanoids play important roles in plant growth and defense against biotic and abiotic stresses (Dixon et al. 2013; Vogt 2010). The biosynthetic pathway of plant phenylpropanoids starts from the shikimate pathway to lignin, flavonoids, hydroxycinnamoyl esters, and other phenolics (Weisshaar and Jenkins 1998; Grace and Logan 2000; Dixon et al. 2013; Vogt 2010). Lignin is the second most abundant biopolymer next to cellulose and plays crucial role in sustaining the structural integrity of cell wall and stiffness and strength of plant stem (Boerjan et al. 2003; Grima-Pettenati and Goffner 1999; Whetten and Sederoff 1995). Additionally, lignin is essential for transporting water and solutes in plants and defending plants from pathogenic attack (Grima-Pettenati and Goffner 1999; Whetten and Sederoff 1995; Boerjan et al.
Flavonoids, the biggest group of plant polyphenols, play important roles in plant’s interaction with different environmental stresses and plant pollination. In general flavonoid compounds can protect plants from damages caused by UV irradiation, insects, and pathogens (Mierziak et al. 2014; Winkel-Shirley 2002). Furthermore, flavonoids are potent nutrients to provide important healthy benefits to humans, such as antioxidative activity (Atmani et al. 2009; Ververidis et al. 2007).

Flavonoids and lignin are derived from the general phenylpropanoid pathway starting with phenylalanine or tyrosine, two products of the shikimate pathway. The three consecutive entry enzymes of the pathway are phenylalanine ammonia lyase (PAL) (Camm and Towers 1973; Wanner et al. 1995) that catalyzes L-phenylalanine to produce cinnamic acid, cinnamate 4-hydroxylation (C4H) that converts cinnamic acid to 4-coumaric acid, and 4-coumarate-CoA ligase (4CL) that transforms 4-coumaric acid to 4-coumaroyl-CoA in the presence of CoA (Bell-Lelong et al. 1997; Schilmiller et al. 2009) (Hu et al. 1998; Schilmiller et al. 2009) (Fig. 1A). Tyrosine ammonia lyase (TAL) transforms L-tyrosine to 4-coumaric acid (Camm and Towers 1969).

To date, 4CL has been characterized to be a small family that plays diverse catalytic roles involved in the metabolic pathways toward flavonoids and lignin (Hamberger and Hahlbrock 2004; Cukovica et al. 2001; Hu et al. 1998) (Fig. 1A). Arabidopsis thaliana includes more than four 4CL isoforms (Li et al. 2015b; Ehlting et al. 1999b). Populus trichocarpa was reported to have 17 4CL isoforms (Rui et al. 2010). Five 4CL members were reported in rice (Jinshan et al. 2011). In all investigated plants, 4CL homologs are classified into two classes, class I and class II (Fig. 1B). On the one hand, 4CL members in class I have been characterized to preferentially involve in the biosynthesis of lignin and wound responses (Lavhale et al. 2018). Examples of members include At4CL1 from Arabidopsis (Li et al. 2015a), Pta4CL from P. trichocarpa (Shi et al. 2010), and Os4CL3, Os4CL4 and Os4CL5 from rice (Gu et al. 2011), and Gm4CL1 and Gm4CL2 from soybean (Lindermayr et al. 2002b). These 4CL members transform caffeic acid and ferulic acid to form their CoA ester toward the lignin biosynthesis and their expression is induced by mechanical wounding (Soltani et al. 2006). Gm4CL1 and Gm4CL2 were demonstrated to transform 4-coumaric, caffeic, ferulic, and sinapic acids to their corresponding CoA esters with a high affinity (Lindermayr et al. 2002a). On the other hand, the class II members have been characterized to preferably associate with the biosynthesis of flavonoids and their transcription is induced by UV irradiation (Cukovica et al. 2001; Lindermayr et al. 2002b; Gui et al. 2011; Hamberger and Hahlbrock 2004; Hu et al. 1998; Costa et al. 2005). Examples include At4CL3 from Arabidopsis (Ehlting et al. 1999b), Os4CL2 from rice (Gui et al. 2011), Gm4CL3 and Gm4CL4 from soybean (Sun et al. 2013), and Pt4CL2 from P. trichocarpa (Harding et al. 2002). At4CL3 was reported to preferentially transform 4-coumaric acid to 4-coumaroyl-CoA (Ehlting et al. 1999a; Hamberger and Hahlbrock 2004). The recombinant Gm4CL3 and Gm4CL4 were demonstrated to transform 4-coumaric and caffeic acids to their corresponding CoA esters (Lindermayr et al. 2002a).

Manipulation of 4CL has been reported to alter either lignin or flavonoid biosynthesis. An ectopic expression of anti-4CL (At4CL1) in A. thaliana plants reduced the lignin contents via a significant decrease of S-lignin (Lee et al. 1997). An overexpression of Pt4CL1 in tobacco plants increased the lignin by 25% in xylem of stem (Lu et al. 2004). A class II 4CL gene of lobolly pine, Pinta4CL3, was overexpressed in P. trichocarpa to obtain multiple transgenic plants (Chen et al. 2014). Metabolic profiling indicated the significant increase of hydroxycinnamoyl-quinate/shikimate esters, caffeoyl-quinate/shikimate esters, and cinnamoyl-quinate in leaves of transgenic populus plants. In addition, the levels of kaempferol, rutin (quercetin-3-O-rutinoside), and quercetin-3-O-glucuronide were increased in transgenic leaves. By contrast, the overexpression of Pinta4CL3 did not alter the levels of coniferyl or sinapyl alcohols toward the formation of lignin. Recently, the differential roles of class I and II members gained genetic evidence from Arabidopsis. The knockout of At4CL1 (class I type) was reported to reduce lignin by 20%, while the knockout of At4CL3 (class II type) was reported to decrease the anthocyanin content by 70% (Li et al. 2015a). The genetic evidence shows the functional differentiation of two isomers in A. thaliana.

Tea plant (Camellia sinensis) is rich in flavonoids and other polyphenols in leaves and buds with multiple benefits to human health (Sueoka et al. 2001; Liu et al. 2012; Jiang et al. 2015). Tea is one of the most popular nonalcoholic beverages worldwide. To date, the biosynthesis of tea flavonoids has gained intensive studies in both genomics and functional genomes (Wei et al. 2018). Genes encoding flavonoid biosynthesis have been functionally characterized to understand their roles associated with the high production of tea flavonoids (Wang et al. 2014, 2018; Liu et al. 2019; Zhao et al. 2017). These accomplishments greatly enhance the understanding of the biosynthesis of tea flavonoids. In comparison, the entry pathway genes of tea phenylpropanoids, such as PAL, C4H, and 4CL, are open for functional characterization, which is essential to elucidate the entire pathway of tea flavonoids. Herein, we report functional characterization.

© Springer
Fig. 1 The step catalyzed by 4-coumarate: CoA ligase (4CL) toward the biosynthesis of lignin and flavonoids and a phylogenetic tree of the 4CL family. A Biosynthetic pathways start from the general phenylpropanoid pathway to lignin and flavonoids. Class I and class II 4CL enzymes are shown to catalyze the step toward lignin and flavonoids. B A phylogenetic tree was constructed with amino acid sequences of Cs4CL1, Cs4CL2, and 17 other 4CL proteins. The tree shows that class I subgroup 4CL and class II subgroup 4CL differentially catalyze the step toward lignin and flavonoid. Abbreviations, PAL: phenylalanine ammonia-lyase, C4H: cinnamate-4-hydroxylase, 4CL: 4-coumarate-CoA ligase, CHS: chalcone synthase, CHI: chalcone isomerase, F3H: flavanone 3-hydroxylase, F3′H: flavonoid 3′-hydroxylase, F3′5′H: flavonoid 3′, 5′-hydroxylase, DFR: dihydroflavonol-4-reductase, ANS: anthocyanidin synthase, anthocyanidin reductase (ANR), LAR: leucoanthocyanidin reductase
of two tea 4CL genes, Cs4CL1 and Cs4CL2. Molecular and phylogenetic analysis indicated that Cs4CL1 belonged to class I and was highly expressed in roots and stems, while Cs4CL2 was clustered in class II and was highly expressed in buds and juvenile leaves. Transgenic, biochemical, and transcriptional data indicated that Cs4CL1 appeared to be preferably associated with the biosynthesis of lignin, while Cs4CL2 appeared to be rather involved in the biosynthesis of flavonoids.

### Materials and methods

#### Plant materials

More than 500 5-year old shrubs of Shu Chazao tea (C. sinensis var. sinensis cv. “Shu Chazao”) are grown in the research station at Anhui Agricultural University for different research purposes. In middle spring, shoot buds, the 1st, 2nd, and 3rd young leaves on new sprouts, fully expended mature leaves, young stems, and roots were collected in liquid nitrogen and then stored at −80°C until experimental purposes.

#### Construction of binary vectors and plant transformation

The cloning method of promoters followed a genome-walker method according to the Genome-Walker Kit (Clontech). The primers (Table S1) were designed for PCR to amplify 1000 nucleotides of promoter sequences. The amplified PCR products were cloned to plasmid pEASY-T1 and then sequenced at BGI to confirm the accuracy. Identification of regulatory elements in the promoter sequences was completed using three online promoter analysis tools at PLACE (http://www.dna.affrc.go.jp/htdocs/PLACE/), Plantcare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), and PlantPan (http://plantpan.itps.ncku.edu.tw/).

#### Gene cloning and phylogenetic analysis

Total RNA was isolated from plant tissues using an RNAIsolate for Plant Tissue Kit (Takara, Dalian, China) according to the manufacturer’s protocol. The first strand cDNA was synthesized using a PrimeScript® RT Reagent Kit (Takara, Dalian, China). According to an EST sequence of Cs4CL1, primer pairs (Supplementary Table S1) were designed for RACE polymerization chain reaction (PCR) to clone the full length of Cs4CL1. Based on available cDNA sequence of Cs4CL2, forward and reverse primer pairs were designed through Primer Premier 5 software according to the ORF (open read frame) sequences (Supplementary Table S1). The ORFs of Cs4CL1 and Cs4CL2 were amplified with a PCR program, which consisted of 98°C for 30 s, 29 cycles of 98°C for 10 s, 60°C for 20 s, and 72°C for 30 s, and then a 10 min extension at 72°C. The amplified PCR products were purified and then cloned to the entry vector pDONR207 using the Gateway® BP Clonase® Enzyme mix (Invitrogen, Carlsbad, CA). After sequencing confirmed the accuracy of the inserts in the recombinant pDONR207 plasmids, ORFs were cloned to the PCB2004 binary vector via an exchange reaction using Gateway® LR Clonase™ enzyme mix (Invitrogen, Carlsbad, CA). After sequencing confirmed the accuracy of the inserts in the recombinant plasmids, PCB2004-Cs4CL1 and PCB2004-Cs4CL2, were introduced into Agrobacterium tumefaciens strain EHA105 for tobacco transformation. The positive EHA105 colonies were selected on agar-solidified medium containing 50 mg/L kanamycin. The transformation of tobacco plants was completed via a leaf disc method (Li et al. 2017). The positive transgenic plants were screened on agar-solidified MS medium supplemented with 25 mg/L of phosphinothricin. Multiple phosphinothricin-resistant plants were verified to be transgenic by both genomic DNA-based PCR and RT-PCR, and then planted in pot soil in the greenhouse.

Cloning and sequence analysis of Cs4CL1 and Cs4CL2 promoters

The resulting ORF sequences were deduced to obtain amino acid sequences of Cs4CL1, Cs4CL2, 12 other Cs4CL or Cs4CL-like proteins, and 17 4CL homologs of other plants via using the MEGA6.1 software. The Gene-Bank IDs of all 4CL homologs were listed in supplementary Table S2. The tree nodes were evaluated with a bootstrap for 1000 replicates and the evolutionary distances were estimated with a p-distance method.
Subcellular localization

The stop codons were removed from the ORFs of both Cs4CL1 and Cs4CL2 by PCR and then the resulting fragments were cloned to the pGWB5 vector, in which each was fused at the 5′end of an EGFP reporter gene using a Gateway cloning system as described above. Two recombinant constructs, pGWB5-Cs4CL1-EGFP and pGWB5-Cs4CL2-EGFP, were obtained and then introduced into A. tumefaciens EHA105. The positive colonies were activated and then used to infect leaves of 45 days old N. benthamiana. After infection for 48 h, leaves were used to examine the EGFP fluorescence under an Olympus FV1000 confocal microscope (Olympus, Tokyo, Japan).

Quantitative reverse transcription PCR (RT-qPCR) and semi-quantitative RT-PCR analysis of gene expression in different tea tissues, treatments, and transgenic tobacco plants

RT-qPCR was performed to understand gene expression in different tissues using SYBR-Green PCR Mastermix (Invitrogen, Carlsbad, CA) on a CFX96™ instrument (Bio-RAD, California, USA) by following the manufacturer’s instructions. Gene-specific primer pairs and thermal programs for Cs4CL1 and Cs4CL2 (Table S1) were designed for PCR. A tea CsACTIN (KA280216.1) gene was used as reference for normalization of RT-qPCR. Amplified products were monitored using an optical reaction module and amplified values were normalized against the CsACTIN (KA280216.1). Tobacco NtACTIN (LOC107809070) was used as a control of transgene expression (Pang et al. 2007). For identification of positive transgenic tobacco lines, both genomic DNA-based PCR and semi-quantitative RT-PCR were completed as previously described (Zhao et al. 2013). All primers are listed in Table S1.

Heterologous expression in Escherichia coli and recombinant protein purification

The ORFs of both Cs4CL1 and Cs4CL2 were ligated to a pet-SUMO vector (Life technology biology, USA) with a T4-ligase according to the manufacturer’s protocol. The resulting positive Cs4CL1-petSUMO and Cs4CL2-petSUMO plasmids were transformed into a competent E. coli BL21 strain. Positive colonies were obtained after screening on agar-solidified Luria–Bertani (LB) medium supplemented with 100 μg/ml kanamycin. One positive colony for each construct was inoculated into 200 ml liquid LB medium supplemented with 100 μg/ml kanamycin in a 500 ml E-flask. Meanwhile, a colony harboring the pet-SUMO vector alone was cultured as a control. The flasks were placed on a rotary shaker with a speed of 250 rpm at 37 °C. After the concentrations of optical density of suspension cultures reached around 1.00 at 600 nm, the flasks were divided into two groups. One group was added with IPTG to a final concentration of 0.2 mM in the suspension cultures. The other group was not added IPTG in the suspension cultures as a control. Moreover, IPTG was added in the suspension cultures from E. coli with the pet-SUMO vector as another control. The temperature was then changed to 28 °C to induce protein expression. After 24 h of induction, the suspension cultures were transferred to tubes, which were centrifuged for 10 min at 5000 rpm. After E. coli cells were harvested, crude proteins were extracted and examined with electrophoresis on a 12% SDS polyacrylamide gel and staining of Coomassie brilliant blue as described previously (Zhao et al. 2017). Recombinant Cs4CL1 and Cs4CL2 were purified with an affinity chromatography consisting of Ni resin (New England Biolabs, MA, USA). Two buffers were used for affinity chromatography. Buffer A consisted of 20 mM pH 7.4 Tris–HCl and 200 mM NaCl. Buffer B was composed of buffer A and 100 mM imidazole. A column was firstly washed and equilibrated with buffer A. After the crude protein extracts were loaded onto the column, buffer A was used to elute all unbound proteins. Then, buffer B was used to elute recombinant Cs4CL1 and Cs4CL2, the quality of which was further examined with electrophoresis on a 12% SDS polyacrylamide gel and staining of Coomassie brilliant blue.

Enzymatic assays

Enzymatic assay was performed in order to characterize the catalytic activity of Cs4CL1 and Cs4CL2 according to a published protocol with minor modifications (Knobloch and Hahlbrock, 1975). Briefly, coenzyme A and three substrates, 4-coumaric acid, caffeic acid, and ferulic acid, were purchased from Sigma (St. Louis, USA) for catalytic assays. Prior to optimize reaction conditions, we tested enzymatic activity with the three different substrates in a 500 μl volume. The reaction mixture contained 0.5 M pH 7.5 phosphate buffer, 0.3 mM coenzyme A (CoA), 5 mM ATP, 5 mM MgCl2, 5 μg purified recombinant enzyme, and 0.4 mM substrates (4-coumaric acid, caffeic acid, and ferulic acid) in a 1.5 ml tube. The reactions were incubated at 30 °C for 30 min and then terminated with the addition of 50 μl methanol. After the reactions were centrifuged for 10 min at 12,000 rpm, the supernatants were transferred to new tubes for immediate HPLC analysis. To further optimize pH and temperature values of Cs4CL1 and Cs4CL2, 4-coumaric acid was used for assays, while other conditions were not changed. To estimate Km and Vmax values, reactions were carried out in a 2 ml reaction volume containing 0.5 M pH7.5 PBS buffer, 0.03 to 0.3 mM CoA, 5 mM ATP, 0.04–0.4 mM substrates (4-coumaric
acid, caffeic acid, and ferulic acid), 2–20 μg recombinant Cs4CL1 and Cs4CL2 protein. The resulting data were used to develop Lineweaver–Burk plots to characterize kinetics and estimate the Km and Vmax values of Cs4CL1 and Cs4CL2.

HPLC analysis of enzymatic reaction products

The enzymatic reaction products were analyzed on a Shimadzu LC20-AT instrument with a full wavelength detection. The elution of products was monitored at 333 nm for 4-coumaroyl-CoA, 350 nm for caffeoyl-CoA, and 346 nm for feruloyl-CoA. The mobile phase consisted of solvent A (0.5% acetic acetate in double deionized water) and B (100% HPLC-grade acetonitrile). A gradient elution program used consisted of 30–90% solvent B from 0 to 23 min, 90–30% solvent B from 23 to 29 min, and then followed by a 10 min column washing with 30% buffer B.

Measurement of total lignin in transgenic and wild type tobacco plants

Lignin measurement was carried out by following the Syros method (Halpin et al. 1994). Briefly, roots, stems, and leaves were harvested from 90-day old plants grown in the pot soil, pooled together, and then ground into fine powder in liquid nitrogen. One hundred mg of powder was suspended in 1 ml 50% ethyl alcohol (in water) in a 2 ml tube. After 3 h of extraction at 80 °C, 1 ml methanol was added into the mixture, which was incubated for 1 h at 80 °C. Tubes were centrifuged at 12,000 rpm for 10 min to obtain supernatant and residue phases. The supernatant was removed to a waste container. The remaining residues were fully dried at 60 °C in an oven. Ten-milligram of dry powder was weighed to a 1.5 ml tube, suspended in 5 ml 25% (w/w) acetyl bromide in acetic acid, and treated for 30 min at 70 °C. The mixtures were added with 0.2 ml 70% perchloric acid in water, mixed thoroughly, and then continuously treated at 70 °C for 30 min. After cooling to the room temperature, tubes were centrifuged at 3000×g for 15 min to obtain supernatants and residues. The supernatant of each treatment was pipetted into a new 50 ml tube, which was added with 5 ml of 2 M NaOH. Then, glacial acetic acid was immediately added to the mixture to adjust the final volume to 25 ml, followed by mixing thoroughly. One ml of mixture was used to measure absorbent value at 280 nm on UV–visible Hitachi U-5100 spectrophotometer (Hitach, Tokyo, Japan). According to a previously published method (Piquemal et al. 2002), the absorbent values were converted to the lignin content in samples.

Extraction of polyphenolic compounds from transgenic and wild type tobacco plants

Leaves from 90-day old wild type, Cs4CL1 and Cs4CL2 transgenic tobacco plants were collected to analyze polyphenols. Phenolic compounds were extracted using the following procedure. Samples were ground into fine powder in liquid nitrogen. Powdered samples (150 mg) were suspended in 1 ml 80% methanol: 1% hydrochloric acid in a 1.5 ml tube. Tubes were completely vortexed and then placed in the room temperature for 20 min, followed by 10 min of centrifugation at 12,000 rpm. The supernatant was transferred to a new tube. This extraction was repeated once to obtain a final volume of 2 ml extract for each sample. All extracts were then filtered through a 0.22 μm membrane into a new tube prior to ultra-high-performance liquid chromatography (UPLC)-MS/MS analysis described below.

Analysis of metabolites by UPLC-MS/MS

Products from enzymatic assays of Cs4CL1 and Cs4CL2 and polyphenolic metabolites extracted from wild type and Cs4CL1 and Cs4CL2 transgenic tobacco plants were analyzed using UPLC-MS/MS on an Agilent LC–MS system (Palo Alto, CA, USA). Compounds were separated in an Agilent 20RBAX RRHD Eclipse Plus C18 column (particle size: 1.8 μm, length: 100 mm, and internal diameter: 2.1 mm Palo Alto, CA, USA). The column oven, mobile gradient, and electrospray ionization technique were as described previously (Jiang et al. 2013a).

Statistical evaluation

All assays described above were performed with either three repetitive experiments or three biological replicates. Mean values were calculated and then statistically evaluated by Student t-test using the SPSS software. P-values less than 0.05 were evaluated to be statistically significant.

Results

Cloning of two 4CL genes from tea plants

We collected different tea organs and then pooled them together to extract total RNA samples, which were used to construct a cDNA library for EST sequencing and then to clone 4CL genes. After sequence assembly, a 4CL EST sequence was annotated from the contigs. Based on this sequence, a 3′ and 5′ rapid application of cDNA ends (RACE) was completed to amplify a full length cDNA, namely Cs4CL1 (ID#: KY615680). Sequence analysis revealed that the open read frame (ORF) of Cs4CL1
included 1629 bp nucleotides, which were deduced to encode 543 amino acids. The molecular mass and the isoelectric point (IP) of Cs4CL1 were predicted to be 59.55 KD and 5.72, respectively. Sequence analysis indicated that Cs4CL1 was different from a reported 4CL sequence (ID#: ABA040922.1) curated in the GenBank. Next, based on the sequence of ABA040922.1, we designed gene special primers and used PCR to amplify the full length of a cDNA from “Shu Chazao”, the sequence of which was 95% identical to ABA040922.1. a. We named this cDNA as Cs4CL2 (ID#: KY615679). The ORF of Cs4CL2 included 1713 bp nucleotides, which were deduced to encode 570 amino acids with a molecular mass of 61.4 KD and IP at 5.74. An amino acid sequence comparison showed that Cs4CL1 and Cs4CL2 shared 61% identity (Fig. S1). In addition, Cs4CL1 had 66% identity to Arabidopsis thaliana 4CL1 (At4CL1) and Cs4CL2 had 68% identity to the At4CL3 (Fig. S1). In addition, the analysis of functional domains revealed that both Cs4CL1 and Cs4CL2 had the AMP binding domain BOX I (SSGTTLPKGV) and catalytic domain BOX II (GEICIRG) (Fig. S1). These features suggested that Cs4CL1 and Cs4CL2 might catalyze a coenzyme CoA ligation reaction in tea plants.

A phylogenetic tree was constructed with amino acid sequences of Cs4CL1, Cs4CL2, 12 Cs4CL or Cs4CL-like proteins, and 17 4CL homologs of other plants. The resulting tree was clustered into three subgroups, class I, class II, and a 4CL-like group. Cs4CL1 was clustered in the class I subgroup associated with the formation of lignin and the Cs4CL2 was clustered in the class II subgroup associated with the biosynthesis of flavonoids (Fig. 1B). This result suggests that Cs4CL1 and Cs4CL2 might diverge in their functions.

**Features of regulatory elements in promoters of Cs4CL1 and Cs4CL2**

The genomic sequences of tea plants was recently published (Xia et al. 2017). This advantage allowed us to use genome-walker technology to clone the promoters of Cs4CL1 and Cs4CL2. As a result, 1537 bp nucleotides at the immediate upstream of the Cs4CL ORF was obtained and 987 bp nucleotides at immediate upstream of the Cs4CL2 ORF was cloned. The promoter sequences of Cs4CL1 and Cs4CL2 are shown in Figure S5. Further sequence analysis via PLACE (http://www.dna.affrc.go.jp/htdocs/PLACE/), Plantcare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) and PlantPan (http://plantpan2.itps.ncku.edu.tw/) identified four types of different regulatory elements, including AC-like elements, light sensitive elements (LSEs), meristem specific elements (MSEs), and defense and stress relative elements. The results showed that the Cs4CL1 promoter contained one AC-like element (the binding site of MYB transcription factors) and the Cs4CL2 promoter included five AC-like elements (Fig. 2A). The two promoters each contained nine LSEs. The LSEs of Cs4CL1 included AAAC, BOX4, G-BOX, CGA, GT1, I-BOX, MNF1, SP1 and TCT-MOTIF (Fig. 2A). The LSEs of Cs4CL2 included BOX4, BOX1, CAG, GATA, GT1, I-BOX, SP1, TCCC, and TCT. Only three, BOX4, I-BOX, and SP1, were shared by the two promoters. The Cs4CL1 promoter had two MSEs, CAT-BOX and CCGTCCC-BOX, while the Cs4CL2 promoter did not have these in the cloned sequence. The Cs4CL1 and Cs4CL2 promoters had eight and two defense and stress relative elements, respectively (Fig. 2A).

**Expression profiles in different tissues and three different treatments**

The expression profiles of two genes in seven tea tissues were featured with qRT-PCR analysis. The resulting data indicated that the expression levels of Cs4CL1 were the highest in roots, followed by stems, mature leaves, and young leaves and buds (Fig. 2B). By contrast, a relatively opposite expression trend of Cs4CL2 was observed in these tissues, the highest in buds, followed by the 1st, 2nd, 3rd, and mature leaves, then by stems and roots (Fig. 2B). These data suggest that two genes might be differentially associated with the biosynthesis of phenylpropanoids in tea tissues.

Based on the regulatory elements in the promoters of Cs4CL1 and Cs4CL2 that are responsive to different stressful stimuli, three treatments, UV-B, mechanical wounding, and cold treatments, were designed to understand their effects on the expression of the two genes in young leaves (Fig. 2C). RT-qPCR was completed to understand gene expression levels in treatments of 1-h UV-B, 6-h wounding, and 6-h 4 °C. The results showed that the expression of the two genes differentially responded to the three treatments. In the 60 min treatment of UV-B irradiation, the expression of Cs4CL2 quickly increased in 15 min and decreased at 30, 45, and 60 min, while the expression levels of Cs4CL1 slightly decreased in the first 45 min and then increased slightly at 60 min. During the 6-h wounding treatment, the data indicated that the expression levels of Cs4CL1 reduced in the first 30 min and then increased from 0.5 to 6 h, while the Cs4CL2 did not transcriptionally respond to this treatment. During the 6-h cold treatment, the expression levels of Cs4CL2 did not change significantly at five time points tested, while the expression levels of Cs4CL1 decreased either significantly or slightly.

**Recombinant Cs4CL1 and Cs4CL2 uses three hydroxycinnamate substrates**

Recombinant Cs4CL1 and Cs4CL2 were used to test different substrates to understand the catalytic activity of the
Fig. 2  Regulatory elements in the promoters of Cs4CL1 and Cs4CL2 and expression profiles in tissues and three different treatments. A Schemes show the locations of AC-elements (MYB-putative binding site), light sensitive elements, defense and stresses responsive elements, and meristem specific elements in 1100 and 1000 bp upstream (promoter) sequences of the open reading frames of Cs4CL1 and Cs4CL2, respectively. Promoter analysis was completed with three types of software. B qRT-PCR results show the expression profiles of Cs4CL1 and Cs4CL2 in the buds, 1st, 2nd, 3rd and mature leaves, stems, and roots. qRT-PCR results indicate effects of UV-B light versus white (control) (C), mechanical wounding vs. non-wounding (control) (D), and 4 °C vs. 25 °C (control) (E) on the expression of Cs4CL1 and Cs4CL2 in leaves. In the three-paired conditions, qRT-PCR was completed to obtain the expression levels of each gene at each time point for both each treatment (UV-B, wounding, and 4 °C) and each control condition. Then, the ratios of UV/white, wounding/non-wounding, and 25 °C/4 °C were calculated for each time point. The resulting ratios were used to characterize the effects of each treatment on the expression of two genes. Values represent the mean ± standard deviation (SD) from three times of experiments each with three biological replicates. The significant expression differences of each gene were evaluated by comparing *P < 0.05, **P < 0.01, and ***P < 0.001.
two enzymes. The ORFs of Cs4CL1 and Cs4CL2 were cloned into the prokaryotic expression vector Pet-SUMO and then expressed to induce their recombinant proteins in *E. coli*. The recombinant Cs4CL1 and Cs4CL2 were further purified by a His-Tag-column system (Fig. S2). A mixture of three hydroxycinnamate molecules, 4-coumaric acid, ferulic acid, and caffeic acid, was used for this catalytic test. HPLC analysis showed that the enzymatic incubations of both the recombinant Cs4CL1 and Cs4CL2 with the three compounds produced new compound peaks (Fig. 3A and B), while these compounds were not formed from all negative controls, such as boiled recombinant Cs4CL1 and Cs4CL2 (Fig. 3C). Further UPLC-MS/MS indicated that three peaks had the mass to charge ratios of 912, 930, and 944 [m/z]−, which were corresponding to coumaryl-CoA (CM-CoA), caffeoyl-CoA (CF-CoA), and feruloyl-CoA (FR-CoA). MS/MS fragmentation analysis identified two main feature fragments 428 and 407 m/z from 912 [m/z]−, 428 and 423 m/z from 930 [m/z]−, and 428 and 437 m/z from 944 [m/z]− (Fig. 3D), which are the main characteristic fingerprints of CM-CoA, FR-CoA, and CF-CoA (Lavhale et al. 2021; Liu et al. 2017). Based on these main fragments, the three peaks were annotated to be CM-CoA, FR-CoA, and CF-CoA from 4-coumaric acid, caffeic acid, and ferulic acid, respectively. These data demonstrated that two enzymes used these three hydroxycinnamate molecules as substrates.

**Fig. 3** HPLC–MS profiles showing the conversion of 4-coumaric acid, ferulic acid, and caffeic acid to their corresponding products by recombinant Cs4CL1 and Cs4CL2. A HPLC peak profiles show that the recombinant Cs4CL1 converts 4-coumaric acid (CM), ferulic acid (FR), and caffeic acid (CF) to 4-coumaric acid-CoA (CM-coa), ferulic acid-CoA (FR-coa), and caffeic acid-CoA (CF-coa), respectively. B HPLC peak profiles show that the recombinant Cs4CL2 converts 4-coumaric acid (CM), ferulic acid (FR), and caffeic acid (CF) to 4-coumaric acid-CoA (CM-coa), ferulic acid-CoA (FR-coa), and caffeic acid-CoA (CF-coa), respectively. C HPLC peak profiles show that the boiled recombinant Cs4CL1 and Cs4CL2 proteins cannot convert 4-coumaric acid, ferulic acid, and caffeic acid to their corresponding CoA products. Crude protein extracts of *E. coli* harboring a control vector and of IPTG-lacking *E. coli* harboring recombinant Cs4CL1 or Cs4CL2 vector were also used for enzymatic assays. The results were the same as those of the boiled Cs4CL1 and Cs4CL2. D HPLC peak profiles show 4-coumaric acid, ferulic acid, and caffeic acid standards. E Tandem mass fragment profiles show main characteristic fragments of 4-coumaric acid-CoA (CM-coa, 912.0), ferulic acid-CoA (FR-coa, 930.0), and caffeic acid-CoA (CF-coa, 944.0).
Optimization of pH and temperature values and kinetic analysis

The reaction temperatures and pH values were tested from 0 to 60 °C and from pH 5.5 to 8.0, respectively. The substrate used for optimization was 4-coumaric acid. The resulting data showed that the temperature optimum of two enzymes was 50 °C (Fig. S3). The optima of pH values for Cs4CL1 and Cs4CL2 were pH6.5 and pH7.5, respectively (Fig. S3).

The kinetics of Cs4CL1 and Cs4CL2 were characterized in the condition of optimum pH and temperature. Three substrates tested were 4-coumaric (CM), ferulic (FR) and caffeic (CF) acids. Double reciprocal plots established with different concentrations showed that the catalysis of Cs4CL1 and Cs4CL2 followed the Michaelis–Menten kinetics (Fig. S4). The Km values of recombinant Cs4CL1 and Cs4CL2 to these three substrates were estimated to be 8.86–19.99 (µM) and 12.61–16.08 (µM), respectively (Table 1). The Km values indicated that the two enzymes had the highest affinity to CF, followed by FR, and then CM. In addition, the Vmax values of Cs4CL1 and Cs4CL2 were 0.54–1.0 and 0.55–1.0 mmol/min, g. The ratios of Vmax/Km (Kcat) were 57–60 for CsCL1 and 43.8–62.2 for CsCL2 (Table 1). The Kcat values indicated that the orders of the turnover rates were CF, CM, and FR for Cs4CL1 and CM, FR, and CF for Cs4CL2.

Subcellular localization of Cs4CL1 and Cs4CL2

Cs4CL1 and Cs4CL2 were fused at N-terminal of EGFP in the vector PGWB5 for subcellular localization analysis. In addition, EGFP and an endoplasmic reticulum (ER) marker protein were used as cytosolic and non-cytosolic localization controls. Confocal microscopic observation was carried out to compare the sub-localization of EGFP, ER-marker-EGFP, Cs4CL1-EGFP and Cs4CL2-EGFP. Multiple leaf samples of N. benthamiana infected by Agrobacterium were examined carefully and greenish florescence of GFP was photographed. The resulting images showed that EGFP alone was localized in the cytosol, while the fused ER-Marker-EGFP was barely observed in the cytosol (Fig. 4). In comparison with the localization of the two controls, the localizations of greenish florescence from both the Cs4CL1-EGFP and Cs4CL2-EGFP fusion proteins were similar to that of EGFP alone (Fig. 4). These data indicated that the Cs4CL1-EGFP and Cs4CL2-EGFP fusion proteins were mainly localized in the cytosol of epidermal cells (Fig. 4).

Effects of overexpression of both Cs4CL1 and Cs4CL2 on lignin, chlorogenic acids, and three flavonols in transgenic tobacco plants

The ORFs of both Cs4CL1 and Cs4CL2 were overexpressed in tobacco plants. Three transgenic lines for each gene (Fig. 5A) was selected for analysis of lignin, flavonoids, and other polyphenolic compounds. In comparison with wild type plants, no phenotypic alterations were observed in transgenic plants (Fig. 5A). The estimation of the total lignin from all tissues showed that the overexpression of both Cs4CL1 and Cs4CL2 increased the contents of lignin in transgenic plants compared with wild type plants (Fig. 5B). UPLC-MS analysis identified quercetin-3-O-rutinoside (rutin), kaempferol-3-O-rutinoside, kaempferol-3-O-glucoside, and chlorogenic acid (Jiang et al. 2013b). Compared to the contents in wild type plants, the

| Isoform | Substrate | Km (µM) | Vmax (mmol/min g) | Vmax/Km (µM) |
|---------|-----------|---------|------------------|---------------|
| Cs4CL1  | 4-Coumarate | 19.99   | 1.00             | 50            |
|         | Ferulate   | 15.98   | 0.75             | 47            |
|         | Caffeate   | 8.86    | 0.54             | 60            |
| Cs4CL2  | 4-Coumarate | 16.08   | 1.00             | 62.2          |
|         | Ferulate   | 14.64   | 0.75             | 52.6          |
|         | Caffeate   | 12.61   | 0.55             | 43.8          |

Table 1 Kinetic values of Cs4CL1 and Cs4CL2 to three substrates

Fig. 4 Subcellular localization of Cs4CL1 and Cs4CL2. Confocal microscopy images show the subcellular localizations of GFP alone (cytosolic localization control), endoplasmic reticulum (ER)-localized-marker (non-cytosolic control), 4CL1-EGFP, and 4CL2-EGFP in epidermal cells of leaves of Nicotiana benthamiana.

Springer
contents of chlorogenic acid were increased 2.5 to 3.0 folds and 2.8–3.5 folds in Cs4CL1 and Cs4CL2 transgenic plants, respectively (Fig. 5D). The contents of rutin were slightly or obviously increased in seedlings of three Cs4CL1 transgenic plants and were significantly increased in seedlings of three Cs4CL2 transgenic ones. Due to the lack of standard samples, the peak areas of kaemferol-3-O-rutinoside and kaempferol-3-O-glucoside were used to compare their levels in transgenic and wild type seedlings. The resulting data indicated that the level of kaemferol-3-O-glucoside were significantly or slightly higher in all transgenic plants of Cs4CL1 and Cs4CL2 than in wild type ones (Fig. 5E). The level of kaemferol-3-O-rutinoside was significantly increased in Cs4CL2 transgenic plants (Fig. 5F).

Discussion

This study shows two 4CL isomers that differentially involve in the biosynthesis of tea polyphenols. Tea plant is highly rich in plant phenylpropanoids (Dixon et al. 2013; Weisshaar and Jenkins 1998; Jiang et al. 2013b). Although the late pathway genes toward flavan-3-ols and anthocyanins have gained intensive studied in the characterization of tea polyphenol biosynthesis (Liu et al. 2012; Wang et al. 2019), the entry pathway genes from phenylalanine to 4-coumaroyl CoA (Fig. 1A) remain for functional characterization to understand their roles in the biosynthesis of flavonoids, lignin, and other phenolic acids such as chlorogenic acid reported here. 4CL is a key enzyme that catalyzes one key entry step toward the biosynthesis of both lignin and flavonoids. It has gained functional characterization in other plants (Cukovica et al. 2001). For example, in Arabidopsis, two subgroups of 4CL isoforms have been classified to involve in the formation of phenylpropanoids. One subgroup is classified as class I, such as At4CL1 that is preferably involved in the biosynthesis of lignin. The other subgroup is classified as class II, such as At4CL3 that is relatively involved in the biosynthesis of flavonoids. These two classes are further characterized to be differentiation in gene expression patterns in different tissues, to have different kinetics to substrates, and to variously respond to environmental stresses (Hamberger and Hahlbrock 2004; Li et al. 2015a). To date, homologs in tea plant remain for functional characterization to understand their metabolic roles involved in the lignin and flavonoids pathways. To understand the functions of 4CL homologs in tea plant, we started to use molecular cloning technologies to clone and characterize two 4CL isomers, Cs4CL1 and Cs4CL2, before the tea genome sequence was reported (Wei et al 2018). Cs4CL1 is a class I member (Fig. 1B) highly expressed in roots and stems (Fig. 2B), which are rich in lignin, while Cs4CL2 is a class II member (Fig. 1B) highly expressed in young leaves (Fig. 2B), which are rich in flavonoids. Although, to date, no tea mutants are available and tea genetic transformation is impossible, we used in vitro enzyme assays and tobacco plants to show that the two isomers are involved in the biosynthesis of tea phenylpropanoids and their involvement is characterized by their metabolic preferences. In addition, to gain a better understanding of 4CL isomers in tea plants, we further mined the genome sequences after the sequencing was reported by our institute in 2018 (Wei et al. 2018). Based on the annotation of the tea genome, 14 sequences were annotated to be potential 4CL homologs, two of which were Cs4CL1 and Cs4CL2 (Table S2). Although 12 others have not be functionally characterized, these findings suggest that the 14 paralogs may be associated with the high production of plant phenylpropanoids in tea plants and further functional characterization of 12 other members is necessary to characterize the entry steps of the tea phenylpropanoid pathway.

In vitro assay, expression induction, and transgenic analyses of Cs4CL1 and Cs4CL2 support differentially metabolic contributions of various 4CL isomers in plants. To date, the substrate preference and metabolic contributions of 4CL homologs from different plants has gained appropriately characterization. In Arabidopsis thaliana, four At4CL isomers were functionally characterized by both enzymatic and genetic analyses. At4CL1 and At4CL2 were reported to prefer to catalyze both 4-coumaric and caffeic acids involved in the biosynthesis of lignin. At4CL1 was found to account for the major catalytic activity in plants, At4CL3 was reported to mainly prefer to catalyze 4-coumaric acid to contribute the biosynthesis of flavonoids, and At4CL4 was shown to only moderately involve in the biosynthesis of lignin (Lee et al. 1997; Li et al. 2015b). Four 4CLs were cloned from soybean (Glycine max), namely Gm4CL1, Gm4CL2, Gm4CL3, and Gm4CL4. Recombinant Gm4CL1 and Gm4CL2 showed higher affinities to multiple substrates such as 4-coumaric, caffeic, ferulic, and sinapic acids, while recombinant Gm4CL3 and Gm4CL4 mainly preferred to 4-coumaric and caffeic acids (Lindermayr et al. 2002a). Different enzyme kinetics of Pt4CL1 (class II) and Pt4CL2 (class I) from aspen were also characterized by in vitro assays. Pt4CL1 was shown to prefer to catalyze 4-coumaric acid, while Pt4CL2 showed a higher affinity to caffeic acid (Hu et al. 1998; Harding et al. 2002). Herein, our data showed that although both recombinant Cs4CL1 and Cs4CL2 had the same trend of Km values in catalyzing 4-coumaric acid, ferulic acid and caffeic acid (Table 1), the Kcat values of the enzymes were different. Cs4CL1 had the highest Kcat value to caffeic acid, while Cs4CL2 had the highest Kcat value to 4-coumaric acid. To further understand the potential difference of the metabolic contributions of Cs4CL1 and Cs4CL2 in plants, we overexpressed them in tobacco plants. Although the overexpression of each gene
alone increased the contents of lignin and chlorogenic acid in tobacco plants (Fig. 5B, D), their differential effects on the levels of three flavonols were revealed by metabolic analysis. The overexpression of Cs4CL2 could consistently increase the contents of rutin and kaempferol-3-O-rutinoside in different transgenic lines, while the overexpression of Cs4CL1 did not (Fig. 5C and F). This metabolic differentiation might result from their catalytic turnover difference that Cs4CL1 preferred to caffeic acid, while Cs4CL2 preferred to 4-coumaric acid (Table 1). In addition, to understand the potential activities of the two isomers in green tea, we analyzed regulatory elements of promoters and then treated plants with UV-B irradiation, wounding, and 4 °C temperature. The Cs4CL2 promoter has five AC-elements, while the Cs4CL1 promoter has one. As well understood, the AC-element plays important roles in responses of gene expression to UV-B irradiation. MYB transcription factors that respond to UV-B irradiation bind it and activate gene expression to lead to the increase of flavonoids in plant leaves (Qian et al. 2020; Zhao et al. 2020). In our experiments, the UV-B treatment indicated that the expression of Cs4CL2 but not Cs4CL1 was quickly increased in leaves (Fig. 2C). In contrast, the Cs4CL1 expression responded to a mechanical wounding quicker and more obvious than the Cs4CL2 (Fig. 2D). The responses of two genes to the 4 °C treatment were also different (Fig. 2E). These data suggest that on the one hand, both Cs4CL1 and Cs4CL2 contribute the biosynthesis of lignin and chlorogenic acid; on the other hand, two differentiate each other in the biosynthesis flavonoids and in responses of different stresses.

In conclusion, tea plant is an important beverage crop due to its richness in flavonoids. Our findings provide insight into the differential roles of two 4CL isomers in the entry pathway steps of tea flavonoid biosynthesis. These findings are anticipated to enhance the metabolic engineering for value-increased beverage products in the future.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s11103-022-01269-6.

**Author contributions** ML, LG, and TX conceived the experiments. ML, LG, YW, XJ, and YL performed the experiments. ML and DX prepared the manuscript. All authors read and approved the manuscript.

**Funding** Funding was provided by National Science Foundation of China (Grant Nos. 31870677 and 31870676).

**Data availability** Enquiries about data availability should be directed to the authors.

**Declarations**

**Competing interests** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**References**

Atmani D, Chaher N, Atmani D, Berboucha M, Debache N, Boudaoud H (2009) Flavonoids in human health: from structure to biological activity. Curr Nutr Food Sci 5(4):225–237

Bell-Lelong DA, Cusumano JC, Meyer K, Chapple C (1997) Cinnamate-4-hydroxylase expression in Arabidopsis (regulation in response to development and the environment). Plant Physiol 113(3):729–738

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54(1):519–546

Camell EL, Towers GHN (1969) Phenylalanine and tyrosine ammonia-lyase activity in sporobolomyces roseus. Phytochemistry 8(8):1407–1413. https://doi.org/10.1016/S0031-9422(00)85907-5

Camell EL, Towers GN (1973) Phenylalanine ammonia lyase. Phytochemistry 12(5):961–973

Chen HY, Babst BA, Nyamdari B, Hu H, Sykes R, Davis MF, Harding SA, Tsai CJ (2014) Ectopic expression of a loblolly pine class II 4-coumarate:CoA ligase alters soluble phenylpropanoid metabolism but not lignin biosynthesis in populus. Plant Cell Physiol 55(9):1669–1678

Costa MA, Bedgar DL, Moinuddin SG, Kim K-W, Cardenas CL, Cochrane FC, Shockey JM, Helms OL, Amakura Y, Takahashi H (2005) Characterization in vitro and in vivo of the putative multigene 4-coumarate: CoA ligase network in Arabidopsis: syringyl lignin and sinapate/sinapyl alcohol derivative formation. Phytochemistry 66(17):2072–2091

Cukovic D, Ehlting J, Van Ziffle JA, Douglas CJ (2001) Structure and evolution of 4-coumarate: coenzyme A ligase (4CL) gene families. Biol Chem 382(4):645–654

Dixon RA, Liu C, Ji HJ (2013) Metabolic engineering of anthocyanins and condensed tannins in plants. Curr Opin Biotechnol 24(2):329–335

Ehlting J, Büttner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E (1999a) Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. Plant J 19(1):9–20

Ehlting J, Büttner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E (1999b) Three 4-coumarate: coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. Plant J 19(1):9–20

Grace SC, Logan BA (2000) Energy dissipation and radical scavenging by the plant phenylpropanoid pathway. Philos Trans R Soc Lond B 355(1402):1499–1510

Grima-Pettenati J, Goffner D (1999) Lignin genetic engineering revisited. Plant Sci 145(2):51–65
Halpin C, Knight ME, Foxon MM, Boudet AM, Boon Gui J, Shen J, Li L (2011) Functional characterization of evolutionarily divergent 4-coumarate: coenzyme A ligase in rice. Plant Physiol 157(2):574–586

Hanpin C, Knight ME, Foxon GA, Campbell MM, Boudet AM, Boon JJ, Chabbert B, Schuch W (1994) Manipulation of lignin quality by downregulation of cinnamyl alcohol dehydrogenase. Plant J 6(3):339–350

Hamberger B, Halbrook K (2004) The 4-coumarate: CoA ligase gene family in Arabidopsis thaliana comprises one rare, sinapate-activating and three commonly occurring isoenzymes. Proc Natl Acad Sci 101(7):2209–2214

Harding SA, Leshechkin J, Chiang VL, Tsai CJ (2002) Differential substrate inhibition couples kinetically distinct 4-coumarate: coenzyme A ligases with spatially distinct metabolic roles in quaking aspen. Plant Physiol 128(2):428–438

Hu W-J, Kawaoka A, Tsai C-J, Lung J, Osakabe K, Ebinuma H, Chiang VL (1998) Compartmentalized expression of two structurally and functionally distinct 4-coumarate: CoA ligase genes in aspen (Populus tremuloides). Proc Natl Acad Sci 95(9):5407–5412

Jiang X, Liu Y, Li W, Zhao L, Meng F, Wang Y, Tan H, Yang H, Wei C, Wan X (2013a) Tissue-specific, development-dependent phenolic compounds accumulation profile and gene expression pattern in tea plant (Camellia sinensis). PLoS ONE 8(4):e62315

Jiang X, Liu Y, Li W, Zhao L, Meng F, Wang Y, Tan H, Yang H, Wei C, Wan X, Gao L, Xia T (2013b) Tissue-specific, development-dependent phenolic compounds accumulation profile and gene expression pattern in tea plant [Camellia sinensis]. PLoS ONE 8(4):e62315

Jiang X, Liu Y, Wu Y, Tan H, Meng F, Wang Y, Li M, Zhao L, Liu L, Qian Y (2015) Analysis of accumulation patterns and preliminary study on the condensation mechanism of proanthocyanidins in the tea plant [Camellia sinensis]. Sci Rep 5:8742

Jinshan G, Junhui S, Laigeng L (2011) Functional characterization of evolutionarily divergent 4-coumarate: coenzyme A ligase in rice. Plant Physiol 157(2):574–586

Lavhale SG, Kalunke RM, Giri AP (2018) Structural, functional and evolutionary diversity of 4-coumarate-CoA ligase in plants. Planta 248(5):1063–1078

Lavhale SG, Joshi RS, Kumar Y, Giri A (2021) Functional insights into two Ocimum kilimandscharicum 4-coumarate-CoA ligases involved in phenylpropanoid biosynthesis. Int J Biol Macromol 181:202–210

Lee D, Mayor K, Chapple C, Douglas C (1997) Down-regulation of 4-coumarate: CoA ligase (4CL) in Arabidopsis: effect on lignin composition and implications for the control of monolignol biosynthesis. Plant Cell 9(1):985–981

Lei ZY, Zhao P, Cao MJ, Cui R, Chen X, Xiong LZ, Zhang QF, Oliver DJ, Xiang CB (2007) High-throughput binary vectors for plant gene function analysis. J Integr Plant Biol 49(4):556–567

Li Y, Im Kim J, Pysh L, Chapple C (2015a) Four isoforms of Arabidopsis 4-coumarate: CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. Plant Physiol 169(4):2409–2421

Li Y, Kim JI, Pysh L, Chapple C (2015b) Four isoforms of Arabidopsis thaliana 4-coumarate: CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. Plant Physiol. https://doi.org/10.1104/pp.15.00838

Li M, Li Y, Guo L, Gong N, Pang Y, Jiang W, Liu Y, Jiang X, Zhao L, Wang Y (2017) Functional characterization of tea (Camellia sinensis) MYB4a transcription factor using an integrative approach. Front Plant Sci 8:943

Lindermayr C, Möllers B, Fliegmann J, Uhlmann A, Lottspeich F, Meirnberg H, Ebel J (2002a) Divergent members of a soybean (Glycine max L.) 4-coumarate:coenzyme A ligase gene family. Eur J Biochem 269(4):1304–1315

Lindermayr C, Möllers B, Fliegmann J, Uhlmann A, Lottspeich F, Meirnberg H, Ebel J (2002b) Divergent members of a soybean (Glycine max L.) 4-coumarate: coenzyme A ligase gene family: primary structures, catalytic properties, and differential expression. Eur J Biochem 269(4):1304–1315

Liu Y, Gao L, Liu L, Yang Q, Lu Z, Nie Z, Wang Y, Xia T (2012) Purification and characterization of a novel galloyltransferase involved in catechin galloylation in the tea plant (Camellia sinensis). J Biol Chem 287(53):44406–44417

Liu T, Yao R, Zhao Y, Xu S, Huang C, Luo J, Kong L (2017) Cloning, functional characterization and site-directed mutagenesis of 4-coumarate: coenzyme A ligase (4CL) involved in coumarin biosynthesis in Pseudocedrus pruerauritum Dunn. Front Plant Sci 8:4

Liu Y, Jiang H, Zhao Y, Li X, Dai X, Zhuang J, Zhu M, Jiang X, Wang P, Gao L (2019) Three Camellia sinensis glutathione S-transfereases are involved in the storage of anthocyanins, flavonols, and proanthocyanidins. Planta 250:1163

Lu H, Zhao Y, Jiang X (2004) Stable and specific expression of 4-coumarate:coenzyme A ligase gene (4CL1) driven by the xylem-specific Pto4CL1 promoter in the transgenic tobacco. Biootechnol Lett 26(14):1147–1152

Mierziak J, Kostyn K, Kulma A (2014) Flvanoids as important molecules of plant interactions with the environment. Molecules 19(10):16240–16265

Pang Y, Pecl BJ, Wright E, Wang ZY, Dixon RA (2007) Early steps in proanthocyanidin biosynthesis in the model legume Medicago truncatula. Plant Physiol 145(3):601–615

Piqemual J, Lapierre C, Myton K, O’Connell A, Schuch WR, Grima-Pettenati J, Boudet AM (2002) Down-regulation of cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. Plant J 13(1):71–83

Qian C, Chen Z, Liu Q, Mao W, Chen Y, Tian W, Liu Y, Han J, Ouyang X, Huang X (2020) Coordinated transcriptional regulation by the UV-B photoreceptor and multiple transcription factors for plant UV-B responses. Mol Plant 13(5):777–792. https://doi.org/10.1016.j.molp. 2020.02.015

Rui S, Ying-Hsuan S, Quanzi L, Steffen H, Ronald S, 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(1):144–163

Schilmiller AL, Stout J, Weng JK, Humphreys J, Ruegger MO, Chapple C (2009) Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis. Plant J 60(5):771–782

Shi R, Sun Y-H, 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(1):144–163

Soltani BM, Ehling J, Hamberger B, Douglas C (2006) Multiple cis-regulatory elements regulate distinct and complex patterns of developmental and wound-induced expression of Arabidopsis thaliana 4CL gene family members. Planta 224(5):1226–1238

Sueoka N, Suganuma M, Sueoka E, Okabe S, Matsuyama S, Imai K, Nakachi K, Fujiki H (2001) A new function of green tea: prevention of lifestyle-related diseases. Ann N Y Acad Sci 928(1):274–280

Sun H, Li Y, Feng S, Zou W, Guo K, Fan C, Si S, Peng L (2017) Functional characterization of a novel galloyltransferase involved in catechin galloylation in the tea plant [Camellia sinensis]. J Biol Chem 287(53):44406–44417

Ververidis F, Trantas E, Douglas C, Vollmer G, Kretzschmar G, Panopoulos N (2007) Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. Biotechnol J 2(10):1214–1234

Vogt T (2010) Phenylpropanoid biosynthesis. Mol Plant 3(1):2–20. https://doi.org/10.1093/mp/ssp106
Wang Y-S, Xu Y-J, Gao L-P, Yu O, Wang X-Z, He X-J, Jiang X-L, Liu Y-J, Xia T (2014) Functional analysis of flavonoid 3′, 5′-hydroxylase from tea plant (*Camellia sinensis*): critical role in the accumulation of catechins. BMC Plant Biol 14(1):347

Wang P, Zhang L, Jiang X, Dai X, Xu L, Li T, Xing D, Li Y, Li M, Gao L (2018) Evolutionary and functional characterization of leucoanthocyanidin reductases from *Camellia sinensis*. Planta 247(1):139–154

Wang P, Liu Y, Zhang L, Wang W, Hou H, Zhao Y, Jiang X, Yu J, Tan H, Wang Y (2019) Functional demonstration of plant flavonoid carbocations proposed to be involved in the biosynthesis of proanthocyanidins. Plant J 101:18–36

Wanner LA, Li G, Ware D, Somssich IE, Davis KR (1995) The phenylalanine ammonia-lyase gene family in Arabidopsis thaliana. Plant Mol Biol 27(2):327–338

Wei C, Yang H, Wang S, Zhao J, Liu C, Gao L, Xia E, Lu Y, Tai Y, She G, Sun J, Cao H, Tong W, Gao Q, Li Y, Deng W, Jiang X, Wang W, Chen Q, Zhang S, Li H, Wu J, Wang P, Li P, Shi C, Zheng F, Jiang J, Huang B, Shan D, Shi M, Fang C, Yue Y, Li F, Li D, Wei S, Han B, Jiang C, Yin Y, Xia T, Zhang Z, Bennetzen JL, Zhao S, Wan X (2018) Draft genome sequence of *Camellia sinensis* var. sinensis provides insights into the evolution of the tea genome and tea quality. Proc Natl Acad Sci 115(18):E4151–E4158

Weissshaar B, Jenkins GI (1998) Phenylpropanoid biosynthesis and its regulation. Curr Opin Plant Biol 1(3):251–257

Whetten R, Sederoff R (1995) Lignin biosynthesis. Plant Cell 7(7):1001

Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol 5(3):218–223

Xia E-H, Zhang H-B, Sheng J, Li K, Zhang Q-J, Kim C, Zhang Y, Liu Y, Zhu T, Li W, Huang H, Tong Y, Nan H, Shi C, Shi C, Jiang J-J, Mao S-Y, Jiao J-Y, Zhang D, Zhao Y, Zhao Y-J, Zhang L-P, Liu Y-L, Liu B-Y, Yu Y, Shao S-F, Ni D-J, Eichler EE, Gao L-Z (2017) The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis. Mol Plant 10(6):866–877

Zhao L, Gao L, Wang H, Chen X, Wang Y, Yang H, Wei C, Wan X, Tao X (2013) The R2R3-MYB, bHLH, WD40, and related transcription factors in flavonoid biosynthesis. Funct Integr Genomics 13(1):75–98

Zhao X, Wang P, Li M, Wang Y, Jiang X, Cui L, Qian Y, Zhuang J, Gao L, Xia T (2017) Functional characterization of a new tea (*Camellia sinensis*) flavonoid glycosyltransferase. J Agric Food Chem 65(10):2074–2083

Zhao B, Wang L, Pang S, Jia Z, Wang L, Li W, Jin B (2020) UV-B promotes flavonoid synthesis in *Ginkgo biloba* leaves. Ind Crop Prod 151:112483. https://doi.org/10.1016/j.indcrop.2020.112483

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.