Insights into acylation mechanisms: co-expression of serine carboxypeptidase-like acyltransferases and their non-catalytic companion paralogs

Shengbo Yao1,†, Yajun Liu2,†, Juhua Zhuang1, Yue Zhao1, Xinlong Dai1, Changjuan Jiang2, Zhihui Wang1, Xiaolan Jiang1, Shuxiang Zhang2, Yumei Qian3, Yuling Tai2, Yunsheng Wang2, De-Yu Xie4,* and Tao Xia1,*

1State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, 230036, Anhui, China, 2School of Life Science, Anhui Agricultural University, Hefei, 230036, Anhui, China, 3School of Biological and Food Engineering, Suzhou University, Suzhou, 234000, Anhui, China, and 4Department of Plant and Microbial Biology, North Carolina State University, Raleigh, North Carolina, 27695, USA

SUMMARY

Serine carboxypeptidase-like acyltransferases (SCPL-ATs) play a vital role in the diversification of plant metabolites. Galloylated flavan-3-ols highly accumulate in tea (Camellia sinensis), grape (Vitis vinifera), and persimmon (Diospyros kaki). To date, the biosynthetic mechanism of these compounds remains unknown. Herein, we report that two SCPL-AT paralogs are involved in galloylation of flavan-3-ols: CsSCPL4, which contains the conserved catalytic triad S-D-H, and CsSCPL5, which has the alternative triad T-D-Y. Integrated data from transgenic plants, recombinant enzymes, and gene mutations showed that CsSCPL4 is a catalytic acyltransferase, while CsSCPL5 is a non-catalytic companion paralog (NCCP). Co-expression of CsSCPL4 and CsSCPL5 is likely responsible for the galloylation. Furthermore, pull-down and co-immunoprecipitation assays showed that CsSCPL4 and CsSCPL5 interact, increasing protein stability and promoting post-translational processing. Moreover, phylogenetic analyses revealed that their homologs co-exist in galloylated flavan-3-ol- or hydrolyzable tannin-rich plant species. Enzymatic assays further revealed the necessity of co-expression of those homologs for acyltransferase activity. Evolution analysis revealed that the mutations of the CsSCPL5 catalytic residues may have taken place about 10 million years ago. These findings show that the co-expression of SCPL-ATs and their NCCPs contributes to the acylation of flavan-3-ols in the plant kingdom.

Keywords: tea, serine carboxypeptidase-like acyltransferase, flavan-3-ol gallates, co-expression, galloylation, non-catalytic companion paralog.
characterized SCPL-ATs have been grouped in clade IA with a GDSYS diagnostic pentapeptide motif sequence (Mugford & Milkowski, 2012).

The SCPL-ATs have been shown to catalyze the acylation of various substrates, such as terpenoids, flavonoids, and alkaloids. These known SCPL-ATs include a sinapoylgucose:malate acyltransferase (SMT) (Lehfeldt et al., 2000), a sinapoylgucose:choline acyltransferase (SCT) (Shirley et al., 2001), two sinapoylgucose:sinapoylgucose acyltransferases (SSTs) (Fraser et al., 2007), a sinapoylgucose:anthocyanin acyltransferase (SAT) (Fraser et al., 2007), and a phenylglucose:flavonol acyltransferase (FP) (Tohge et al., 2016) from A. thaliana; an isobutyrylgucose:isobutyrylgucose acyltransferase (GAT) from Lycopersicon pennelli (Li et al., 1999); and an N-methyl anthraniloylgucose:avenacin acyltransferase (SAD7) from Avena strigosa (Mugford et al., 2009); an indoleacetylglucose:inositol acyltransferase (IAA-AT) from Onzya sativa (Clarkowska et al., 2018); a phenyllactylglucose:tropine acyltransferase (LS) from Atropa belladonna (Qiu et al., 2020); and a chlorogenic acid:caffeic acid acyltransferase (CAS) from Echinacea purpurea (Fu et al., 2021). In addition, a few more putative SCPL-ATs have been shown to transfer acyl groups, such as a galloylgucose:epicatechin acyltransferase (EC-GT) catalyzing 3-O-galloylation of epicatechin (EC) from Camellia sinensis (Liu et al., 2012) and two p-hydroxybenzoylgucose:anthocyanin acyltransferases (AA7G-AT and AA7GBG-AT) from Delphinium grandiflorum (Nishizaki et al., 2013). However, the majority of SCPL-AT candidates in most sequenced plants remain to be functionally determined to characterize whether and how they are involved in the acylation of metabolites in plants. For example, four SCPL homologs were cloned from galloylated catechin-rich fruits of grape (Vitis vinifera) and persimmon ( Diospyros kaki); nevertheless, to date, their function has not yet been elucidated (Akagi et al., 2009; Bontpart et al., 2018).

Published studies suggest that most SCPL-ATs undergo complex post-translational processing and secretion into organelles (mostly vacuoles) (Bontpart et al., 2015; Mugford & Milkowski, 2012). This brings challenges to the functional identification of SCPL-ATs. The precursor protein encoded by SCPL needs to undergo a series of post-translational processing steps to form a catalytically active mature protein, which includes removal of the N-terminal signal peptide, disulfide bond formation, glycosylation modifications, proteolytic cleavage of the internal spacer peptide, and the formation of heterodimers (Mugford et al., 2008; Mugford & Milkowski, 2012). However, some SCPL-ATs, such as AtSMT, do not go through this process (Hase et al., 2002). The N-terminal signal peptide, which guides the protein to be processed in the secretory pathway toward vacuoles, is present in the characterized SCPL-AT sequences in plants according to bioinformatic prediction (Fu et al., 2021; Hause et al., 2002; Mugford et al., 2013).

Flavan-3-ol gallates include compounds such as epigallocatechin gallate (EGCG) and epicatechin gallate (ECG), which are galloylated products of flavan-3-ol, and are a main class of flavan-3-ol derivatives (Figure 1a). These compounds are rich in numerous crops, such as tea (Camellia sinensis), grape, and persimmon (Akagi et al., 2010; Jiang et al., 2013; Obreque-Slier et al., 2010). Furthermore, EGCG and ECG are regularly consumed from tea, wine (Rousserie et al., 2019), and fruits, such as grape and persimmon. These compounds give fruits and drinks a bitter and astringent taste (Zhuang et al., 2020). More importantly, galloylated flavan-3-ols have multiple benefits to human health, such as anti-inflammatory (Fechtner et al., 2017), antioxidant (Lambert & Elias, 2010), mitochondrial modulation (Oliveira et al., 2016), anti-metabolic syndrome (Legeay et al., 2015), anti-neurodegeneration (Singh et al., 2016), anti-cancer (Zhao et al., 2021), and anti-COVID-19 (Zhu & Xie, 2020) effects. Accordingly, much research on the secondary metabolism of tea, grape, persimmon, and other crops has focused on the biosynthesis of these flavan-3-ol gallates.

Tea buds and leaves are rich in EGCG and ECG, which account for more than 10% of the dry mass of these vegetative tissues (Jiang et al., 2013) and are distributed in different subcellular locations rather than being confined to vacuoles (Liu et al., 2009). Although the mechanisms by which these tissues produce EGCG and ECG in large quantities remain to be elucidated, we previously reported an EC-GT purified from tea leaves and demonstrated that it effectively catalyzed the formation of EGCG and ECG from EG and EC, respectively, in the presence of 1-O-β-glucogallin (1G) (Liu et al., 2012). A previous study reported that three recombinant CsSCPL proteins expressed in Escherichia coli are able to catalyze the formation of EGCG and ECG using 1,2,3,4,6-penta-O-galloyl-β-D-glucose (PGG) as galloyl donor (Ahmad et al., 2020). In addition, we recently reported that a plant tannase degalloylated EGCG and ECG to EGC and EC, respectively (Dai et al., 2020), indicating that the accumulation of high contents of EGCG and ECG is controlled by reversible reactions (Figure 1a). Nevertheless, how tea leaves regulate such dynamic reversible galloylation and degalloylation and maintain high production of EGCG and ECG is unknown.

In order to find out which gene encodes EC-GT, we purified the enzyme from tea leaves according to a previously described method (Liu et al., 2012). The purified enzyme was used for protein identification, and then we obtained two SCPL paralogs from this enzyme, named CsSCPL4 and CsSCPL5. Herein, we determined that the co-expression of CsSCPL4 and CsSCPL5 contributes to their acyltransferase activity in vitro and in vivo. CsSCPL4 is an active enzyme,
while CsSCPL5 is non-active, but CsSCPL4 activity depends on CsSCPL5. The pull-down and co-immunoprecipitation (Co-IP) assays effectively demonstrated the interaction between CsSCPL4 and CsSCPL5 pro-peptides. Based on our findings, we name CsSCPL5 a non-catalytic companion paralog (NCCP). Homologs from grape, persimmon, and
oil tea plant (*Camellia oleifera*) are also co-expressed. Accordingly, our findings reveal a novel mechanism by which the co-expression of an active SCPL and its NCCP is likely responsible for acyltransferase activity in plants.

**RESULTS**

Two SCPL-AT paralogs of a known tea epicatechin-galloyltransferase are identified

We previously isolated an EC-GT which catalyzed the galloylation of flavan-3-ols in *C. sinensis* (Liu et al., 2012), but it was unclear which gene encoded this EC-GT. In order to obtain the key genes determining the biosynthesis of galloylated flavan-3-ols in tea plants, the native EC-GT enzyme was purified again from tea leaves by successive steps according to our previous work (Liu et al., 2012). The results of protein purification under enzyme activity monitoring showed that the most active eluate formed one single protein fraction with a molecular weight ranging from 200 to 240 kDa (Figure 1b; Figure S1). Four bands were separated by SDS-PAGE from the above protein fraction, with molecular weights of 36, 34, 22, and 20 kDa (Figure 1b). These four bands were used for protein identification, and two candidate proteins were annotated as SCPLs based on sequence alignment with a public protein database of tea plant.

We cloned these two genes and named them CsSCPL4 (GenBank: MZ462934) and CsSCPL5 (GenBank: MZ462935). The amino acid identity between the two paralogs was approximately 50.2%. The abovementioned 36- and 22-kDa bands corresponded to the predicted large and small subunits of the mature CsSCPL4 protein (Figure 1b; Figures S2 and S3). The 34- and 20-kDa bands corresponded to the predicted large and small subunits of the mature CsSCPL5 protein (Figure 1b; Figures S4 and S5).

Amino acid sequence alignment of CsSCPL4 and CsSCPL5 and two functionally known serine carboxypeptidases revealed conserved functional regions, including the signal peptides in the N-terminus, catalytic triad residues, an oxanion binding site, a pentapeptide motif, disulfide bond-forming Cys residues, large and small subunit sequences, and an excision peptide (Figure 1c). The alignment highlighted that the CsSCPL5 protein lacks the conserved catalytic triad Ser-Asp-His (S-D-H). Instead, the predicted catalytic residues of CsSCPL5 comprise Thr-Asp-Tyr (T-D-Y). Based on these features, we hypothesized that mutation of CsSCPL5 at the conserved catalytic triad would be likely to abolish its acyltransferase activity.

A protein homology search with CsSCPL4 and CsSCPL5 against the tea genome was completed to obtain 46 proteins that were annotated as SCPL proteins. These 46 protein sequences and some SCPL sequences from *A. thaliana* and other plants were used for phylogenetic analysis. Based on a previous report (Fraser et al., 2005), the resulting phylogenetic tree classified these homologs into three known clades: IA, IB, and II (Figure S6a). Those SCPL proteins classified in clade IA belong to the acyltransferase clade. The phylogenetic tree also showed that there were nine CsSCPL proteins that were specifically classified into clade IA and gathered with the SCPL proteins from grape and persimmon. Another phylogenetic tree was built with homologous sequences of the nine CsSCPL proteins. The resulting tree classified CsSCPL4 and CsSCPL5 and their homologous sequences into groups d and e of clade IA, respectively (Figure S6b).

Expression profiles of two paralogs are associated with content profiles of EGCG and ECG in tea tissues

Quantitative real-time PCR (qRT-PCR), immunoblotting, and metabolite measurement were performed to characterize the expression profiles of the two genes, the two proteins, and EGCG and ECG in different tissues. qRT-PCR data showed a simultaneous expression pattern of CsSCPL4 and CsSCPL5 in buds, the first, second, third, fourth, and fifth leaves, and young and old stems, in which the expression level was the highest in buds but low in old stems (Figure 1d). The protein expression patterns of CsSCPL4 and CsSCPL5 were examined in these organs by immunoblot analysis. Two mouse monoclonal antibodies were prepared to specifically target the pro-peptides of CsSCPL4 and CsSCPL5 (54-58 kDa) and the small subunits (20-22 kDa) of the mature proteins of CsSCPL4 and CsSCPL5, respectively (Figure 1e). The immunoblot assay results indicated that the expression trends of the small subunits of the two mature proteins were related to the trends of gene expression and flavan-3-ol gallate accumulation (Figure 1d-f). In particular, the abundance of the small subunits of the CsSCPL4 and CsSCPL5 mature proteins was higher in young leaves than in mature leaves (Figure 1e). This suggests that the proteins are more actively post-translationally processed to form functional mature proteins in young leaves than in mature leaves.

These data indicated a potential association between the expression patterns of the two genes, the post-translational processing levels of the two proteins, and the accumulation of EGCG and ECG.

Co-expression of recombinant CsSCPL4 and CsSCPL5 in tobacco leaves contributes to the galloylation of flavan-3-ols

First, we used an *E. coli* system to express recombinant CsSCPL4 and CsSCPL5 for the analysis of their catalytic activity (Figure S7). However, the purified prokaryote-expressed CsSCPL4, CsSCPL5, and co-expressed CsSCPL4 and CsSCPL5 did not exhibit catalytic activity with βG as acyl donor and EC or EGC as acyl acceptors (Figure S8).

Recently, a study reported the use of *E. coli* to express three recombinant tea CsSCPLs, and their acyltransferase activity...
activity was tested in vitro using PGG and EC or EGC as substrates (Ahmad et al., 2020). However, our results showed that CsSCPL4 and CsSCPL5 alone or co-expressed could not convert PGG and EC or EGC to ECG or EGCG (Figure S9).

Next, according to reports about the successful transient expression of a few plant SCPLs in Nicotiana benthamiana leaves for the analysis of their activity (Mugford et al., 2009; Qiu et al., 2020; Weier et al., 2008), we used agroinfiltration-based infection to introduce CsSCPL4 and CsSCPL5 into the leaves of N. benthamiana. Crude proteins were extracted and used for catalytic analysis with EGC and βG. However, neither CsSCPL4 nor CsSCPL5 could catalyze the formation of EGC (Figure 2b).

Then, based on the abovementioned protein identification results that the native EC-GT fraction was composed of CsSCPL4 and CsSCPL5, we proposed that the co-expression of the two recombinant proteins in tobacco leaves might contribute to acyltransferase activity. To test this hypothesis, CsSCPL4 and CsSCPL5 were co-infiltrated into N. benthamiana leaves (Figure 2a). CsSCPL4 and CsSCPL5 alone and vector control protein extracts were used as controls. Both ultraperformance liquid chromatography (UPLC) and UPLC-tandem mass spectrometry (MS/MS) analyses were conducted to detect metabolites from enzymatic incubations. UPLC recorded at 280 nm showed that the co-expressed CsSCPL4 and CsSCPL5 produced an obvious peak with the same retention time as EGCG, while all controls did not (Figure 2b). MS/MS analysis identified that its mass-to-charge ratio (m/z) was 457.0 [M-H]−, the same as that for EGCG (Figure S10). This co-expression experiment was repeated more than 20 times, and all experiments yielded the same results. Meanwhile, a time course analysis was carried out with the co-expressed CsSCPL4 and CsSCPL5 and substrates. The results showed that as the reaction time was increased to 5 h, the production of EGCG was increased, while the concentrations of the two substrates, EGC and βG, decreased (Figure S11).

To confirm that the co-expressed CsSCPL4 and CsSCPL5 are indeed involved in the biosynthesis of EGCG in vivo, EGC and βG were injected into leaves that transiently expressed the two genes (Figure 2c). Metabolites were extracted from substrate-injected leaves for UPLC-MS/MS analysis. The results showed that EGCG could be detected in extracts from leaves co-expressing CsSCPL4 and CsSCPL5 but not from leaves expressing CsSCPL4 or CsSCPL5 alone (Figure 2d). These results indicated that when the two genes were co-expressed, EGC and βG were catalyzed to EGCG in the infected leaves of N. benthamiana.

Last, we generated stable transgenic Nicotiana tabacum plants expressing CsSCPL4 and CsSCPL5 to substantiate the necessity of the co-expression of the two paralogs for acyltransferase activity. We used one CsSCPL5 transgenic plant to pollinate one CsSCPL4 transgenic plant to produce CsSCPL4 × CsSCPL5 hybrids (Figure 2e). The hybrid seeds were collected and then planted to screen F1 progeny by antibiotic selection and PCR analysis (Figure 2f). Crude proteins were extracted from leaves of transgenic N. tabacum plants and then incubated with EGC and βG. UPLC analysis showed that EGCG was produced by protein extracts of the hybrid N. tabacum but not by protein extracts of CsSCPL4, CsSCPL5, and vector control transgenic plants (Figure 2g). Furthermore, according to our protocol for the purification of EC-GT (Liu et al., 2012), we performed a gradual purification of recombinant proteins from the CsSCPL4 × CsSCPL5 hybrid and then tested their acylation activity. The resulting partial purification products catalyzed the formation of EGCG from EGC and βG with pKat values ranging from 7.99 to 312.51 pKat mg−1 (Figure 2h). However, the proteins purified from CsSCPL4-alone and CsSCPL5-alone transgenic leaves did not show any acyltransferase activity (Figure 2h).

Subsequently, the recombinant protein was detected in transgenic N. benthamiana leaves by immunoblot analysis. Two rabbit polyclonal antibodies were prepared to specifically target the large subunit bands of the mature CsSCPL4 and CsSCPL5 proteins, respectively (Figure 2i). The immunoblot analysis results showed that the large subunit (36 kDa) of the mature CsSCPL4 protein was detected in the transient transgenic leaves co-expressing CsSCPL4 and CsSCPL5, while the CsSCPL4 pro-peptide and the large subunit were not detected in plants expressing CsSCPL4 alone (Figure 2i). The CsSCPL5 pro-peptide (54 kDa) and the large subunit (34 kDa) were detected in both transient transgenic leaves co-expressing CsSCPL4 and CsSCPL5 and leaves transiently expressing CsSCPL5 alone (Figure 2i).

In order to further explore the roles of CsSCPL4 and CsSCPL5 in the formation of mature proteins, the C-terminus of CsSCPL4 was fused with a GFP tag and transiently expressed in N. benthamiana alone or co-expressed with CsSCPL5. The extracted protein from transgenic leaves expressing CsSCPL4-GFP alone still had no galloylation activity, while the activity of extracted co-expressed CsSCPL4-GFP and CsSCPL5 was 11.35 pKat mg−1. The immunoblot analysis results showed that the CsSCPL4-GFP pro-peptides were expressed in both tests. However, the large subunit of CsSCPL4 was detected only in transgenic leaves co-expressing CsSCPL4-GFP and CsSCPL5 (Figure 2i). This experiment directly shows that: (i) the CsSCPL4 protein expressed alone is unstable and easily degraded, and cannot be processed into mature protein; (ii) when the C-terminus of CsSCPL4 is fused with a GFP tag, this increases its stability, but the fusion protein cannot be processed into active mature protein; and (iii) CsSCPL5 increases the stability of CsSCPL4 and promotes its processing.

© 2022 The Authors, The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd., The Plant Journal, (2022), 111, 117–133
The above results suggested that the co-expression of CsSCPL4 and CsSCPL5 is crucial for their galloylation activity. Co-expression of the two genes increased the stability of pro-peptides and mature proteins and promoted the post-translational processing to form a catalytically active protein.

The catalytic role of CsSCPL4 and the non-catalytic companion role of CsSCPL5 differ between protein variants

In order to elucidate the co-expression mechanism of the two proteins, the roles of signal peptides, catalytic residues, and disulfide bonds need to be considered. For this purpose, we performed three types of amino acid alterations to create variants at the signal peptide, the catalytic triad, and the disulfide bond.

In order to confirm the necessity of two signal peptides for the catalytic activity of the enzyme, a signal peptide removal experiment was performed (Figure 3a). The results showed that the co-expressed proteins with one or both lacking the signal peptide showed no catalytic activity (Figure 3b), which indicates that the signal peptide is necessary for the catalytic activity of the enzyme.

To further verify whether the two proteins were processed at different subcellular locations, a signal peptide exchange experiment was designed. The results showed that the exchange of signal peptides does not abolish, but slightly affects the catalytic activity of the enzyme (Figure 3b). It is speculated that the signal peptides of CsSCPL4 and CsSCPL5 guide their pro-peptides to the same subcellular locations to be processed into functional enzymes. The results also showed that the signal peptides of CsSCPL4 and CsSCPL5 are both effective for enzyme activity, and the efficiency of the signal peptide of CsSCPL4 is slightly higher than that of CsSCPL5 (Figure 3a,b).

The alterations of amino acids at the catalytic triad and the disulfide bond are shown in Figure 3(c). The analysis of enzyme activities of these protein variants revealed that: (i) the positive control of CsSCPL4 and CsSCPL5 co-expression showed EGCG production; (ii) co-expression of the three CsSCPL4 variant and CsSCPL5 pairs did not result in EGCG production; (iii) in the three CsSCPL5 variant and CsSCPL4 pairs, co-expression of Cs5(T187A) and CsSCPL4 and co-expression of Cs5(T187S, Y457H) and CsSCPL4 resulted in similar activity as the positive control, and co-expression of Cs5(C293A) and CsSCPL4 did not result in EGCG formation; (iv) co-expression of the Cs4(S190T, H457Y) variant and the Cs5(T187S, Y457H) variant did not result in EGCG production as a result of the exchange of catalytic residues (Figure 3d). These data demonstrated that CsSCPL4 with the catalytic triad S-D-H is necessary for the galloylation of EGC, that CsSCPL5 variants at the triad T-D-Y do not lead to loss of galloylation, and that alteration of the disulfide bond in both CsSCPL4 and CsSCPL5 leads to loss of galloylation activity.

The expression levels of these recombinant protein variants were detected by immunoblotting with the two rabbit polyclonal antibodies against the large subunits of CsSCPL4 and CsSCPL5 (Figure 3e). The pro-peptides and the large subunits of CsSCPL4 and CsSCPL5 were detected in three co-expression combinations that showed enzyme activities, including (i) CsSCPL4 and CsSCPL5, (ii) CsSCPL4 and Cs5(T187A), and (iii) CsSCPL4 and Cs5(T187S, Y457H) (Figure 3e). This suggests that the predicted catalytic residues of CsSCPL5 are not critical for the enzymatic activity and protein expression. With respect to mutation of the catalytic residues of CsSCPL4, small amounts of the large subunit of CsSCPL4 were detected upon Cs4(S190A) and CsSCPL5 co-expression, while no corresponding protein band of CsSCPL4 was detected upon co-expression of Cs4 (S190T, H457Y) and CsSCPL5 (Figure 3e). This suggests that mutation of the catalytic residues of CsSCPL4 not only affects the activity of the enzyme but also affects the post-translational processing and stability of proteins. With respect to mutation of the cysteine residues of CsSCPL4 and CsSCPL5, small amounts of the large subunit of CsSCPL5 were detected in both variants, while the pro-peptide and the large subunit of CsSCPL4 were not.
detected (Figure 3e). It is speculated that elimination of the disulfide bond may cause abnormal processing of the two proteins and result in failure to obtain mature proteins.

In conclusion, the co-expression of these two proteins and their variants demonstrates that CsSCPL4 is catalytic but CsSCPL5 is non-catalytic in the galloylation of EGC to form EGCG. Disulfide bonds are crucial for galloylation. Based on the non-catalytic nature of CsSCPL5 in the reactions and its essential roles in galloylation, we conclude it is an NCCP of CsSCPL4.

**Figure 3.** Effects of signal peptides and amino acid changes at the catalytic triads and disulfide bonds on the acyltransferase activity of co-expressed CsSCPL4 and CsSCPL5.

(a) Diagram showing the signal peptide (SP) removal and exchange between CsSCPL4 and CsSCPL5.

(b) Acrlytransferase assay results demonstrating that the signal peptides affect the formation of epigallocatechin gallate (EGCG) from epicatechin gallate (EGC) and G. nd, not detected. The values are the mean ± SD of three biological replicates.

(c) Diagram showing amino acid changes created in the catalytic triads and disulfide bonds.

(d) pKat values showing the effects of different variants of CsSCPL4 and CsSCPL5 on their activity when co-expressed. nd, not detected. The values are the mean ± SD of three biological replicates.

(e) Immunoblot images showing the co-expression of CsSCPL4 and CsSCPL5 variants in the infected Nicotiana benthamiana leaves. Anti-Cs4-L and anti-Cs5-L correspond to the two prepared rabbit polyclonal antibodies. Cs4-L, CsSCPL4 large subunit; Cs5-L, CsSCPL5 large subunit.

Demonstration of the interaction between CsSCPL4 and CsSCPL5 via pull-down and Co-IP

Although the abovementioned results demonstrated that the co-expression of CsSCPL4 and CsSCPL5 was likely responsible for galloylation of EGC and CsSCPL5 is an NCCP, it remained to be determined whether the two proteins directly interact with each other. To answer this question, we used a pull-down assay to observe the interaction between CsSCPL4 and CsSCPL5. Four paired cDNAs, (i)
MBP and GST, (ii) GST and MBP-CsSCPL5, (iii) GST-CsSCPL5 and MBP, and (iv) GST-CsSCPL5 and MBP-CsSCPL4, were co-expressed in E. coli (Figure 4a), extracted, and then loaded onto a Glutathione resin column for affinity chromatography. The eluted proteins were analyzed by SDS-PAGE and then immunoblotted with antibodies against GST, MBP, CsSCPL4, and CsSCPL5. MBP-CsSCPL4 and GST-CsSCPL5 were detected in the eluate of co-expressed MBP-CsSCPL4 and GST-CsSCPL5, and all controls were negative. This result demonstrated that MBP-CsSCPL4 was bound to GST-CsSCPL5 (Figure 4a; Figure S12).

To further determine whether the interaction of CsSCPL4 and CsSCPL5 could occur in plant cells, a Co-IP assay was performed by co-expressing CsSCPL4-GFP and CsSCPL5 in the leaves of N. benthamiana. For the Co-IP assay, crude proteins were extracted and then immunoprecipitated with anti-GFP magnetic beads to purify CsSCPL4-GFP. The immunoblot results show that the pro-peptides and the large subunits of CsSCPL4-GFP and CsSCPL5 can be detected in the crude proteins, while only the pro-peptide bands of CsSCPL4-GFP and CsSCPL5 were detected in the purified proteins, and all controls were negative (Figure 4b). Those results indicated that the CsSCPL4-GFP and CsSCPL5 pro-peptides interact in N. benthamiana leaf cells. In conclusion, the pull-down and Co-IP results showed that CsSCPL4 and CsSCPL5 interact before processing. However, whether the two mature proteins also interact is unclear. Since the currently prepared anti-CsSCPL4 and anti-CsSCPL5 antibodies are not suitable for IP, the Co-IP assay could not be conducted in tea plants.

**Flavan-3-ols regiospecificity and different acyl donors of co-expressed recombinant CsSCPL4 and CsSCPL5**

Different acyl acceptors and acyl donors were used to test the substrate diversity of the co-expressed recombinant CsSCPL4 and CsSCPL5. In addition to the EGC described above, EC, (±)-catechin, and (±)-gallocatechin were used as acyl acceptors. The UPLC-MS/MS results showed that the co-expressed CsSCPL4 and CsSCPL5 produced EGCG and EGC (Figure S10), but did not produce galloylated compounds from (±)-catechin or (±)-gallocatechin (Figure S13). This result indicates that the co-expressed CsSCPL4 and CsSCPL5 proteins have regiospecificity toward EGC and EC with the 2R,3R-2,3-cis stereo configuration. Likewise, in addition to jG, caffeoyl glucoside
(caffeoyl 1-O-β-D-glucose [Caf-Glu]) was used to test the acyl donor diversity. EC and Caf-Glu were incubated with the co-expressed CsSCPL4 and CsSCPL5, and the enzymatic product of EC caffeoyl ester (EC 3-O-caffeate [EC-Caf]) was detected by UPLC-MS/MS (Figure S10). However, the co-expressed CsSCPL4 and CsSCPL5 did not produce EGCG or ECG with substrates PGG and EGC or EC (Figure S14).

A dual role of βG and kinetics of co-expressed CsSCPL4 and CsSCPL5 toward EGC and βG

Although βG was used as an acyl donor in enzymatic assays, it could also act as an acyl acceptor to compete with EGC (Figure S15a). When the enzymatic assays with co-expressed CsSCPL4 and CsSCPL5 substrates and βG lasted longer, for example 3 h, the results showed that in addition to the formation of EGCG, UPLC-MS/MS analysis could detect another peak with an m/z of 483.0 [M-H]-, which was identified to be digalloyl glucose (DGG), a monomer of hydrolyzable tannins (HTs), based on the MS/MS fragments (Figures S10 and S15b). No DGG was detected when CsSCPL4 or CsSCPL5 was expressed alone (Figure S10). Based on these data, βG played a dual role. This phenomenon has also been observed in enzymatic reactions of other acyltransferases, such as AtSMT, which is associated with the formation of 1,2-di-O-sinapoyl-β-D-glucose and free sinapate (Stehle et al., 2008). However, in the reaction system of EGCG formation catalyzed by co-expressed CsSCPL4 and CsSCPL5, DGG is only a by-product whose production is inhibited when using EGC as a substrate (Figure S15c).

According to the Michaelis–Menten kinetics equation, the $K_m$ and $V_{\text{max}}$ values of co-expressed CsSCPL4 and CsSCPL5 toward EGC were 0.83 mM and 573 pKat mg⁻¹, and toward βG these values were 1.09 mM and 625 pKat mg⁻¹ (Figure S15c,d). The velocity of the DGG versus [βG] plot showed a linear increase in DGG production as the concentrations of βG were increased (Figure S15e). However, it is difficult to obtain the kinetic constant of βG because of the low enzyme activity in this reaction.

CsSCPL4 and CsSCPL5 homologs identified in other flavan-3-ol gallate- or HT-rich plants

Next, we aimed to investigate whether co-expression of SCPL4 and SCPL5 is prevalent in plants rich in galloylated flavan-3-ols or HTs. We first examined 15 plants that had homologs of CsSCPL4 and CsSCPL5 in SCPL clade IA groups d and e, and then rebuilt a snapshot of an unrooted tree to highlight their taxonomic classification (Figure 5a). These 15 plant species are from six orders, including Ericales (C. sinensis, C. oleifera, D. kaki), Vitales (V. vinifera, Vitis rotundifolia), Myrtales (Punica granatium, Eucalyptus grandis), Fagales (Castanea mollissima, Myrica rubra, Quercus suber, Juglans regia), Malpighiales (Manihot esculenta, Hevea brasiliensis, Jatropha curcas), and Sapindales (Mangifera indica). These plant species are rich in galloylated flavan-3-ols or HTs (Table S1). These homologous sequences of CsSCPL4 and CsSCPL5 were used for alignment analysis (Figures S16 and S17). All sequences have signal peptides in the N-terminus. The catalytic triad S-D-H is conserved in all CsSCPL4 homologs. Except for CsSCPL5, DkSCPL1, and VvGAT2, as described above, all CsSCPL5 homologs have the conserved catalytic triad S-D-H (Figure 5b). Although most of these homologs remain to be functionally characterized, these data indicate the prevalence of these genes in plant species rich in galloylated flavan-3-ols or HTs.

In particular, the SCPL5 homologs in C. oleifera and in V. rotundifolia both contain a normal S-D-H catalytic triad (Figure 5b). From an evolutionary perspective, this suggests that mutations at the catalytic residues of CsSCPL5 and VvGAT2 occurred recently (Figure 5c). Based on the previous literature on the evolution of Camellia (Li et al., 2021; Yu et al., 2017) and phylogenetic relationships of plant species (Figure 5c), we speculate that this mutation may have taken place about 10 million years ago.

DkSCPL1 and DkSCPL2 from persimmon and VvGAT1 and VvGAT2 from grape require co-expression for their acyltransferase activity

To understand whether the co-expression of the two protein pairs is also required for acyltransferase activity, we cloned the DkSCPL2/DkSCPL1 and VvGAT1/VvGAT2 genes and then infiltrated each single gene or pair into the leaves of N. benthamiana as described above. The enzymatic assays showed that neither transiently expressed DkSCPL1 nor transiently expressed DkSCPL2 could catalyze EGC and βG to form EGCG, while the two paralogs when co-expressed catalyzed the formation of EGCG (Figure 6a). This result showed that the co-expression of DkSCPL1 and DkSCPL2 is likely responsible for their acyltransferase activity.

In contrast to the positive results of the co-expression of DkSCPL1 and DkSCPL2, neither VvGAT1 or VvGAT2 alone nor VvGAT1 and VvGAT2 co-expressed could catalyze the formation of EGCG from EGC and βG. To further test the activity of these two grape paralogs, the signal peptides of VvGAT1 and VvGAT2 were substituted with those of CsSCPL4 and CsSCPL5 to create chimeric Cs4SP-VvGAT1 and Cs5SP-VvGAT2 (Figure S18), respectively. The two chimeric paralogs were transiently co-expressed to test their activity as described above. The results showed that Cs4SP-VvGAT1 and Cs5SP-VvGAT2 when co-expressed catalyzed the formation of EGCG from EGC and βG (Figure 6a). This result indicates that the co-expression of VvGAT1 and VvGAT2 contributes to the galloylation of flavan-3-ols and that the signal peptides also play a role in the use of N. benthamiana leaves to express active enzymes.
and CsSCPL5 homologs identified in the genome of C. oleifera. Despite their close phylogenetic relationship, UPLC-MS/MS analysis revealed apparent differences in the phenolic profiles. Young twigs of tea were rich in catechin aglycones and galloylated catechins, while young twigs of oil tea were rich in HTs but...

Figure 5. Homologs of CsSCPL4 and CsSCPL5 identified in 15 plants rich in flavan-3-ol gallates or hydrolyzable tannins (HTs). (a) A phylogenetic tree shows CsSCPL4 homologs in group d and CsSCPL5 homologs in group e of 15 plants. (b) Amino acid sequence alignment showing that the catalytic triad residues (S-D-H) are fully conserved in group d but varied in group e. Blue color shows conserved amino acids, and red color indicates varied amino acids. (c) Phylogenetic tree showing the divergence time and evolutionary relationship in different plant species. *indicates the genus Vitis; Star symbol indicates the genus Camellia. Ng, Neogene; Pg, Paleogene; J, Jurassic; Tri, Triassic; P, Permian; Pe, Pennsylvanian; Mis, Mississippian; S, Silurian; O, Ordovician; C, Cambrian.

Co-expression of two SCPL-ATs for galloylation

Camellia oleifera (oil tea) and C. sinensis (tea) belong to two subgroups of Camellia. Despite their close phylogenetic relationship, UPLC-MS/MS analysis revealed apparent differences in the phenolic profiles. Young twigs of tea were rich in catechin aglycones and galloylated catechins, while young twigs of oil tea were rich in HTs but...
lacked or had only trace levels of galloylated catechins (Figure 6b; Table S2). To better understand the formation of HTs, we mined the transcriptomes of oil tea and identified \( \text{CsSCPL4} \) and \( \text{CsSCPL5} \) homologs, named \( \text{CoSCPL4} \) and \( \text{CoSCPL5} \). Interestingly, the catalytic triads of both \( \text{CoSCPL4} \) and \( \text{CoSCPL5} \) were composed of the conserved S-D-H (Figure 5b). To test their functions, \( \text{CoSCPL4} \) and \( \text{CoSCPL5} \) were transiently expressed alone and \( \text{CoSCPL4} \) and \( \text{CoSCPL5} \) were transiently co-expressed in leaves of \( \text{N. benthamiana} \). Crude proteins were extracted for enzymatic assays with EGC and \( \beta\text{G} \) as substrates. UPLC-MS/MS analysis of enzymatic products showed that \( \text{CoSCPL4} \) and \( \text{CoSCPL5} \) when co-expressed not only converted EGC (an acyl acceptor) and \( \beta\text{G} \) (an acyl donor) into EGCG but also catalyzed the formation of DGG from \( \beta\text{G} \) (Figure 6c). In contrast, \( \text{CoSCPL4} \) or \( \text{CoSCPL5} \) alone could not catalyze the formation of EGCG and DGG. These data indicate that co-expression of \( \text{CoSCPL4} \) and \( \text{CoSCPL5} \) is likely responsible for their acyltransferase activity.

Figure 6. Functional analysis of the co-expression of two SCPL homologs from \( \text{Diospyros kaki} \), \( \text{Vitis vinifera} \), and \( \text{Camellia oleifera} \).

(a) Extracted ion chromatographs showing epigallocatechin gallate (EGCG) formation in the enzymatic assay. Each single gene and each pair of genes were transiently expressed in leaves of \( \text{Nicotiana benthamiana} \). Cs4SP-VvGAT1, the CsSCPL4 signal peptide was fused to the N-terminus of VvGAT1; Cs5SP-VvGAT2, the CsSCPL5 signal peptide was fused to the N-terminus of VvGAT2.

(b) Total ion current chromatograms showing different phenolic compounds in \( \text{C. sinensis} \) and \( \text{C. oleifera} \). Compounds labeled with HT1-7 are seven hydrolyzable tannins listed in Table S2.

(c) Extracted ion chromatographs showing EGCG and digalloyl glucose (DGG) formation in the enzymatic assays. EGCG was detected in the enzymatic assay with the substrates EGC and \( \beta\text{G} \). DGG was detected in the enzymatic assay with the substrate \( \beta\text{G} \). Co4 + Co5-TE, transient co-expression of \( \text{CoSCPL4} \) and \( \text{CoSCPL5} \); Co4-TE, transient expression of \( \text{CoSCPL4} \) alone; Co5-TE, transient expression of \( \text{CoSCPL5} \) alone.
DISCUSSION

CsSCPL5 is an NCCP of CsSCPL4

We named CsSCPL5 an NCCP of CsSCPL4 based on its function in the galloylation of EGC, EC, and βG. Phylogenetic analysis classified CsSCPL4 in group d and CsSCPL5 in group e, which are both included in clade IA of the SCPL family (Figure 5a; Figure S6). Transcriptional and protein analyses revealed that the two paralogs were co-expressed with a similar pattern, and their co-expression was closely associated with the profiles of EGCG and ECG in different tea tissues (Figure 1d-f). Enzymatic analyses showed that neither CsSCPL4 nor CsSCPL5 alone had catalytic activity, and in N. benthamiana leaves only their co-expression resulted in efficient acyltransferase activity (Figure 2).

One study reported that three recombinant CsSCPL proteins expressed in E. coli, named CsSCPL11 (GenBank: QID89725), CsSCPL13 (GenBank: QID89726), and CsSCPL14 (GenBank: QID89727), showed galloylation activity with PGG and EC or EGC as substrates (Ahmad et al., 2020). However, our experimental results showed that prokaryote-expressed recombinant CsSCPL4 and CsSCPL5 did not have any catalytic activity (Figures S8 and S9), and a series of experimental results showed that the processed mature CsSCPL protein was the main form for catalytic activity (Figures 2i and 3e). Amino acid sequence alignment revealed that the three CsSCPLs reported by Ahmad et al. (2020) were inconsistent with CsSCPL4 and CsSCPL5; particularly, CsSCPL11 is an incomplete CsSCPL4 sequence which lacks the serine (S) residue in the catalytic triad (S-D-H) (Figure S19).

The catalytic triad residues S-D-H of SCPL family enzymes are conserved and necessary for transferase activities. The predicted catalytic triads of CsSCPL4 and CsSCPL5 are composed of S-D-H and T-D-Y, respectively (Figure 1c). Based on this, it is speculated that CsSCPL5, which lacks the conserved catalytic triad, lost its catalytic activity and performs other functions. In a series of CsSCPL4 and CsSCPL5 co-expression experiments, all mutations of the catalytic residues of CsSCPL4 abolished enzyme activity, while all mutations of the predicted catalytic residues of CsSCPL5 did not. Even if the catalytic triads were exchanged between CsSCPL4 and CsSCPL5, the co-expression variants had no enzyme activity (Figure 3d). Therefore, our results demonstrated that CsSCPL4 is the catalytic enzyme, while CsSCPL5 is an NCCP of CsSCPL4.

The functional role of CsSCPL5

To understand whether this enzymatic event is prevalent in plants, we studied their homologs in flavan-3-ol gallate- or HT-rich plants. Homologs of CsSCPL5 and their similar roles were shown in oil tea, grape, and persimmon (Figure 6). Likewise, other EGCG- or HT-rich plants investigated in this study have CsSCPL4 and CsSCPL5 homolog pairs (Figure 5a).

Due to the lack of the conserved catalytic triad S-D-H (Figure 5b), we speculated that SCPL5 orthologous proteins in tea plant, grape, and persimmon are pseudoenzymes. So, is being a pseudoenzyme a necessary condition for being an NCCP? Based on heterologous expression and in vitro enzymatic assays of C. oleifera SCPL genes (Figure 6c), the answer is negative. The SCPL5 homologous protein from C. oleifera has complete S-D-H residues, but still acts as an NCCP. Considering the fact that the SCPL5 orthologs are all pseudoenzymes in three plant species rich in flavan-3-ol gallates, it is speculated that the emergence of pseudoenzymes is closely related to the high accumulation of flavan-3-ol gallates in tea, grape, and persimmon.

As more genomes of organisms are sequenced, many proteins are annotated as pseudoenzymes due to the lack of essential catalytic residues or domains (Eyers & Murphy, 2016; Ribeiro et al., 2019). Studies show that genes encoding pseudoenzymes usually come from multi-gene families. When genes are replicated or expanded, the selection pressure disappears, which makes it possible for genes encoding pseudoenzymes to appear (Eyers & Murphy, 2016). Based on results of the CoSCPL4 and CoSCPL5 co-expression enzymatic assays (Figure 6c) and the evolutionary status of C. oleifera (Figure 5c), we believe that the emergence of the SCPL5 ortholog as an NCCP occurred earlier than the emergence of SCPL5 pseudoenzymes.

CsSCPL5 may serve as an NCCP by acting as a pseudoenzyme or ‘moonlighting’ protein. Recent studies have shown that pseudoenzymes have quite diverse functions, such as allosteric regulation, supporting signal complex assembly, binding and inhibiting active enzymes, stabilizing the folded state of active enzymes, regulating protein localization, recruiting or isolating substrates, and regulating the folding of conventional enzymes (Jeffery, 2019). Typical functional mechanisms of plant pseudoenzymes include the following: (i) form heteromultimers with functional homologous enzymes or other functional enzymes to increase their efficiency (Ban et al., 2018; Jiang et al., 2015); (ii) participate in stabilization of protein complexes (Dell’Aglio et al., 2017; Moccand et al., 2014); and (iii) bind to and stabilize the intermediate compounds of metabolic pathways (Ban et al., 2018).

Post-translational processing is a process that many plant SCPL-ATs must undergo to form mature proteins (Li & Steffens, 2000; Mugford et al., 2009; Qiu et al., 2020). Most mature SCPL proteins are heterodimers with large and small subunits. Only a few reported mature SCPL enzymes act as monomers, such as the AtSMT protein (Hause et al., 2002). Based on our immunoblot assay results using antibodies against CsSCPL4 and CsSCPL5 in tea plant (Figure 1e) and in transgenic N. benthamiana expressing the two genes (Figure 2i), we believe that CsSCPL4 must undergo post-translational processing to form large and small subunits and become catalytically functional...
active. We speculate that CsSCPL5 is likely to play a key role in increasing stability of the CsSCPL4 pro-peptide or mature protein and promote the post-translational processing to form a catalytically active protein.

Many pseudoenzymes have been reported to perform various functions by interacting with metal ions, small ligands, proteins, and other biological macromolecules (Todd et al., 2002). The pull-down and Co-IP experiments confirmed the interaction between the CsSCPL4 and CsSCPL5 pro-peptides (Figure 4). Alteration of the amino acids at the catalytic triad of CsSCPL4 and the disulfide bonds of CsSCPL4 and CsSCPL5 affected the post-translational processing and stability of the two proteins (Figure 3e), indicating the complexity of post-translational processing of the two proteins. In short, although the interaction between the CsSCPL4 and CsSCPL5 pro-peptides has been discovered, the mechanism of protein interaction is still poorly understood.

There are still many incomprehensible experimental results. For example, why do mutation of the catalytic residues of CsSCPL4 and elimination of the disulfide bond affect the amounts of the two mature proteins? We speculate that it is likely that the alterations in these residues change the features of the unprocessed protein, such as hydrophobicity and the steric structure, which further affect the interaction between the two proteins, reducing the stability of the two proteins and preventing them from being processed into mature protein.

EXPERIMENTAL PROCEDURES

Plant materials and chemical reagents

Tea (C. sinensis L. cv. ‘Shuchazao’), oil tea (C. oleifera Abel.), grape (V. vinifera L.), and persimmon (D. kaki Thunb.) plants were cultivated in the experimental garden at Anhui Agricultural University (Anhui, China). Fresh samples were collected, rapidly frozen in liquid nitrogen, and stored at −80°C until use. Nicotiana tabacum cv. G28 (tobacco) and N. benthamiana were grown in a greenhouse with a light intensity of 150–200 μmol m−2 sec−1, a photoperiod of 16/8 h light/dark, and a temperature of 22 ± 2°C.

Authentic standard compounds, EC, EGC, ECG, EGCG, PGG, catechin, and galloacetin were purchased from Sigma-Aldrich (St. Louis, MO, USA). 1-O-β-Glucogallin and caffeoyl glucoside were biosynthesized with CsUGT84A22 according to the methods in our previous report (Cui et al., 2016). EC-Caf was provided by Prof. Guanhui Bao. Mouse monoclonal anti-GFP was purchased from Abmart Inc. (Shanghai, China). Rabbit polyclonal anti-plant actin was purchased from Sangon Biotech Co., Ltd. (Shanghai, China). Mouse monoclonal anti-GST, mouse monoclonal anti-MBP, and horseradish peroxidase-conjugated secondary antibodies were purchased from Proteintech, Inc. (Wuhan, China).

Purification of galloyltransferase and protein MS identification

The native EC-GT enzyme was extracted and purified from young leaves of the tea plant following our previously published protocol (Liu et al., 2012). The resulting EC-GT protein bands on SDS-PAGE gels were excised and then digested with trypsin for protein MS identification by BGI Genomics Co., Ltd. (Shenzhen, China). Briefly, the digested proteins were analyzed through an LC-MS/ MS system with an LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) in positive ion mode. The identified peptide fragments were blasted against the C. sinensis genome database (http://tpia.teaplant.org/) to obtain gene sequences.

Gene cloning

RNA isolation and reverse transcription were completed as reported previously (Wang et al., 2018). Four cDNA libraries were developed from the leaves of tea and oil tea plants and the fruits of grape and persimmon. After cloning and analysis, eight accurate full cDNA sequences were identified and named CsSCPL4 (GenBank: MZ462934), CsSCPL5 (GenBank: MZ462935), CoSCPL4 (GenBank: MZ462936), CoSCPL5 (GenBank: MZ462937), VvGAT1 (GenBank: XM_002272080.4), VvGAT2 (GenBank: XM_002269403.3), DkSCPL1 (GenBank: AB195285.1), and DkSCPL2 (GenBank: AB473186.1). The gene expression and evolution analyses are included in Data S1.

Quantitative analyses of flavan-3-ol gallates in C. sinensis

Buds, leaves, and new stems at different developmental stages were collected in liquid nitrogen and then freeze-dried. The galloyl compounds in the samples were extracted and quantitatively analyzed in strict accordance with the methods in our previous report (Zhuang et al., 2020).

Transient expression of CsSCPL4 and CsSCPL5 in N. benthamiana

The open reading frames (ORFs) of CsSCPL4 and CsSCPL5 were individually fused into a CaMV35S promoter-driven PCB2004 binary vector by using the Gateway LR Clonase enzyme (Invitrogen, Carlsbad, CA, USA). The two resulting vectors PCB2004-CsSCPL4 and PCB2004-CsSCPL5 were individually introduced into Agrobacterium tumefaciens strain GV3101 (pSoup-19) via electroporation. The positive-colony suspension cultures were centrifuged, and the pellets were resuspended in infiltration buffer (10 mM MES, pH 5.6, 10 mM MgCl2, and 100 μM acetosyringone). The OD600 values of the suspensions were adjusted to 0.6. For co-expression, the two suspensions were mixed in equal volumes and then injected into the leaves of 30-day-old N. benthamiana plants. The infiltrated plants were cultivated for 72 h in the greenhouse, and then leaves were collected and frozen in liquid nitrogen for enzyme assays as described below. The detail steps of crude protein extraction from plants are included in Data S1.

Enzymatic assays

The enzymatic assay was carried out in 100-μl reaction mixtures containing 50 mM phosphate buffer (pH 6.0), 0.4 mM EGCG or EC, 0.4 mM C-G or PGG, 2 mM ascorbic acid, and 60 μg of desalted protein at 30°C for 3 h. All reactions were stopped by adding an equal volume of methanol and samples were centrifuged at 10 000 g at 4°C for 10 min. The supernatant was pipetted into a new tube for UPLC or UPLC-QqQ-MS/MS analysis. Detailed steps are included in Data S1.

Site-directed mutation

Based on the sequence information, gene-specific primers were designed at the codons encoding the catalytic triad and disulfide
Pull-down assay

The ORFs of CsSCPL4 and CsSCPL5 without the signal peptides were individually fused with MBP and GST tags and then cloned into the co-expression vector pRSETduet-1. The resulting vector MBP-Cs4 + GST-Cs5 was transformed into the E. coli Rosetta strain for co-expression of the two proteins. In addition, the same steps were completed to produce MBP + GST, MBP + GST-Cs5, and MBP-Cs4 + GST as controls. The GST-tagged recombinant proteins were purified through Glutathione resin (Takara, Beijing, China). Strict washing was performed to exclude any remaining potential contamination. SDS-PAGE and immunoblot analysis were conducted with crude proteins and purified proteins to verify the interaction between MBP-CsSCPL4 and GST-CsSCPL5. The immunoblot protocol is included in Data S1.

Co-IP assay

Five binary vectors (PCB2004, PCB2004-CsSCPL4, PCB2004-CsSCPL5, pGWB5-GFP, and pGWB5-CsSCPL4-GFP) were included with the following four different combinations to show the interaction between CsSCPL4 and CsSCPL5: (i) PCB2004 as vector control, (ii) GFP + CsSCPL4 + CsSCPL5, (iii) GFP + CsSCPL5, and (iv) CsSCPL4-GFP + CsSCPL5. They were transformed into N. benthamiana leaves by agroinfiltration. After 72 h, the infiltrated leaves were collected and ground in liquid nitrogen for crude protein extraction. The resulting protein extracts were incubated with 25 μl anti-GFP magnetic beads (Abmart, Shanghai, China) at 4 °C for 4 h on a rotator. Then, the beads were washed four times with TBS buffer following the manufacturer’s guide. Finally, the immunoprecipitates on beads were re-suspended in 100 μl 1x SDS sample loading buffer and heated at 100°C for 8 min. The supernatants were used for immunoblot analysis with anti-GFP, anti-CsSCPL4, and anti-CsSCPL5.

Comparative analyses of galloyl compounds in C. sinensis and C. oleifera

The top bud and two new leaves of new short twigs from C. sinensis and C. oleifera were harvested into liquid nitrogen, freeze-dried, and ground into fine powder samples. The galloyl compounds in the samples were extracted with 80% methanol and further used for UPLC-Q-TOF-LC/MS analysis according to our recently reported protocol (Zhuang et al., 2020).

Statistical analysis

The statistical significance of the data was determined by SPSS software version 16.0 (SPSS, Inc., Chicago, IL, USA). Multiple comparison analysis was performed with Duncan's test. Significant differences are represented by different lowercase letters. A P-value of <0.05 was considered to indicate statistical significance.

ACKNOWLEDGMENTS

We thank the National Natural Science Foundation of China (32072621, 31970676, 31970677, 31902069), the Joint Funds of the National Natural Science Foundation of China (U21A20223), the Collegiate Natural Science Foundation of Anhui Province, China (03087060), and the Natural Science Foundation of Anhui Province, China (1908085MC100).

© 2022 The Authors. The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd., The Plant Journal, (2022), 111, 117–133
REFERENCES

Ahmad, M.Z., Li, P., She, G., Xia, E., Benedito, V.A., Wan, X.C. et al. (2020) Genome-wide analysis of serine carboxypeptidase-like acyltransferase gene family for evolution and characterization of enzymes involved in the biosynthesis of galloylated catechins in the tea plant (Camellia sinensis). Frontiers in Plant Science, 11, e848.

Akagi, T., Suzuki, Y., Ikegami, A., Kamitakahara, H., Takano, T., Nakatsubo, F. et al. (2007) Related Arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. Plant Physiology, 144, 1986-1999.

Ban, Z., Qin, H., Mitchell, A.J., Liu, B., Zhang, F., Weng, J.-K. et al. (2005) Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends in Plant Science, 10, 542-549.

Jeffery, C.J. (2019) The demise of catalysis, but new functions arise: pseudoenzymes as the phosphexes of the protein world. Biochemical Society Transactions, 47, 371-379.

Jiang, W., Yin, Q., Wu, R., Zheng, G., Liu, J., Dixon, R.A. et al. (2015) Role of a chalcone isomerase-like protein in flavonoid biosynthesis in Arabidopsis thaliana. Journal of Experimental Botany, 66, 7165-7179.

Kiyosue, M., TAKEIZumi, E. & Jakubowska, A. (2013) Discovery of flavan-3-ols and galloylated derivatives in C. sinensis and C. oleifera leaves through UPLC-TOF-LC/MS analysis.

The primers used in this study.

The Plant Journal, 111, 117–133.
Mugford, S.T., Qi, X., Bakht, S., Hill, L., Wegel, E., Hughes, R.K. et al. (2009) A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in Oats. Plant Cell, 21, 2473-2484.

Nishizaki, Y., Yasunaga, M., Okamoto, E., Okamoto, M., Hirose, Y., Yamaguchi, M. et al. (2013) p-Hydroxybenzoyl-glucose is a zwiter donor for the biosynthesis of 7-polyacylated anthocyanin in Delphinium. Plant Cell, 25, 4150-4165.

Obreque-Slier, E., Pena-Neira, A., Lopez-Solis, R., Zamora-Marin, F., Ricardo-da-Silva, J.M., Daglia, M., Rastrelli, L. & Nabavi, S.M. (2016) The sng2 mutant of Arabidopsis is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose: malate sinapoyltransferase gene provides signatures of natural light selection in Brassicaceae. Nature Communications, 7, 12399.

Shirley, A.M., McMichael, C.M. & Chapple, C. (2001) Arabidopsis defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose: choline sinapoyltransferase. The Plant Journal, 28, 83-94.

Singh, N.A., Mandal, A.K. & Khan, Z.A. (2016) Potential neuroprotective properties of epigallocatechin-3-gallate (EGCG). Nutrition Journal, 15, 60.

Stehle, F., Brandt, W., Milkowski, C. & Strack, D. (2008) Activities of Arabidopsis sinapoylgucose: malate sinapoyltransferase shed light on functional diversification of serine carboxypeptidase-like acyltransferases. Phytochemistry, 69, 1826-1831.

Todd, A.E., Orenge, C.A. & Thornton, J.M. (2002) Sequence and structural differences between enzyme and nonenzyme homologs. Structure, 10, 1435-1451.

Tohge, T., Wendenburg, R., Ishihara, H., Nakabayashi, R., Watanabe, M., Sulpice, R. et al. (2016) Characterization of a recently evolved flavonoid-phenylacyltransferase gene provides signatures of natural light selection in Brassicaceae. Nature Communications, 7, 12399.

Wang, W., Zhou, Y., Wu, Y., Dai, X., Liu, Y., Gian, Y. et al. (2018) Insight into catechins metabolic pathways of Camellia sinensis based on genome and transcriptome analysis. Journal of Agricultural and Food Chemistry, 66, 4281-4293.

Weier, D., Mittasch, J., Strack, D. & Milkowski, C. (2008) The genes BnSCT1 and BnSCT2 from Brassica napus encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression. Planta, 227, 375-385.

Wilson, A.E., Matei, H.D. & Tian, L. (2016) Glucose ester enabled acylation in plant specialized metabolism. Phytochemistry Reviews, 15, 1057-1074.

Yu, X.Q., Gao, L.M., Soltis, D.E., Soltis, P.S., Yang, J.B., Fang, L. et al. (2017) Insights into the historical assembly of East Asian subtropical evergreen broadleaf forests revealed by the temporal history of the tea family. The New Phytologist, 215, 1235-1248.

Zhang, J., Dai, X., Zhu, M., Zhang, S., Dai, Q., Jiang, X. et al. (2020) Evaluation of astringent of green tea through mass spectrometry-based targeted metabolic profiling of polyphenols. Food Chemistry, 305, 125070.