Phytochemistry 164 (2019) 50–59

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Biosynthesis of methyl (*E*)-cinnamate in the liverwort *Conocephalum salebrosum* and evolution of cinnamic acid methyltransferase

Chi Zhang[a], Xinlu Chen[a], Barbara Crandall-Stotler[b], Ping Qian[c], Tobias G. Köllner[d], Hong Guo[e,f], Feng Chen[a,*]

[a] Department of Plant Sciences, University of Tennessee, Knoxville, TN 37996, USA
[b] Department of Plant Biology, Southern Illinois University, Carbondale, IL 62901, USA
[c] Shandong Agricultural University, Chemistry and Material Science Faculty, Tai'an 271018, Shandong, China
[d] Department of Biochemistry, Max Planck Institute for Chemical Ecology, Hans-Knöll-Strasse 8, D-07745 Jena, Germany
[e] Department of Biochemical, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA
[f] UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA

ARTICLE INFO

Keywords:
Conocephalum salebrosum
Conocephalaceae
Liverworts
SABATH methyltransferase
Specialized metabolism
Convergent evolution

ABSTRACT

Methyl (*E*)-cinnamate is a specialized metabolite that occurs in a variety of land plants. In flowering plants, it is synthesized by cinnamic acid methyltransferase (CAMT) that belongs to the SABATH family. While rarely reported in bryophytes, methyl (*E*)-cinnamate is produced by some liverworts of the *Conocephalum conicum* complex, including *C. salebrosum*. In axenically grown thalli of *C. salebrosum*, methyl (*E*)-cinnamate was detected as the dominant compound. To characterize its biosynthesis, six full-length *SABATH* genes, which were designated *CsSABATH1-6*, were cloned from *C. salebrosum*. These six genes showed different levels of expression in the thalli of *C. salebrosum*. Next, *CsSABATH1-6* were expressed in *Escherichia coli* to produce recombinant proteins, which were tested for methyltransferase activity with cinnamic acid and a few related compounds as substrates. Among the six SABATH proteins, CsSABATH6 exhibited the highest level of activity with cinnamic acid. It was renamed CsCAMT. The apparent *Km* value of CsCAMT using (*E*)-cinnamic acid as substrate was determined to be 50.5 μM. In contrast, CsSABATH4 was demonstrated to function as salicylic acid methyltransferase and was renamed CsSAMT. Interestingly, the *CsCAMT* gene from a sabinene-dominant chemotype of *C. salebrosum* is identical to that of the methyl (*E*)-cinnamate-dominant chemotype. Structure models for CsCAMT, CsSAMT and one flowering plant CAMT (ObCCMT1) in complex with (*E*)-cinnamic acid and salicylic acid were built, which provided structural explanations to substrate specificity of these three enzymes. In phylogenetic analysis, CsCAMT and ObCCMT1 were in different clades, implying that methyl (*E*)-cinnamate biosynthesis in bryophytes and flowering plants originated through convergent evolution.

1. Introduction

Plants produce diverse volatile specialized metabolites that include carboxyl methyl esters. These methyl esters play important roles in various biological processes (Chen et al., 2003a; Knudsen et al., 1993; Seo et al., 2001). Methyl (*E*)-cinnamate, methyl ester of (*E*)-cinnamic acid, is one such compound. Its odor is defined as "balsamic, strawberry, fruity, cherry". Methyl (*E*)-cinnamate occurs as a scent compound in many flowers such as orchids (Kaiser, 1993) and *Narcissus* species (Arai, 1994). It is also made by some fruits such as strawberry (Gomes da Silva and Chaves das Neves, 1999) and plum (Ismail et al., 1980). Its occurrence in leaves of some plants is also known, such as in basils (Viña and Murillo, 2003) and *Eucalyptus* species (Curtis et al., 1990). In flowers and fruits, methyl (*E*)-cinnamate is probably involved in attracting pollinators and seed dispersers, respectively (Dodson et al., 1969; Eltz and Lunau, 2005). In leaves, this compound most likely functions as a chemical defense (Hattori et al., 1992). It has been shown to inhibit the growth of pathogenic fungi and bacteria (Ali et al., 2010). The molecular basis of methyl (*E*)-cinnamate biosynthesis has so far only been studied in basil (Kapteyn et al., 2007), where cinnamic acid/*p*-coumaric acid methyltransferases (ObCCMTs) catalyze the formation of methyl (*E*)-cinanmate using *S*-adenosyl-L-methionine (SAM) as methyl donor (Kapteyn et al., 2007).

ObCCMTs belong to the family of methyltransferases known as SABATH (D'Auria et al., 2003). Most substrates of SABATH methyltransferases are important phytohormones including indole-3-acetic

* Corresponding author.
E-mail address: fengc@utk.edu (F. Chen).

https://doi.org/10.1016/j.phytochem.2019.04.013
Received 11 February 2019; Received in revised form 16 April 2019; Accepted 24 April 2019

acid, gibberellins, salicylic acid and jasmonic acid (Chen et al., 2003a; Qin et al., 2005; Seo et al., 2001; Varbanova et al., 2007; Zhao et al., 2007). Other substrates of SABATH methyltransferases include benzoic acid (Murfitt et al., 2000), *p*-methoxybenzoic acid (Koeduka et al., 2016), farnesoic acid (Yang et al., 2006), anthranilic acid (Köllner et al., 2010), nicotinic acid (Hippauf et al., 2010) and theobromine (Kato et al., 2000; McCarthy and McCarthy, 2007). Such diverse substrates within a same protein family make SABATH proteins useful models for studying substrate specificity evolution (Huang et al., 2012). Within the *SABATH* family, indole-3-acetic acid methyltransferase gene (*IAMT*) has been proposed to be an ancient member in seed plants (Zhao et al., 2008). *ObCCMTs* are closely related to *IAMT* (Kapteyn et al., 2007), making it tempting to speculate that *ObCCMTs* may have evolved from *IAMT* through gene duplication and functional divergence.

Liverworts are considered extant plants that are most closely related to ancestral land plants (Bowman et al., 2017). Like flowering plants, liverworts produce a diverse array of specialized metabolites, especially terpenoids (Chen et al., 2018). While rarely reported, methyl (*E*)-cinnamate is known to be made by liverworts within the *Conocephalum conicum* complex (Ghani et al., 2016; Harinantenaina et al., 2007; Toyota, 2000; Wood et al., 1996). Species of the *Conocephalum conicum* (L.) Underw. (Conocephalaceae) complex, known as great scented liverworts, are widely distributed in North America, Europe and East Asia (Szweykowski et al., 2005). They can be easily recognized by their snake skin-like surface and aromatic odor (Atherton et al., 2010). The *C. conicum* complex is comprised of two morphologically distinct species, namely *C. conicum* and *C. salebrosum* Szweyk., Buczkowska & Odrzykoski and multiple cryptic species (Szweykowski et al., 2005). Although when Szweykowski et al. (2005) named *C. salebrosum*, they indicated that it mostly corresponded to *C. conicum* genotype S, North American specimens with genotype A from North Carolina were also included. This species level relationship between genotypes A and S (=*C. salebrosum*) was further confirmed in Ludwiczuk et al. (2013). In an early study of presumed *C. conicum* populations in the in United States, methyl (*E*)-cinnamate was identified as a major constituent from a sample of unknown natural origin as well as a population from Southern Illinois, but was not detected in the extracts of wild *C. conicum* from coastal northern California (Wood et al., 1996). In subsequent studies by a Japanese group, three chemotypes of *C. conicum* were identified: methyl (*E*)-cinnamate-dominant, sabinene-dominant and bornyl acetate-dominant (Toyota, 2000; Toyota et al., 1997). In this study, we aim at elucidating the molecular basis of methyl (*E*)-cinnamate biosynthesis in the North American member of the *C. conicum* complex (i.e., *C. salebrosum*), comparing it to that in flowering plants and gaining understanding about the molecular basis for methyl (*E*)-cinnamate variations among *C. conicum* populations.

2. Results

2.1. Production of axenic culture of C. salebrosum and chemical profiling

The sample collected from Illinois (IL) morphologically matched *C. salebrosum*, a segregate species of *C. conicum* (Szweykowski et al., 2005). This population is the chemotype with methyl (*E*)-cinnamate as the dominant compound described by Wood et al. (1996), and was therefore chosen as the model chemotype for most of the experiments in this study. Live plants were collected in the field and grown in axenic culture. Chemical profiling analysis was performed using gas chromatography–mass spectrometry (GC-MS) to determine the volatile components in the axenically grown gametophytes of this species of the *C. conicum* complex. From organic extraction, a total of 10 volatiles were detected (Fig. 1A), of which methyl (*E*)-cinnamate was the dominant component, accounting for 50.5% of the total volatiles. The monoterpene sabinene was the second most abundant with a concentration of 20% of that of methyl (*E*)-cinnamate. The thallus of *C. salebrosum*, when growing in the wild, possesses a characteristic aromatic odor. To determine whether *C. salebrosum* grown axenically on culture medium emits volatiles, we performed headspace collection of explants of *C. salebrosum*, which was analyzed by GC-MS. Three dominant compounds were detected (Fig. 1B). All of them were also detected by organic extraction. Similar to organic extraction, methyl (*E*)-cinnamate was the most abundant compound in the volatile bouquet of *C. salebrosum* (Fig. 1B).

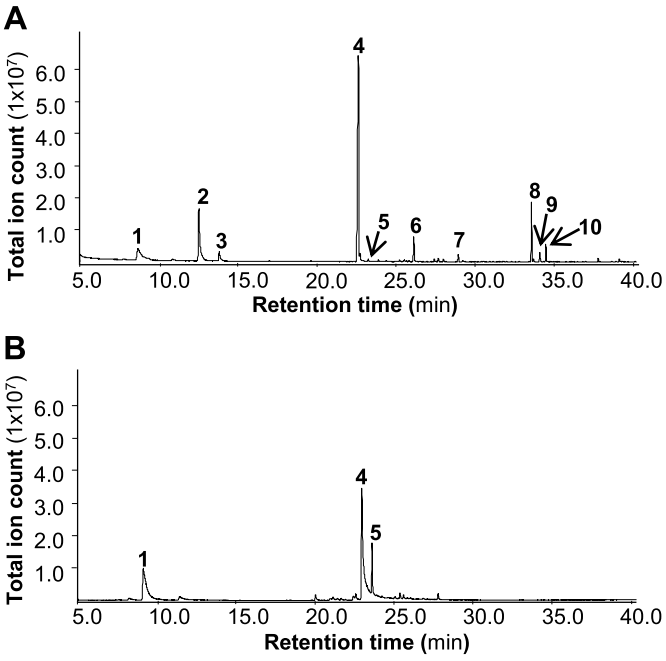

Fig. 1. Chemical analysis of *C. salebrosum*. (A) GC chromatogram of the organic extraction of *C. salebrosum*. 1*, sabinene; 2, 1-octanol; 3, 1-octenyl acetate; 4*, methyl (*E*)-cinnamate; 5, unidentified sesquiterpene hydrocarbon; 6, 10-*epi*-cubebol; 7, 1-*epi*-cubenol; 8, 1-octadecyne; 9, (*E*)-2-tetradecen-1-ol; 10, phytol. (B) GC chromatogram of headspace collection of *C. salebrosum*. 1*, sabinene; 4*, methyl (*E*)-cinnamate; 5, unidentified sesquiterpene hydrocarbon. "*" indicates compounds that were verified by authentic standards.

2.2. Isolation of SABATH genes from C. salebrosum and gene expression analysis

We hypothesized that in the *C. salebrosum* methyl (*E*)-cinnamate is synthesized by the action of SABATH methyltransferase. To identify *SABATH* genes, we analyzed the transcriptome of *C. conicum* prepared from vegetative thalli (https://sites.google.com/a/ualberta.ca/onekp), which was produced by the 1 KP consortium (Matasci et al., 2014). Via a HMMR search (Finn et al., 2011) using Methyltransf_7 Pfam profile (Finn et al., 2015) as query, a total of nine unigenes encoding SABATH proteins were identified. Six of them appear to be in full-length when compared to other SABATH proteins and their corresponding orthologs in *C. salebrosum* were designated CsSABATH1-6 (Table S1). Using RT-PCR, full-length cDNAs for all six CsSABATH genes were cloned from the thallus of the axenically grown *C. salebrosum*. The lengths of six proteins range from 383 to 443 amino acids (Fig. 2). Sequence similarities among CsSABATHs range from 37% to 52%. Their similarities to ObCCMT1, one known CAMT from basil, are between 36% and 48%. The amino acid residues for binding the methyl donor SAM among CsSABATHs and ObCCMT1 are identical. The residues that interact with the carboxyl moiety of the substrate are also conserved. However, the residues that interact with the aromatic moiety are significantly different (Fig. 2).

With six *SABATH* genes cloned from *C. salebrosum*, next we compared their expression levels using semi-quantitative reverse transcription polymerase chain reaction (RT-PCR) analysis. At 25 cycles of

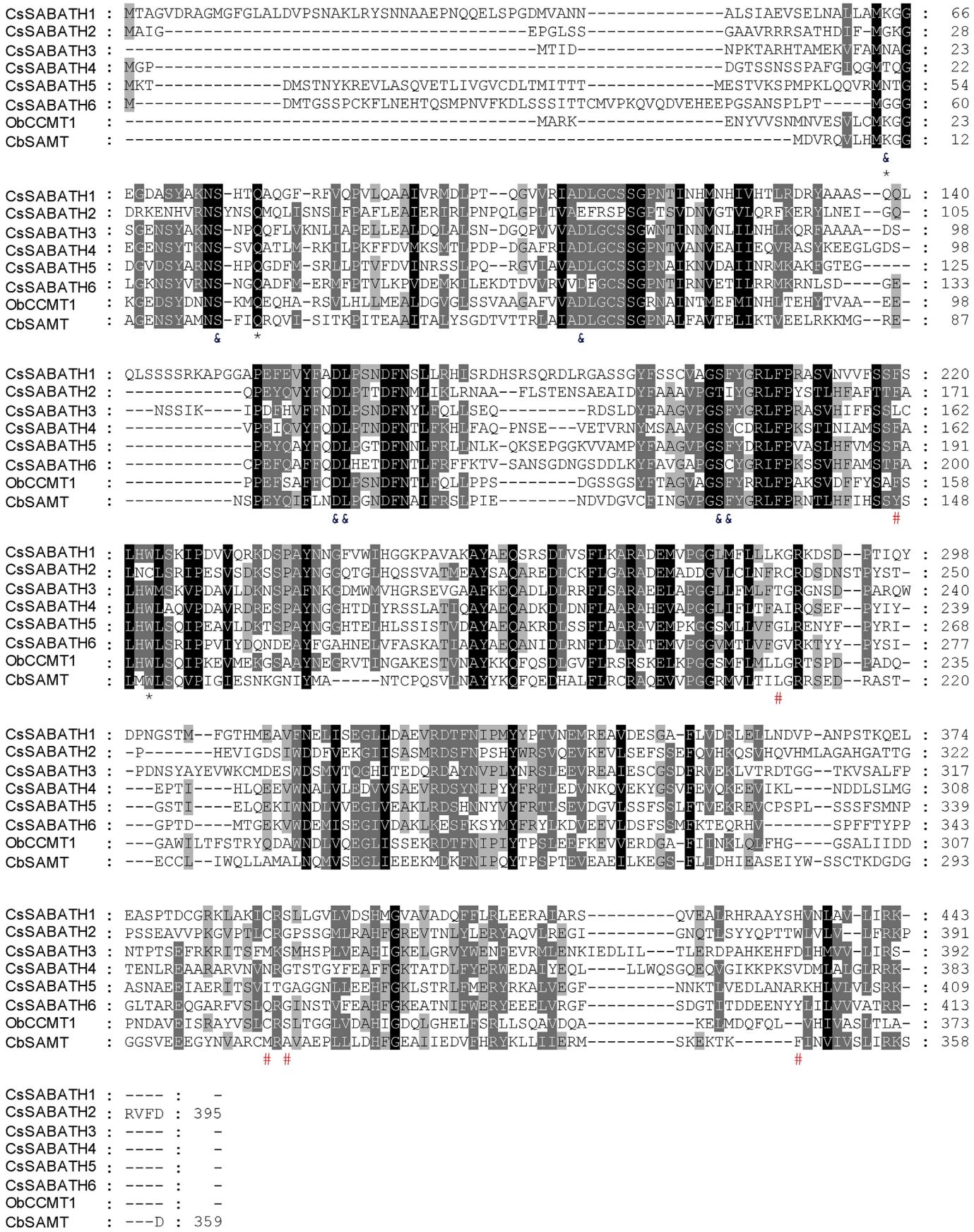

Fig. 2. Sequence alignment of CcSABATHs with ObCCMT1 and CbSAMT. Conserved residues are in shade with the more conserved the darker. Residues indicated with "&" are SAM/SAH-binding residues. Residues indicated with "*" are residues that interact with the carboxyl moiety of substrate. Residues indicated with "#" interact with the aromatic moiety of substrate and are important for substrate selectivity. CcSABATHs, *C. conicum sensu latu* SABATHs; ObCCMT1, *Ocimum basilicum* cinnamate/*p*-coumarate carboxyl methyltransferase 1; CbSAMT, *Clarkia breweri* salicylic acid methyltransferase.

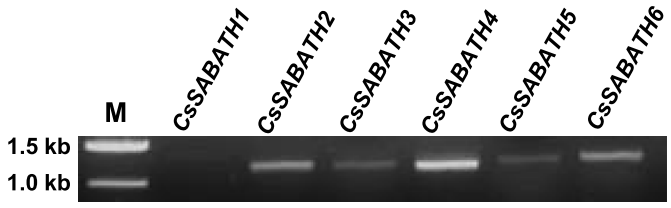

Fig. 3. Expression of six *CsSABATHs* genes in sterile thalli measured by semi-quantitative RT-PCR. Same amount of cDNA was used in each reaction. This experiment was repeated three times with similar results. "M" indicates molecular marker.

PCR reactions, transcripts for five of the six genes could be detected from thallus tissue, the exception being *CsSABATH1* (Fig. 3). *CsSABATH4* showed the highest level of expression. While *CsSABATH2* and *CsSABATH6* showed also higher expression levels. *CsSABATH3* and *CsSABATH5* were found to be expressed at relatively low levels (Fig. 3).

2.3. Biochemical characterization of CsSABATHs expressed in E. coli

Gene expression analysis (Fig. 3) suggested that *CsSABATH2*, *CsSABATH4* and *CsSABATH6* are the most probable candidates as *CAMT*. To validate this prediction, we characterized the proteins encoded by all six *CsSABATH* genes using a biochemical approach. Full-length cDNA for each of the six *CsSABATH* genes was cloned by RT-PCR into a protein expression vector without any tag and expressed in *E. coli* to produce recombinant proteins. Each of the recombinant CsSABATH enzymes was assayed with (*E*)-cinnamic acid and an additional nine carboxyl acids: indole-3-acetic acid, gibberellin A3, salicylic acid, jasmonic acid, anthranilic acid, nicotinic acid, benzoic acid, *p*-coumaric acid and abscisic acid (Table 1). Among the six CsSABATHs, CsSABATH6 was highly active using (*E*)-cinnamic acid as substrate. It also showed activity with benzoic acid and *p*-coumaric acid, which was about 32% and 18% of the activity with (*E*)-cinnamic acid. CsSABATH4 also showed some activity with (*E*)-cinnamic acid, but the activity was very low. Instead, CsSABATH4 was most active with salicylic acid as substrate. Its activity with (*E*)-cinnamic acid was only about 6% of that with salicylic acid. CsSABATH6 and CsSABATH4 were renamed CsCAMT (GenBank accession no. MK673137) and CsSAMT (GeneBank accession no. MK673138), respectively. CsSABATH1, CsSABATH2, CsSABATH3 and CsSABATH5 did not show activity with any of the substrates tested.

2.4. Biochemical properties of CsCAMT

With the identification of CsCAMT in *C. salebrosum*, next, we performed biochemical assays to determine its detailed biochemical properties using a purified CsCAMT containing a N-terminal His-tag. The optimum pH for CsCAMT was determined to be 6.5 (Fig. 4A). For temperature stability, CsCAMT was relatively stable when the assay temperatures were below 42 °C (Fig. 4B). CsCAMT activity can be affected by various ions (Fig. 4C). Na^+ and Mg^{2+} ions slightly stimulated the activity of CsCAMT. K^+, NH_4^+, Ca^{2+} and Mn^{2+} had a mild inhibitory effect. Fe^{2+}, Zn^{2+}, Cu^{2+} or Fe^{3+} completely inhibited the activity of CsCAMT (Fig. 4C). The kinetic property of CsCAMT was also measured. The apparent *Km* value of CsCAMT using cinnamic acid as substrate was determined to be 50.5 μM (Fig. 4D). The product of CsCAMT was verifed to be methyl (*E*)-cinnamate (see Fig. 5).

Table 1
The relative activities of CsSABATH4 and CsSABATH6 from *C. salebrosum* with ten carboxyl acids.

	CsSABATH4 (CsSAMT)	CsSABATH6 (CsCAMT)
(*E*)-cinnamic acid	6%[a]	**100%**
p-Coumaric acid	4%	18%
Nicotinic acid	0%	3%
Anthranilic acid	12%	2%
Benzoic acid	51%	32%
Salicylic acid	**100%**	0%
Abscisic acid	0%	0%
Jasmonic acid	1%	0%
Gibberellic acid	0	0%
Indole-3-acetic acid	5%	3%

[a] Values are averages of three independent measurements using radiochemical assays. The final concentration of each substrates tested was 1 mM. The highest level of activity was arbitrarily set at 100%.

2.5. Comparison of CAMT genes in different chemotypes of the C. conicum complex

As previously reported, some populations of *C. conicum* showed a monoterpene-dominant chemical profile (Craft et al., 2016; Ludwiczuk et al., 2013; Toyota, 2000). To gain some understanding into the molecular basis of such variations, we collected a population of *C. salebrosum* from Tennessee (TN). This population exhibited sabinene-dominant chemotype (Fig. S1). Next, a putative ortholog of *CsCAMT* gene was cloned from the sabinene-dominant plants and fully sequenced. It is identical to the *CsCAMT* gene isolated from methyl (*E*)-cinnamate-dominant plants.

2.6. Structural modeling of CsCAMT, CsSAMT and ObCCMT1

The structure models for CsCAMT, CsSAMT and one basil CAMT (ObCCMT1) (Kapteyn et al., 2007) were built using the X-ray structure of CbSAMT (Zubieta et al., 2003) as the template. The active sites of the models for CsCAMT and CsSAMT in complex with (*E*)-cinnamic acid are plotted in Fig. 6A and B, respectively. In each of these cases, the model was first superposed with the protein structure of CbSAMT that is complexed with salicylic acid and *S*-adenosyl-L-homocysteine (SAH) at the active site. (*E*)-cinnamic acid was then docked into the active site such that its cinnamate carboxylate group was superposed with that of salicylic acid to form the reactive configuration for the methyl transfer. Fig. 6A shows that (*E*)-cinnamic acid can fit well into the active site of CsCAMT with its carboxyl moiety located at the suitable position for accepting the methyl group from SAM. The result is consistent with the experimental observation that (*E*)-cinnamic acid is the preferred substrate for CsCAMT. In contrast, when (*E*)-cinnamic acid was docked into the active site of CsSAMT (Fig. 6B) in the similar fashion, the ring of (*E*)-cinnamic acid becomes too close to some of the residues of CsSAMT. For instance, the distances between two of the carbon atoms on the ring are only about 2.4 Å and 2.1 Å to the C_α and C_β carbons of Ile237, respectively. Such close contacts would create significant repulsions between the enzyme and (*E*)-cinnamic acid and is likely to destroy the reactive configuration for the methyl transfer.

The active sites of the models for CsCAMT and CsSAMT in complex with salicylate are shown in Fig. 6C and D. Fig. 6D shows that, unlike (*E*)-cinnamic acid, salicylic acid can fit into the active site of CsSAMT without the close contacts observed for (*E*)-cinnamic acid. This is expected, as salicylic acid is significantly smaller than (*E*)-cinnamic acid. Trp159 of CsCAMT (Fig. 6C) and Trp157 of CsSAMT (Fig. 6D) are equivalent to Trp151 of CbSAMT, one of the two key residues for ensuring proper orientation and proximity of salicylic acid to SAM for the methylation reaction in CbSAMT (Zubieta et al., 2003). It is of interest to note from Fig. 6C that, unlike in CbSAMT (Trp151) and CsSAMT (Trp157), Trp159 in CsCAMT seems to have the potential to interact not only with the carboxyl group of salicylic acid, but also with the 2-OH group. Such additional interaction with the 2-OH group might distort the optimal arrangement between the methyl acceptor and donor and therefore interferes with the methyl transfer. Consistent with this suggestion, Table 1 shows that CsCAMT does not have activity on salicylic acid, even though it is active on benzoic acid.

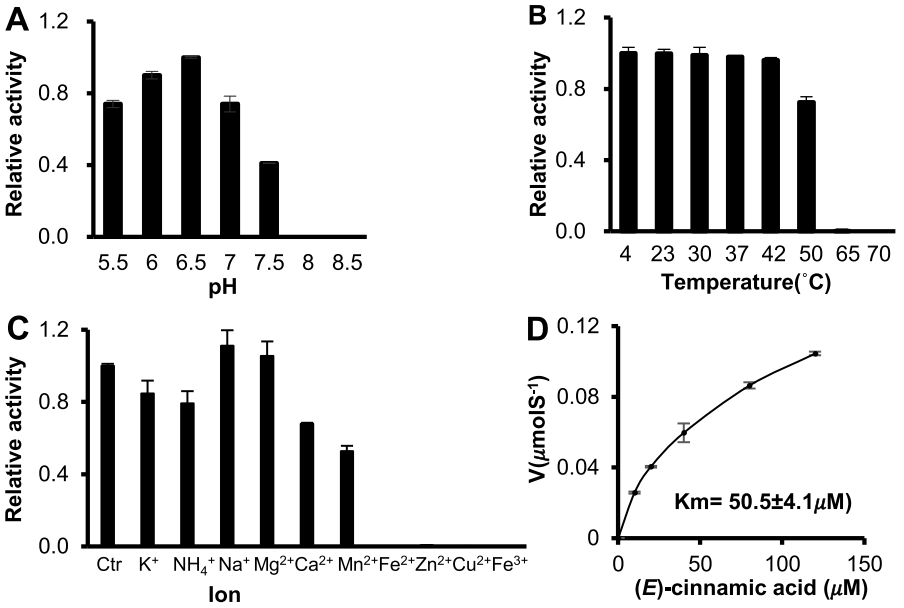

Fig. 4. Biochemical properties of CsCAMT using (*E*)-cinnamic acid as substrate. (A) pH optimum. Level of CsCAMT activity in the buffer of pH 6.5 was arbitrarily set at 1.0. (B) Thermostability. The activity of CsCAMT incubated at 4 °C for 30 min was arbitrarily set at 1.0. (C) Effect of ions. Individual ions were added to reactions in the form of chloride salts at final concentration of 5 mM. Level of activity without any ion added served as a control (Ctr) and was arbitrarily set at 1.0. (D) Steady-state kinetic measurement of CsCAMT. Each value was the mean of three independent measurements and bar indicates standard deviation.

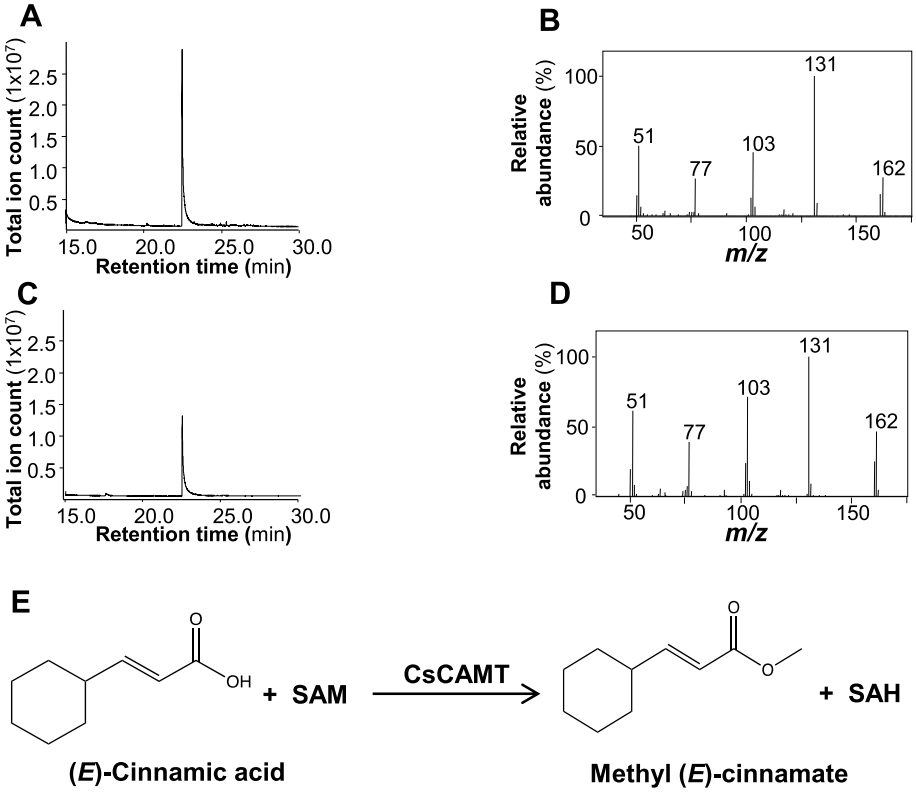

Fig. 5. Authentic methyl (*E*)-cinnamate and the product of CsCAMT methylation. (A) GC chromatogram of authentic methyl (*E*)-cinnamate; (B) Mass spectrum of authentic methyl (*E*)-cinnamate; (C) GC chromatogram of the product of CsCAMT; (D) Mass spectrum of the product of CsCAMT; (E) Schematic methylation reaction catalyzed by CsCAMT. SAM stands for S-adenosyl-L-methionine. SAH stands for *S*-adenosyl-L-homocysteine.

In Fig. 6E, the structure model for ObCCMT1 complexed with (*E*)-cinnamic acid is plotted on the left, and the superposition of the active sites for CsCAMT and ObCCMT1 is shown in Fig. 6F. Similar to the case of CsCAMT, (*E*)-cinnamic acid can fit into the active site of ObCCMT1 without close contacts with the active site residues (Fig. 6E), in agreement with an earlier study on this enzyme (Kapteyn et al., 2007). It is of interest to note from Fig. 6F that a number of the active site residues for ObCCMT1/CsCAMT are conserved in these two enzymes (e.g., Tyr29/Tyr20, Gln36/Gln27, Asp68/Asp59, Ser72/Ser63, Asn76/Asn67, Asp108/Asp98, Leu109/Leu99, Tyr141/Tyr139, Phe157/Phe155, His160/His158, Trp161/Trp159 and Tyr272/Tyr268).

2.7. Relatedness of CsCAMT with other SABATH methyltransferases

An approximately-maximum-likelihood phylogenetic tree was constructed using the six CsSABATHs, three ObCCMTs and SABATH proteins from four sequenced plants that include *Arabidopsis thaliana*, *Oryza sativa*, *Physcomitrella patens* and *Selaginella moellendorffii* to understand their evolutionary relatedness (Fig. 7). All SABATHs in seed plant except gibberellic acid methyltransferases from *A. thaliana* (AtGAMTs) resolved in a single clade. ObCCMTs from basil resolved in a clade with IAMTs from *A. thaliana* (AtIAMT) and *O. sativa* (OsIAMT). CsCAMT and CsSAMT resolved with other uncharacterized CsSABATHs from *C. salebrosum* and together they are resolved in a clade with

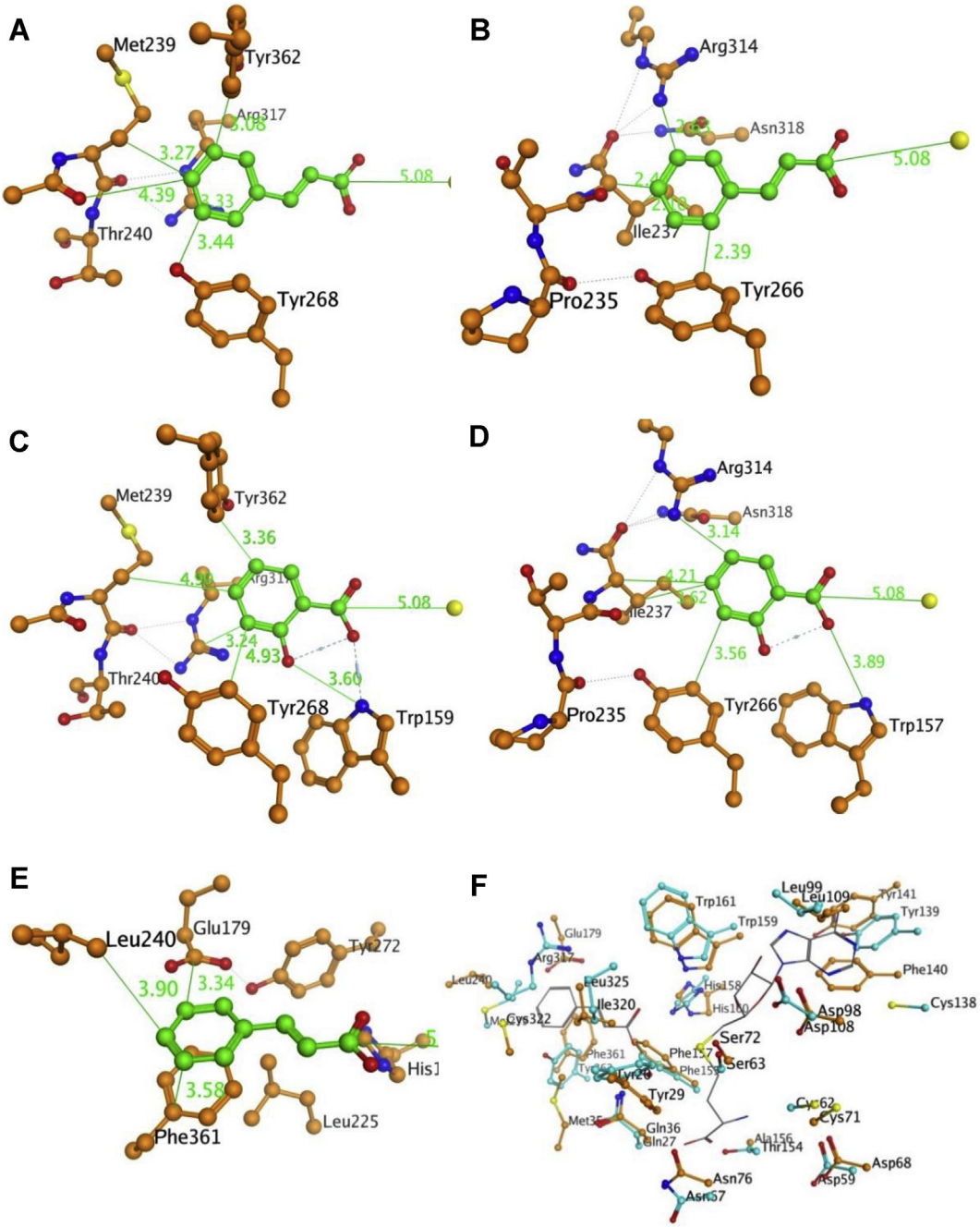

Fig. 6. Active sites of the structure models of CsCAMT, CsSAMT and ObCCMT1 with (*E*)-cinnamic acid or salicylic acid docked into the active sites. (A) CsCAMT (orange) complexed with (*E*)-cinnamic acid (green); the sulfur atom from SAH is shown as a yellow ball. Distances between the substrate and enzyme/SAH atoms are given in Å. (B) CsCAMT complexed with (*E*)-cinnamic acid. (C) CsCAMT complexed with salicylic acid. (D) CsSAMT complexed with salicylic acid. (E) ObCCMT1 complexed with (*E*)-cinnamic acid. (F) Superposition of ObCCMT1 (orange) and CsCAMT (cyan) with some active-site residues shown. (*E*)-cinnamic acid and SAH are shown in line except for the sulfur atom. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

AtGAMTs. Surprisingly, liverwort SABATHs do not resolve with the other major SABATHs in seed plants, including IAMTs and CCMTs.

3. Discussion

With the isolation and characterization of six *SABATH* genes from *C. salebrosum*, this is the first report on functional study of *SABATH* genes in liverworts. Particularly, we successfully identified *CsCAMT* (*CsSABATH6*) as the gene with the capability for synthesis of methyl (*E*)-cinanmate in *C. salebrosum* (Table 1). In addition, *CsSABATH4* was determined to encode SAMT (Table 1). CsCAMT has a higher affinity with the substrate (*E*)-cinnamic acid than ObCCMT1 from basil (Kapteyn et al., 2007). Its Km value (50.5 μM) is about the half of that for ObCCMT1 (124 μM). CsCAMT and ObCCMT1 also showed different selectivity towards *p*-coumaric acid and benzoic acid. ObCCMT1 had about 30% and 10% of activities of (*E*)-cinnamic acid with *p*-coumaric acid and benzoic acid, respectively (Kapteyn et al., 2007). CsCAMT exhibited the opposite trend: its activities with *p*-coumaric acid and benzoic acid were 18% and 32% respectively of that with (*E*)-cinnamic acid (Table 1). The identification of CsCAMT and CsSAMT (CsSABATH4) together with the availability of basil ObCCMT1 makes it possible to understand the structural basis underlying their substrate

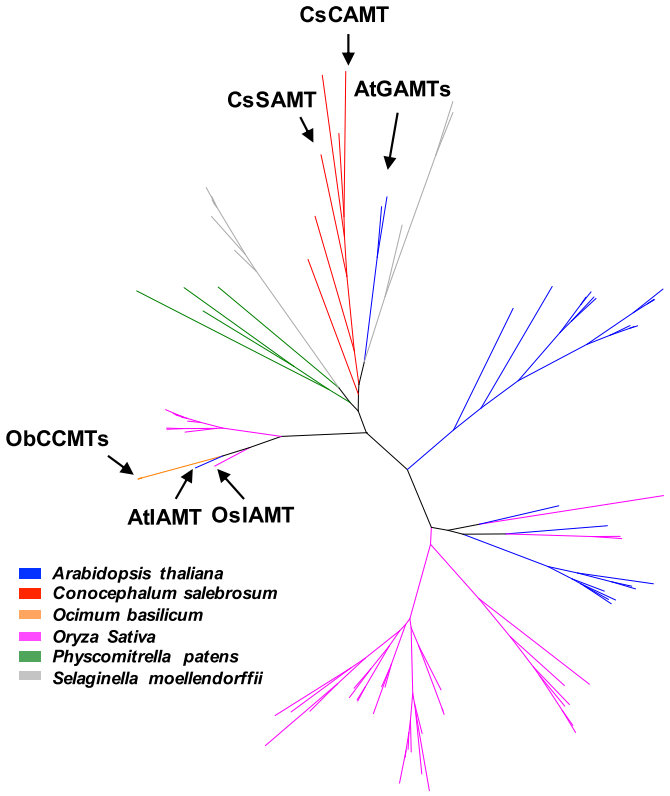

Fig. 7. Phylogenetic tree of the SABATH family of methyltransferases in *C. salebrosum*, *Arabidopsis thaliana*, *Oryza sativa*, *Physcomitrella patens*, *Selaginella moellendorffii* and ObCCMTs. The SABATH proteins from each species were color-coded. CCMTs: Cinnamate/*p*-coumarate methyltransferases; CAMT: cinnamic acid methyltransferase; IAMT: indole-3-acetic acid methyltransferase; SAMT: salicylic acid methyltransferase. AtIAMT, AtGAMTs, CsCAMT, CsSAMT, OsIAMT and ObCCMTs were indicated with arrows. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

specificities.

We demonstrated from computer modeling that while the active site of CsCAMT can accommodate (*E*)-cinnamic acid and generate the reactive configuration for the methyl transfer (Fig. 6A), CsSAMT may not be able to do so. This is because some of the active site residues are too close to the ring of (*E*)-cinnamic acid (Fig. 6B), and such close contacts would create significant repulsions between CsSAMT and (*E*)-cinnamic acid, leading to the non-reactive configuration for the methyl transfer. This suggestion is supported by the lack of the activity of CsSAMT on (*E*)-cinnamic acid (see Table 1). Similar discussion has been made in our earlier quantum mechanical/molecular mechanical (QM/MM) study for understanding the substrate specificity of CbSAMT (Yao et al., 2010). Interestingly, as observed in other characterized SABATH methyltransferases, our results on ObCCMT1/CsCAMT (Fig. 6F) show that both SAM/SAH-binding residues and carboxyl moiety interacting residues are very conservative, but residues that interact with the aromatic moiety of the substrate are different (Fig. 2). This indicates that similar interaction between active sites and substrate is achieved by the combination of different residues.

As indicated by our phylogenetic study (Fig. 7), ObCCMTs share a common ancestor with IAMTs from Arabidopsis and rice. Although IAMT in basil has not been identified, our results suggest that ObCCMTs and IAMTs, which belong to an ancient SABATH group in seed plants evolved from a common ancestral type. However, CsCAMT together with other CsSABATH members share a common ancestor with *Arabidopsis* GAMTs, which is beyond our expectation. This demonstrates that CsCAMT may have evolved from other uncharacterized SABATH methyltransferases and emerged independently from ObCCMT in basil. The phylogeny in combination with structural modeling implies that the evolution of methyl (*E*)-cinnamate synthesis in the liverwort and flowering plants is the result of convergent evolution.

C. conicum and *C. salebrosum* usually inhabit shady moist environments, which can be hotbeds for microbe proliferation. Several studies showed that methyl (*E*)-cinnamate has antimicrobial activity (Gilles et al., 2010; Padalia et al., 2017). This could be a good reason for some populations of *C. salebrosum* steadily producing volatile methyl (*E*)-cinnamate in relatively large amounts (Fig. 1B). Aside from the capability of inhibiting the growth of fungi and bacteria, methyl (*E*)-cinnamate has larvicidal activity (Fujiwara et al., 2017). Larvae of the moth *Epimartyri pardella* have been reported to feed on *C. conicum* (Davis and Landry, 2012). This can be another rational cause for liverwort to defend itself from herbivores. Ghani et al. (2016) have shown that production of methyl (*E*)-cinnamate increases in populations of *C. conicum* from Japan when they are exposed to stressful growth conditions, suggesting that gene expression may be influenced by environment, while Szweykowski et al. (2005) note that *C. salebrosum* often grows in habitats that vary from flooded to seasonally dry conditions.

The *in vitro* product of CsSAMT, methyl salicylate, has been shown to be a critical signaling molecule in plant defenses, particularly as the mobile signal in systematic acquired resistance against pathogens (Park et al., 2007; Shulaev et al., 1997). It may also affect plant defense against insects in both direct and indirect ways (Ament et al., 2010; Zhu and Park, 2005). The identification of CsSAMT indicates that methyl salicylate might play similar roles liverworts. As in the case of CsCAMT in the phylogenetic analysis, CsSAMT is more closely related to other CsSABATH members rather than to other characterized SAMTs or BSMTs (benzoic acid/salicylic acid MT) in seed plants. This observation implies that SAMTs might have evolved independently in liverworts and in flowering plants as well.

Besides methyl (*E*)-cinnamate, *C. salebrosum* produces a number of terpenoids (Fig. 1). Some progress has been made recently in our understanding of terpenoid biosynthesis in several species of bryophytes (Jia et al., 2016; Kumar et al., 2016a; Xiong et al., 2018), in which monoterpenes and sesquiterpenes are synthesized by a newly identified type of terpene synthase genes called MTPSL (Jia et al., 2018). It will be interesting to determine whether the same type of enzyme is responsible for the formation of monoterpenes and sesquiterpenes in *C. salebrosum* and what biological function these terpenoids have.

4. Conclusions

In this study, we identified and characterized the enzyme CsCAMT that is capable for the biosynthesis of methyl (*E*)-cinnamate in the liverwort *C. salebrosum*, a species in the *C. conicum* complex. CsCAMT belongs to the SABATH family of methyltransferases. The *SABATH* family in *C. salebrosum* has at least nine members based on transcriptome analysis, six of which were judged to be full-length and selected for characterization in this study. It is evident that members of the SABATH family in *C. salebrosum* have different substrate specificities, which is similar to the SABATH family in flowering plants such as *Arabidopsis* (Varbanova et al., 2007; Yang et al., 2006), rice (Zhao et al., 2008) and gymnosperms (Chaiprasongsuk et al., 2018). This implies repeated gene duplication followed by functional divergence of the SABATH family. Selected members of the *C. salebrosum* SABATH family are not conserved with their counterparts in flowering plants, evidenced by CAMT and SAMT in this study (Fig. 7). Particularly for CAMT, we can infer that this enzyme in *C. salebrsum* and in flowering plants evolved independently. There are strong similarities between the active site of CsCAMT and that of ObCCMT1, despite their relatively distant relatedness among the SABATH family, indicating similarities in catalytic mechanisms. There are a number of interesting future directions. With CsCAMT and ObCCMT1 resulting from convergent evolution, it will be

interesting to ask whether they evolved from similar or different ancestral activities. Because the two chemotypes of *C. salebrosum* contain the same *CAMT* gene, it will be interesting to ask what causes the difference in the contents of methyl (*E*)-cinnamate: different expression levels of CsCAMT or the different concentrations of its substrate cinnamic acid or both? With methyl (*E*)-cinnamate accumulated in the thallus and also emitted as a volatile compound, it will also be interesting to ask what are the biological functions CsCAMT and its product methyl (*E*)-cinnamate. In addition, for the four CsSABATHs that did not show activity with any of the ten carboxylic acids tested, their *in vivo* substrates and biological functions can be an important future study.

5. Experimental

5.1. Plant culture and axenic culture

Two populations of *C. salebrosum* Szweyk., Buczkowska & Odrzykoski (Conocephalaceae) were used this study. One population was collected from Illinois: Williamson Co., Rocky Bluff Nature Preserve, nr. Devils Kitchen Lake, growing over moist sandstone rocks; 37°38′32.34″ N, 89°05′51.25″ W; elevation 157 m, 16 Oct. 2017, *J. Henry s.n.* [F]). This population is methyl (*E*)-cinnamate-dominant. The other population was collected from Tennessee: Knox Co., Campbell station park, nr. North Fork Turkey Creek, growing over moist rocks; 35°53′13.5″N 84°10′02.4″W and is sabinene-dominant. The Illinois plants collected from the field were rinsed with running water until thalli were clean. Explants were sterilized using 70% ethanol solution for 2 min, then rinsed 3 times with sterilized distilled water. The explants were further surface-sterilized using 10% bleach (v/v) for 5 min, then rinsed 3 times using sterilized distilled. The sterilized explants were transferred onto Hatcher medium which was prepared following the protocol of Dr. Crandall-Stotler's Lab (http://bryophytes.plant.siu.edu).

5.2. Organic extraction and headspace collection

Thallus materials of axenically grown and field-collected *C. salebrosum* were ground in liquid nitrogen and subject to organic extraction using ethyl acetate as solvent as previously described (Chen et al., 2003b). For volatile profiling, a solid phase microextraction (SPME) fiber coated with 100-μm polydimethylsiloxane was inserted into the headspace of a *C. salebrosum* culture plate to start volatile collection. After 1.5 h, the SPME fiber was retracted and inserted into the injector port of GC for compound separation and identification.

5.3. GC-MS analysis

Samples from either organic extraction or headspace collection were analyzed with a Shimadzu 17A gas chromatograph coupled to a Shimadzu QP5050A quadrupole mass selective detector. The GC conditions included splitless injection, Restek R xi-5Sil MS column (Restek, Bellefonte, PA), Helium as the carrier gas and a temperature gradient of 5 °C/min from 40 °C to 240 °C after an initial 3-min hold. Compounds were identified based on comparisons to authentic standards when available or the NIST library.

5.4. Database search and phylogenetic analysis

To identify putative *SABATH* genes from the *C. conicum* complex, the protein sequences of sample ID ILBQ from OneKP database (Matasci et al., 2014) were searched using Pfam profile PF03492 (Finn et al., 2015) via the hmmsearch (Finn et al., 2011). For phylogeny reconstruction, MAFFT (version 7.369b, under L-INS-I strategy) was used to perform multiple protein sequence alignments (Katoh and Standley, 2013). FastTree (version 2.1.10, under the JTT + CAT model with 1000 resamples) was used to construct Approximately-Maximum-Likelihood trees (Price et al., 2010). The trees were further edited using MEGA (version 7.0.21) (Kumar et al., 2016b).

5.5. Cloning full length cDNA of CsSABATHs

Total RNA was extracted from vegetative thalli using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA) and reverse-transcribed into first strand cDNA in a 15 μL reaction volume using the First-strand cDNA Synthesis Kit (Amersham Biosciences, Piscataway, NJ) as previously described (Chen et al., 2003a). *CsSABATH* full-length cDNAs were amplified using forward primer and reverse primers corresponding with the 5′ and 3′ ends of the *CsSABATH* coding region (Table S2), respectively. The PCR was carried out using the following program: 94 °C for 3 min followed by 32 cycles at 94 °C for 30 s, 58 °C for 30 s, 72 °C for 1 min 30 s, and followed by a final extension at 72 °C for 10 min. PCR products were isolated on 1.0% agarose gel and purified using QIAquick Gel Extraction kit (Qiagen, Valencia, CA). The cDNAs were then cloned into the vector pEXP-5-CT/TOPO following the vendor's protocol (Invitrogen, Carlsband, CA). The cloned cDNAs were fully sequenced.

5.6. Semi-quantitative reverse-transcription PCR

Semi-quantitative reverse-transcription PCR (RT-PCR) was performed as previously described (Chen et al., 2003a). The cDNA from a single reverse-transcription reaction was used as template and 6 pairs of specific primers (Table S2) were used in the PCR reaction. PCR was performed using the same program as described in Section 5.5 except that the cycle number was 25. 30 μL PCR products and 3 μL 1 kb ladder marker were loaded in the agarose gel.

5.7. Purification of CsSABATHs expressed in E. coli

CsSABATHs were subcloned into the vector of pET32a (Invitrogen, Carlsband, CA). To express the CsSABATH proteins, the protein expression constructs were transformed into the *E. coli* strain BL21 (DE3) CodonPlus (Stratagene, La Jolla, CA) and cultured under 25 °C. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added into *E. coli* cultures at a concentration of 500 μM to induce protein expression. After 18 hours of cultivation at 25 °C, the bacterial cells were harvested and lysed by sonication. His-tagged CsSABATH proteins were enriched from the *E. coli* cell lysate using Ni-NTA (Invitrogen, Carlsband, CA) following the manufacturer's protocol. An empty pET32a vector without any insert was set to be a negative control.

5.8. Radiochemical methyltransferase activity assay

CsSABATH enzyme assays including the measurements of optimum pH, thermostability, the effect of cations and kinetic parameters were conducted using a radiochemical protocol as previously described (Zhao et al., 2013). A non-radiochemical assay followed by extraction and GC-MS analysis was performed to determine the chemical identity of the product of CsSABATH methylation.

5.9. Protein structure modeling

Based on the crystallographic structure (PDB code: 1M6E) of CbSAMT, the homology models of CsCAMT, CsSAMT and ObCCMT1 were built by using the tools in the MOE program [Molecular Operating Environment (MOE), version 2013.08 (2015) Chemical Computing Group Inc., Montreal]. SAH in each of the models was built based on the superposition of the model with the X-ray structure of the CbSAMT complex containing both SAH and SA. (*E*)-cinnamic acid molecule was manually docked into the active site in each case based on the superposition of the (*E*)-cinnamic acid carboxylate group with that of SA in the CbSAMT complex that is close to the sulfur atom of SAH. Such correct binding mode of the carboxyl moiety in the enzyme-substrate

complex is expected to be necessary for the enzyme's function as methyltransferases (Yao et al., 2010, 2012). The similar idea has been used in our earlier study of other plants (Qian et al., 2015).

Acknowledgements

Chi Zhang has been partly supported by a scholarship from the China Scholarship Council. We thank the 1 KP initiative for the availability of *Conocephalum conicum* transcriptome data. Ping Qian has been supported in part by the Natural Science Foundation of Shandong Province of China (No. ZR2017MB048).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phytochem.2019.04.013.

References

Ali, N., Rahmani, M., Shaari, K., Ali, A., Cheng Lian, G., 2010. Antimicrobial activity of *Cinnamomum impressicostatum* and *C. pubescens* and bioassay-guided isolation of bioactive (E)-methyl cinnamate. J. Biol. Sci. 10, 101–106.

Ament, K., Krasikov, V., Allmann, S., Rep, M., Takken, F.L., Schuurink, R.C., 2010. Methyl salicylate production in tomato affects biotic interactions. Plant J. 62, 124–134.

Arai, T., 1994. Volatile compounds of *Narcissus taztta* var. *chinensis* flowers. Nippon Koryo Kyokai 184, 105–111.

Atherton, I., Bosanquet, S., Lawley, M., 2010. Mosses and Liverworts of Britain and Ireland: a Field Guide. British Bryological Society, Plymouth UK.

Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S.S., Althoff, F., Araki, T., Arteaga-Vazquez, M.A., Balasubrmanian, S., Barry, K., Bauer, D., Boehm, C.R., Briginshaw, L., Caballero-Perez, J., Catarino, B., Chen, F., Chiyoda, S., Chovatia, M., Davies, K.M., Delmans, M., Demura, T., Dierschke, T., Dolan, L., Dorantes-Acosta, A.E., Eklund, D.M., Florent, S.N., Flores-Sandoval, E., Fujiyama, A., Fukuzawa, H., Galik, B., Grimanelli, D., Grimwood, J., Grossniklaus, U., Hamada, T., Haseloff, J., Hetherington, A.J., Higo, A., Hirakawa, Y., Hundley, H.N., Ikeda, Y., Inoue, K., Inoue, S.-I., Ishida, S., Jia, Q., Kakita, M., Kanazawa, T., Kawai, Y., Kawashima, T., Kennedy, M., Kinose, K., Kinoshita, T., Kohara, Y., Koide, E., Komatsu, K., Kopischke, S., Kubo, M., Kyozuka, J., Lagercrantz, U., Lin, S.-S., Lindquist, E., Lipzen, A.M., Lu, C.-W., De Luna, E., Martienssen, R.A., Minamino, N., Mizutani, M., Mizutani, M., Mochizuki, N., Monte, I., Mosher, R., Nagasaki, H., Nakagami, H., Naramoto, S., Nishitani, K., Ohtani, M., Okamoto, T., Okumura, M., Phillips, J., Pollak, B., Reinders, A., Rövekamp, M., Sano, R., Sawa, S., Schmid, M.W., Shirakawa, M., Solano, R., Spunde, A., Suetsugu, N., Sugano, S., Sugiyama, A., Sun, R., Suzuki, Y., Takenaka, M., Takezawa, D., Tomogane, H., Tsuzuki, M., Ueda, T., Umeda, M., Ward, J.M., Watanabe, Y., Yazaki, K., Yokoyama, R., Yoshitake, Y., Yotsui, I., Zachgo, S., Schmutz, J., 2017. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. Cell 171, 287–304.

Chaiprasongsuk, M., Zhang, C., Qian, P., Chen, X., Li, G., Trigiano, R.N., Guo, H., Chen, F., 2018. Biochemical characterization in Norway spruce (*Picea abies*) of SABATH methyltransferases that methylate phytohormones. Phytochemistry 149, 146–154.

Chen, F., D'auria, J.C., Tholl, D., Ross, J.R., Gershenzon, J., Noel, J.P., Pichersky, E., 2003a. An *Arabidopsis thaliana* gene for methylsalicylate biosynthesis, identified by a biochemical genomics approach, has a role in defense. Plant J. 36, 577–588.

Chen, F., Tholl, D., D'auria, J.C., Faroog, A., Pichersky, E., Gershenzon, J., 2003b. Biosynthesis and emission of terpenoid volatiles from Arabidopsis flowers. Plant Cell 15, 481–494.

Chen, F., Ludwiczuk, A., Wei, G., Chen, X., Crandall-Stotler, B., Bowman, J.L., 2018. Terpenoid secondary metabolites in bryophytes: chemical diversity, biosynthesis and biological functions. Crit. Rev. Plant Sci. 37, 210–231.

Craft, J.D., Harrelson, D., Setzer, W.N., 2016. Chemotypic variation of *Conocephalum salebrosum* in the southeastern Appalachian range: a search for cryptic plant biodiversity around the Tennessee River Valley. Nat. Prod. Commun. 11, 1009–1014.

Curtis, A., Southwell, I.A., Stiff, I.A., 1990. *Eucalyptus*, a new source of E-methyl cinnamate. J. Essent. Oil Res. 2, 105–110.

D'Auria, J.C., Chen, F., Pichersky, E., 2003. The SABATH family of MTS in *Arabidopsis thaliana* and other plant species. In: In: Romeo, J.T. (Ed.), Recent Advances in Phytochemistry, vol. 37. Elsevier, Amesterdam, pp. 253–283.

Davis, D.R., Landry, J.-F., 2012. A review of the North American genus *Epimartyria* (Lepidoptera, Micropterigidae) with a discussion of the larval plastron. ZooKeys 37–83.

Dodson, C.H., Dressler, R.L., Hills, H.G., Adams, R.M., Williams, N.H., 1969. Biologically active compounds in orchid fragrances. Science 164, 1243–1249.

Eltz, T., Lunau, K., 2005. Antennal response to fragrance compounds in male orchid bees. Chemoecology 15, 135–138.

Finn, R.D., Clements, J., Eddy, S.R., 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39, W29–W37.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2015. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44, D279–D285.

Fujiwara, G.M., Annies, V., de Oliveira, C.F., Lara, R.A., Gabriel, M.M., Betim, F.C., Nadal, J.M., Farago, P.V., Dias, J.F., Miguel, O.G., Miguel, M.D., Marques, F.A., Zanin, S.M., 2017. Evaluation of larvicidal activity and ecotoxicity of linalool, methyl cinnamate and methyl cinnamate/linalool in combination against *Aedes aegypti*. Ecotoxicol. Environ. Saf. 139, 238–244.

Ghani, N.A., Ludwiczuk, A., Ismail, N.H., Asakawa, Y., 2016. Volatile components of the stressed liverwort *Conocephalum conicum*. Nat. Prod. Commun. 11, 103–104.

Gilles, M., Zhao, J., An, M., Agboola, S., 2010. Chemical composition and antimicrobial properties of essential oils of three Australian *Eucalyptus* species. Food Chem. 119, 731–737.

Gomes da Silva, M.D., Chaves das Neves, H.J., 1999. Complementary use of hyphenated purge-and-trap gas chromatography techniques and sensory analysis in the aroma profiling of strawberries (*Fragaria ananassa*). J. Agric. Food Chem. 47, 4568–4573.

Harinantenaina, L., Kida, S., Asakawa, Y., 2007. Phytochemistry of three selected liverworts: *Conocephalum conicum, Plagiochila barteri* and *P. terebrans*. Arkivoc 7, 22–29.

Hattori, M., Sakagami, Y., Marumo, S., 1992. Oviposition deterrents for the limabean pod borer, *Etiella zinckenella* (Treitschke)(Lepidoptera: Pyralidae) from *Populus nigra* L.c.v. Italica leaves. Appl. Entomol. Zool. 27, 195–204.

Hippauf, F., Michalsky, E., Huang, R., Preissner, R., Barkman, T.J., Piechulla, B., 2010. Enzymatic, expression and structural divergences among carboxyl O-methyltransferases after gene duplication and speciation in *Nicotiana*. Plant Mol. Biol. 72, 311–330.

Huang, R., Hippauf, F., Rohrbeck, D., Haustein, M., Wenke, K., Feike, J., Sorrelle, N., Piechulla, B., Barkman, T.J., 2012. Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates. Proc. Natl. Acad. Sci. U.S.A. 109, 2966–2971.

Ismail, H.M., Williams, A.A., Tucknott, O.G., 1980. The flavour components of plum. Eur. Food Res. Technol. 171, 24–27.

Jia, Q., Li, G., Köllner, T.G., Fu, J., Chen, X., Xiong, W., Crandall-Stotler, B.J., Bowman, J.L., Weston, D.J., Zhang, Y., Chen, L., Xie, Y., Li, F., Rothfels, C.J., Larsson, A., Graham, S.W., Stevenson, D.W., Wong, G.K., Gershenzon, J., Chen, F., 2016. Microbial-type terpene synthase genes occur widely in nonseed land plants, but not in seed plants. Proc. Natl. Acad. Sci. U.S.A. 113, 12328–12333.

Jia, Q., Köllner, T.G., Gershenzon, J., Chen, F., 2018. MTPSLs: new terpene synthases in nonseed plants. Trends Plant Sci. 23, 121–128.

Kaiser, R., 1993. The Scent of Orchids: Olfactory and Chemical Investigations. Elsevier Science Publishers B.V.

Kapteyn, J., Qualley, A.V., Xie, Z., Fridman, E., Dudareva, N., Gang, D.R., 2007. Evolution of cinnamate/*p*-coumarate carboxyl methyltransferases and their role in the biosynthesis of methylcinnamate. Plant Cell 19, 3212–3229.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kato, M., Mizuno, K., Crozier, A., Fujimura, T., Ashihara, H., 2000. Caffeine synthase gene from tea leaves. Nature 406, 956–957.

Knudsen, J.T., Tollsten, L., Bergström, L.G., 1993. Floral scents—a checklist of volatile compounds isolated by head-space techniques. Phytochemistry 33, 253–280.

Koeduka, T., Kajiyama, M., Suzuki, H., Furuta, T., Tsuge, T., Matsui, K., 2016. Benzenoid biosynthesis in the flowers of *Eriobotrya japonica*: molecular cloning and functional characterization of p-methoxybenzoic acid carboxyl methyltransferase. Planta 244, 725–736.

Köllner, T.G., Lenk, C., Zhao, N., Seidl-Adams, I., Gershenzon, J., Chen, F., Degenhardt, J., 2010. Herbivore-induced SABATH methyltransferases of maize that methylate anthranilic acid using S-adenosyl-L-methionine. Plant Physiol. 153, 1795–1807.

Kumar, S., Kempinski, C., Zhuang, X., Norris, A., Mafu, S., Zi, J., Bell, S.A., Nybo, S.E., Kinison, S.E., Jiang, Z., Goklany, S., Linscott, K.B., Chen, X., Jia, Q., Brown, S.D., Bowman, J.L., Babbitt, P.C., Peters, R.J., Chen, F., Chappel, J., 2016a. Molecular diversity of terpene synthases in the liverwort *Marchantia polymorpha*. Plant Cell 28, 2632–2650.

Kumar, S., Stecher, G., Tamura, K., 2016b. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Ludwiczuk, A., Odrzykoski, I.J., Asakawa, Y., 2013. Identification of cryptic species within liverwort *Conocephalum conicum* based on the volatile components. Phytochemistry 95, 234–241.

Matasci, N., Hung, L., Yan, Z., Carpenter, E.J., Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Ayyampalayam, S., Barker, M., Burleigh, J.G., Gitzendanner, M.A., Wafula, E., Der, J.P., dePamphilis, C.W., Roure, B., Philippe, H., Ruhfel, B.R., Miles, N.W., Graham, S.W., Mathews, S., Surek, B., Melkonian, M., Soltis, D.E., Soltis, P.S., Rothfels, C., Pokorny, L., Shaw, J.A., DeGironimo, L., Stevenson, D.W., Villarreal, J.C., Chen, T., Kutchan, T.M., Rolf, M., Baucom, R.S., Deyholos, M.K., Samudrala, R., Tian, Z., Wu, X., Sun, X., Zhang, Y., Wang, J., Leebens-Mack, J., Wong, G.K., 2014. Data access for the 1,000 Plants (1KP) project. GigaScience 3, 17.

McCarthy, A.A., McCarthy, J.G., 2007. The structure of two N-methyltransferases from the caffeine biosynthetic pathway. Plant Physiol. 144, 879–889.

Murfitt, L.M., Kolosova, N., Mann, C.J., Dudareva, N., 2000. Purification and characterization of S-adenosyl-L-methionine: benzoic acid carboxyl methyltransferase, the enzyme responsible for biosynthesis of the volatile ester methyl benzoate in flowers of *Antirrhinum majus*. Arch. Biochem. Biophys. 382, 145–151.

Padalia, R.C., Verma, R.S., Chauhan, A., Goswami, P., Singh, V.R., Verma, S.K., Darokar, M.P., Singh, N., Saikia, D., Chanotiya, C.S., 2017. Essential oil composition and antimicrobial activity of methyl cinnamate-Linalool chemovariant of *Ocimum basilicum* L. from India. Record Nat. Prod. 11, 193.

Park, S.-W., Kaimoyo, E., Kumar, D., Mosher, S., Klessig, D.F., 2007. Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. Science 318, 113–116.

Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490.

Qian, P., Zhao, N., Chen, F., Guo, H., 2015. Understanding substrate specificity of related plant methylesterases (MESs) from computational investigations. Chem. J. Chin. Univ. 36, 2283–2291.
Qin, G., Gu, H., Zhao, Y., Ma, Z., Shi, G., Yang, Y., Pichersky, E., Chen, H., Liu, M., Chen, Z., 2005. An indole-3-acetic acid carboxyl methyltransferase regulates Arabidopsis leaf development. Plant Cell 17, 2693–2704.
Seo, H.S., Song, J.T., Cheong, J.-J., Lee, Y.-H., Lee, Y.-W., Hwang, I., Lee, J.S., Choi, Y.D., 2001. Jasmonic acid carboxyl methyltransferase: a key enzyme for jasmonate-regulated plant responses. Proc. Natl. Acad. Sci. U.S.A. 98, 4788–4793.
Shulaev, V., Silverman, P., Raskin, I., 1997. Airborne signalling by methyl salicylate in plant pathogen resistance. Nature 385, 718.
Szweykowski, J., Buczkowska, K., Odrzykoski, I., 2005. *Conocephalum salebrosum* (Marchantiopsida, Conocephalaceae)–a new Holarctic liverwort species. Plant Systemat. Evol. 253, 133–158.
Toyota, M., 2000. Phytochemical study of liverworts *Conocephalum conicum* and *Chiloscyphus polyanthos*. Yakugaku Zasshi 120, 1359–1372.
Toyota, M., Saito, T., Matsunami, J., Asakawa, Y., 1997. A comparative study on three chemo-types of the liverwort *Conocephalum conicum* using volatile constituents. Phytochemistry 44, 1265–1270.
Varbanova, M., Yamaguchi, S., Yang, Y., McKelvey, K., Hanada, A., Borochov, R., Yu, F., Jikumaru, Y., Ross, J., Cortes, D., 2007. Methylation of gibberellins by Arabidopsis GAMT1 and GAMT2. Plant Cell 19, 32–45.
Viña, A., Murillo, E., 2003. Essential oil composition from twelve varieties of basil (*Ocimum spp*) grown in Colombia. J. Braz. Chem. Soc. 14, 744–749.
Wood, W.F., Lancaster, W.C., Fisher, C.O., Stotler, R.E., 1996. (*E*)-methyl cinnamate: the major volatile from some populations of the liverwort, *Conocephalum conicum*. Phytochemistry 42, 241–242.
Xiong, W., Fu, J., Köllner, T.G., Chen, X., Jia, Q., Guo, H., Qian, P., Guo, H., Wu, G., Chen, F., 2018. Biochemical characterization of microbial type terpene synthases in two closely related species of hornworts, *Anthoceros punctatus* and *Anthoceros agrestis*. Phytochemistry 149, 116–122.
Yang, Y., Yuan, J.S., Ross, J., Noel, J.P., Pichersky, E., Chen, F., 2006. An *Arabidopsis thaliana* methyltransferase capable of methylating farnesoic acid. Arch. Biochem. Biophys. 448, 123–132.
Yao, J., Chu, Y., An, R., Guo, H., 2012. Understanding product specificity of protein lysine methyltransferases from QM/MM molecular dynamics and free energy simulations: the effects of mutation on SET7/9 beyond the Tyr/Phe switch. J. Chem. Inf. Model. 52, 449–456.
Yao, J., Xu, Q., Chen, F., Guo, H., 2010. QM/MM free energy simulations of salicylic acid methyltransferase: effects of stabilization of ts-like structures on substrate specificity. J. Phys. Chem. B 115, 389–396.
Zhao, N., Yao, J., Chaiprasongsuk Li, G., Guan, J., Tschaplinski, T.J., Chen, F., 2013. Molecular and biochemical characterization of the jasmonic acid methyltransferase gene from black cottonwood (*Populus trichocarpa*). Phytochemistry 94, 74–81.
Zhao, N., Ferrer, J.-L., Ross, J., Guan, J., Yang, Y., Pichersky, E., Noel, J.P., Chen, F., 2008. Structural, biochemical, and phylogenetic analyses suggest that indole-3-acetic acid methyltransferase is an evolutionarily ancient member of the SABATH family. Plant Physiol. 146, 455–467.
Zhao, N., Guan, J., Lin, H., Chen, F., 2007. Molecular cloning and biochemical characterization of indole-3-acetic acid methyltransferase from poplar. Phytochemistry 68, 1537–1544.
Zhu, J., Park, K.-C., 2005. Methyl salicylate, a soybean aphid-induced plant volatile attractive to the predator *Coccinella septempunctata*. J. Chem. Ecol. 31, 1733–1746.
Zubieta, C., Ross, J.R., Koscheski, P., Yang, Y., Pichersky, E., Noel, J.P., 2003. Structural basis for substrate recognition in the salicylic acid carboxyl methyltransferase family. Plant Cell 15, 1704–1716.