Characterization of a Cytosolic Acyl-Activating Enzyme Catalyzing the Formation of 4-Methylvaleryl-CoA for Pogostone Biosynthesis in *Pogostemon Cablin*

Jing Chen[1,2], Lang Liu[1,2], Ying Wang[1,2], Zhengguo Li [1,2], Guodong Wang [3], George A. Kraus[4], Eran Pichersky [5] and Haiyang Xu [1,2,*]

[1]School of Life Sciences, Chongqing University, Chongqing 401331, China
[2]Center of Plant Functional Genomics, Institute of Advanced Interdisciplinary Studies, Chongqing University, Chongqing 401331, China
[3]State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[4]Department of Chemistry, Iowa State University, Ames, IA 50011, USA
[5]Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI 48109, USA

*Corresponding author: E-mail, hyxu@cqu.edu.cn

Pogostone, a compound with various pharmaceutical activities, is a major constituent of the essential oil preparation called Pogostemonis Herba, which is obtained from the plant *Pogostemon cablin*. The biosynthesis of pogostone has not been elucidated, but 4-methylvaleryl-CoA (4MVCoA) is a likely precursor. We analyzed the distribution of pogostone in *P. cablin* using gas chromatography-mass spectrometry (GC-MS) and found that pogostone accumulates at high levels in the main stems and leaves of young plants. A search for the acyl-activating enzyme (AAE) that catalyzes the formation of 4MVCoA from 4-methylvaleric acid was launched, using an RNAseq-based approach to identify 31 unigenes encoding putative AAEs including the *PcAAE2*, the transcript profile of which shows a strong positive correlation with the distribution pattern of pogostone. The protein encoded by *PcAAE2* was biochemically characterized *in vitro* and shown to catalyze the formation of 4MVCoA from 4-methylvaleric acid. Phylogenetic analysis showed that *Pc*AAE2 is closely related to other AAE proteins in *P. cablin* and other species that are localized to the peroxisomes. However, *Pc*AAE2 lacks a peroxisome targeting sequence 1 (PTS1) and is localized in the cytosol.

Keywords: 4-Methylvaleric acid ● 4-Methylvaleryl-CoA ● Acyl-activating enzyme ● Branched-chain fatty acid ● *Pogostemon cablin* ● Pogostone

Introduction

Pogostemonis Herba, commonly known as 'Guang-Huo-Xiang' in Chinese and patchouli in English, is produced from the dried aerial part of the plant *Pogostemon cablin* (Blanco) Benth. (Labiatae). As *Pogostemonis Herba* and its essential oil (patchouli oil) possess variety of pharmacological activities (Lu et al. 2011, He et al. 2013, Li et al. 2013, Wang et al. 2016), they have widely been used in many Asian countries as traditional Chinese medicine for the treatment of fatigue, summer heat, nausea, vomit, abdominal distension and other ailments since ancient time (Zhao et al. 2005, Chen et al. 2015, Hu et al. 2017). Patchouli oil has also been widely used for perfumes, cosmetics and food stuffs (Pattnaik et al. 1996, Lin et al. 2014, Yang et al. 2016). Pogostone (formula $C_{12}H_{16}O_4$, **Fig. 1**) is one of the major constituents of patchouli oil and is largely responsible for the intensive aromatic odor and, therefore, designated as a chemical marker for the quality control of patchouli oil (Swamy and Sinniah 2015, Yang et al. 2016). This compound has been demonstrated to exert various bioactive activities including anti-bacterial (Peng et al. 2014), anti-candida (Li et al. 2012), antifungal (Li et al. 2012, Luchesi et al. 2020), anticancer (Cao et al. 2017), insecticidal (Huang et al. 2014), anti-inflammatory (Li et al. 2014), immunosuppressive (Su et al. 2015) and gastroprotective (Chen et al. 2015) activities. The levels of pogostone in patchouli oil, mainly obtained from stems and leaves of Pogostemon plants, vary among samples collected from different geographic locations and at different periods (Luo et al. 2003, Li et al. 2004, Hu et al. 2006).

The biochemical steps involved in the biosynthesis of pogostone are currently unknown. It is likely that one or more polyketide synthases that use various starter molecules are involved (**Fig. 1A**). One obvious such starter is 4-methylvaleryl-CoA (4MVCoA), derived from the branched-chain fatty acid 4-methylvaleric acid (**Fig. 1B**). Acyl-CoA thioesters such as 4MVCoA are formed by members of the acyl-activating enzyme (AAE) superfamily that activate carboxylic acids through a
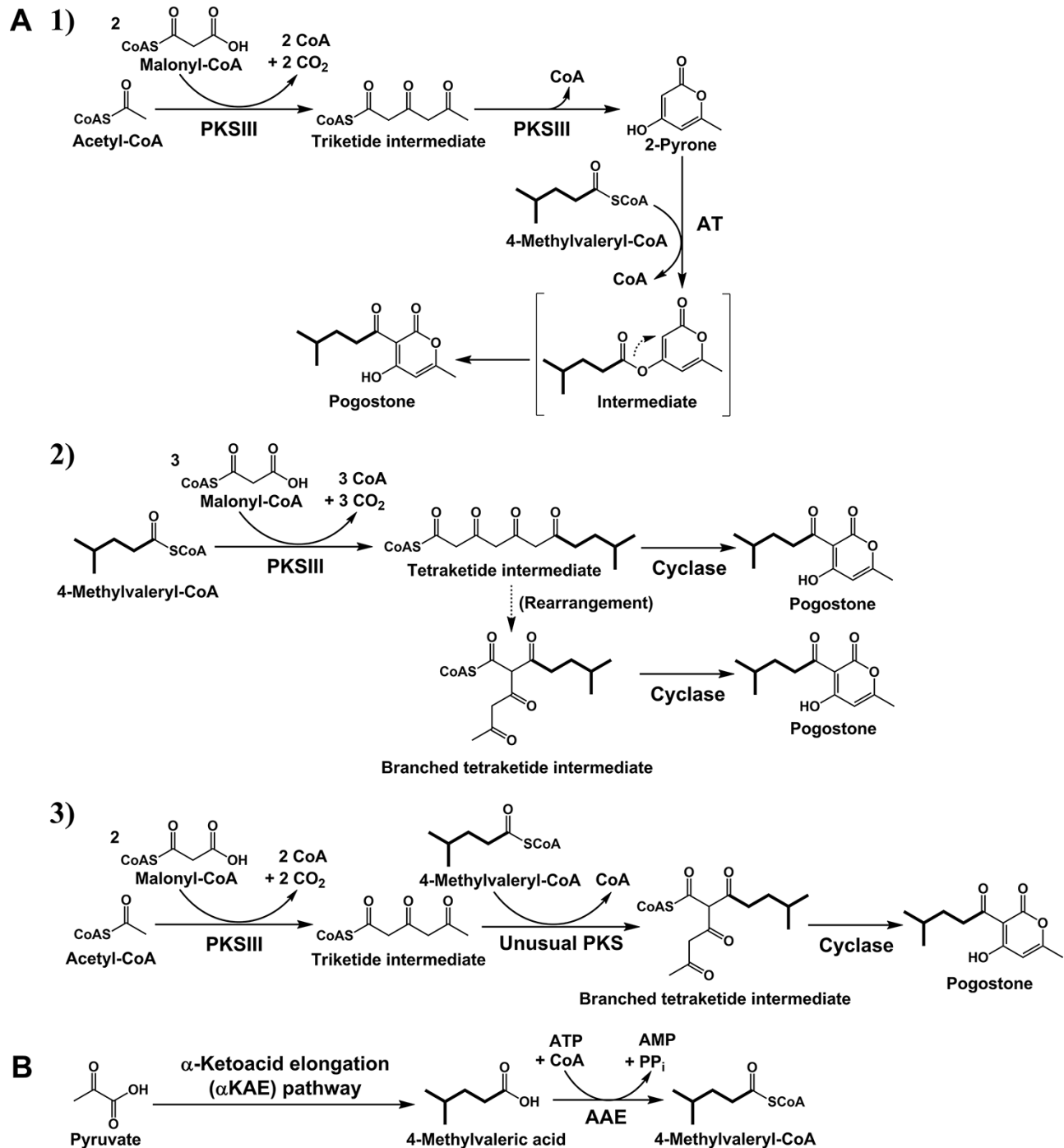
Fig. 1 The proposed biosynthetic pathway of pogostone and its 4MVCoA precursor in *P. Cablin*. A, Proposed alternative pogostone biosynthetic pathways: 1) A pathway involving one type III polyketide synthase (PKSIII) that catalyzes the condensation of acetyl-CoA and malonyl-CoA to form 2-pyrone (4-Hydroxy-6-methyl-2-pyrone), which is then acylated by one typical acyltransferase using 4MVCoA as substrate to form an unstable intermediate, followed by spontaneous rearrangement to form pogostone. 2) A pathway involving one PKSIII that catalyzes the condensation of 4MVCoA and malonyl-CoA to form a tetraketide intermediate, which is then rearranged (spontaneously or by a cyclase) and cyclized by a cyclase to form pogostone. 3) A pathway involving one typical PKSIII that catalyzes the condensation of acetyl-CoA and malonyl-CoA to form a triketide intermediate, followed with condensation catalyzed by an unusual polyketide synthase with 4MVCoA as the extension unit to form a branched tetraketide intermediate, which is then rearranged by a cyclase to form pogostone. B, Proposed biosynthetic pathway of 4MVCoA. 4-Methylvaleric acid is derived from pyruvate through α-ketoacid elongation pathway, and catalyzed by an AAE to form 4MVCoA. PKSIII, type III polyketide synthase; AT, Acyltransferase; AAE. The carbon chain of the branched short-chain fatty acid precursor 4-methylvaleric acid and 4MVCoA is labeled in bold to show its incorporation into pogostone.

two-step mechanism. First, the carboxylic acid group is adenylated to form an enzyme-bound acyl-AMP intermediate by ATP pyrophosphorylysis, and this step is followed by the displacement of AMP with a CoA group to form the relevant acyl-CoA thioester (Shockey and Browse 2011) (**Fig. 1B**). One of the unifying features of this superfamily is the well-conserved 12 amino acid residue AMP-binding motif (PROSITE PS00455), which is used to identify putative genes encoding AAEs (Shockey et al. 2003). Besides carboxylic acids, various plant hormones are also conjugated to amino acids by members of this superfamily (Nobuta et al. 2007, Okrent et al. 2009).

We have begun identifying and characterizing the enzymes involved in the biosynthesis of pogostone. The combined analysis of transcriptome and metabolome of different tissues at different developmental stages has been effectively used to discover enzymes involved in the specialized metabolic pathway in many species such as pyrethrum (Xu et al. 2018, 2019). To facilitate the elucidation of the pogostone biosynthetic pathway, a transcriptome assembly of *P. cablin* was generated in this study from RNAseq analysis of seedling, root, stem and leaf harvested at different developmental stages. GC-MS analysis established that pogostone is largely present in seedlings, the main stems and top leaves of main stem of *P. cablin* at earlier developmental stage and showed a developmentally specific accumulation pattern in top leaves of the main stem of *P. cablin*.

Here, we report the identification and characterization of the enzyme catalyzing the formation of 4MVCoA. Through transcriptome analysis of *P. cablin* RNAseq database, we identified 31 putative *AAE* genes including *PcAAE2*, the transcript profile of which shows a positive correlation with the accumulation pattern of pogostone. Through *in vitro* enzymatic assays, we identified *PcAAE2* as encoding 4MVCoA synthetase. We further showed that *Pc*AAE2 is a cytosolic enzyme that may have evolved from an ancestral peroxisome-targeted AAE.

Results

Distribution of pogostone in *P. cablin* top leaves newly generated from the main stems shows a developmental pattern

The levels of pogostone in *P. cablin* seedlings, roots, main stems and top leaves from the main stems at different developmental stages (**Fig. 2A**) were determined by gas chromatography-mass spectrometry (GC-MS) analysis of methyl tert-butyl ether (MTBE)-extracted macerated tissues. This analysis showed that pogostone is largely produced in *P. cablin* seedlings and in the main stems and the top leaves of the main stem of plants younger than 7 weeks (**Fig. 2B**). Plants older than 7 weeks contained negligible amounts of pogostone compared to younger plants (**Fig. 2B**), which indicates that the distribution of pogostone in *P. cablin* newly generated leaves from the main stem is developmentally controlled (**Fig. 2B**). Besides pogostone, another compound related to pogostone (with one carbon shorter in the side chain and likely derived from isovaleryl-CoA),
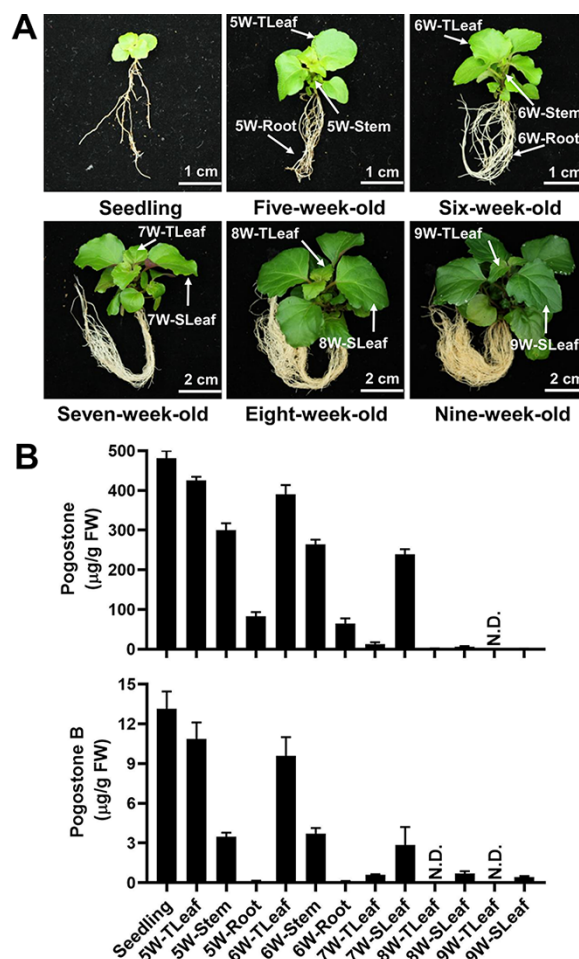

Fig. 2 Distribution analysis of pogostone and pogostone B in different tissues of *P. cablin* at different developmental stages. A, Different developmental stages of *P. cablin*. B, Concentration of pogostone and pogostone B in different tissues of *P. cablin* at different developmental stages. The organic extracts of tissues were analyzed by GC-MS and quantification was achieved by normalization of the peaks to the internal standard tetradecane and comparison with standard curves of authentic pogostone and pogostone B (Data are means ± SD of four independent biological replicates). FW, Fresh weight. The tissues used in this study are described as follows: The seedling used in this study is 4-week-old, 5W-TLeaf, the top leaves from main stem of 5-week-old *P. cablin* plant; 5W-Stem, the main stems of 5-week-old *P. cablin* plant; 5W-Root, the roots of 5-week-old *P. cablin* plant; 6W-TLeaf, the top leaves from main stem of 6-week-old *P. cablin* plant; 6W-Stem, the main stems of 6-week-old *P. cablin* plant; 6W-Root, the roots of 6-week-old *P. cablin* plant; 7W-TLeaf, the top leaves from main stem of 7-week-old *P. cablin* plant; 7W-SLeaf, the leaves from main stem of 7-week-old *P. cablin* plant are beneath and most closest to top leaves; 8W-TLeaf, the top leaves from main stem of 8-week-old *P. cablin* plant; 8W-SLeaf, the leaves from main stem of 8-week-old *P. cablin* plant are beneath and most closest to top leaves; 9W-TLeaf, the top leaves from main stem of 9-week-old *P. cablin* plant; 9W-SLeaf, the leaves from main stem of 9-week-old *P. cablin* plant are beneath and most closest to top leaves. The top leaves used in this study are two symmetrical leaves with the main vein of around 2 cm. it is noteworthy that 7W-SLeaf is converted from 6W-TLeaf after 1 week, similar cases for 8W-SLeaf converted from 7W-TLeaf and 9W-SLeaf converted from 8W-TLeaf.

Table 1 Bioinformatic analysis of the eight AAEs in the clade VI of AAE superfamily of *P. cablin*

Name	BLASTP against Arabidopsis genome (functional annotation)	Identity (%)	Peptide length (amino acids)	Signal peptide prediction (using Target P 2.0 and PTS1 Predictor)
*Pc*AAE1	AT2G17650, isovaleryl-CoA ligase[a]	67	584	mTP (0.6691), cTP (0.0015), luTP (0.0005), PTS1 (−24.573)
*Pc*AAE2	AT2G17650, isovaleryl-CoA ligase[a]	64	543	mTP (0.0021), cTP (0.0296), luTP (0.0038), PTS1 (−25.047)
*Pc*AAE3	AT3G16910, acetate/butyrate-CoA ligase[b]	78	568	mTP (0.0011), cTP (0.0083), luTP (0), PTS1 (12.027)
*Pc*AAE4	AT1G65890, AAE 12	61	590	mTP (0.0117), cTP (0.0144), luTP (0.0008), PTS1 (6.968)
*Pc*AAE5	AT5G16340, AAE 6	63	553	mTP (0.01), cTP (0.1781), luTP (0.0198), PTS1 (6.992)
*Pc*AAE6	AT5G16370, AAE 5	65	551	mTP (0), cTP (0), luTP (0), PTS1 (5.274)
*Pc*AAE7	AT1G20560, butyrate-CoA ligase[a]	70	555	mTP (0.004), cTP (0.006), luTP (0.0135), PTS1 (11.738)
*Pc*AAE8	AT1G21540, AAE 9	65	609	mTP (0.0007), cTP (0.2914), luTP (0.0014), PTS1 (−35.162)

[a,b]The *in vitro* enzymatic activities have been reported (Shockey et al. 2003, Xu et al. 2013). mTP, mitochondrial transit peptide; cTP, chloroplast transit peptide; luTP, thylakoid luminal transit peptide; PTS1, peroxisomal targeting signal 1. The websites of the PTS1 Predictor and TargetP 2.0 are http://mendel.imp.ac.at/pts1/ and http://www.cbs.dtu.dk/services/TargetP/, respectively.

which we named pogostone B, was also detected in *P. cablin* plants (**Supplementary Fig. S1**). Pogostone B showed a very similar distribution pattern to that of pogostone in the tested tissues, albeit present at much lower levels (**Fig. 2B**).

Identification of candidate genes responsible for 4MVCoA biosynthesis

To identify the genes encoding the enzymes involved in the conversion of 4-methylvaleric acid to 4MVCoA, RNAseq libraries were constructed from six different *P. cablin* tissue samples including Seedling, 5W-Root, 5W-Stem, 5W-TLeaf, 8W-TLeaf and 8W-SLeaf from *P. cablin* plants (**Fig. 2A**). The corresponding transcriptome assemblies were then constructed and unigenes were annotated by comparison with genes on the UniProt database and using the BLASTX program. AAE genes of *P. cablin* were identified by using Arabidopsis representative AAE sequences of each of the seven clades of the AAE superfamily, which had been previously defined (Shockey and Browse 2011), to query our transcriptome assembly using TBLASTN software. In addition, we searched the functional annotations of the unigene with keywords such as 'AMP-binding protein'. Since we sought to search for the gene responsible for 4MVCoA biosynthesis from 4-methylvaleric acid, the unigenes encoding putative AAEs catalyzing the conjugation of plant hormones to amino acids were not included in this search. A total of 31 unigenes encoding putative AAEs, containing the conserved AMP-binding motif (PROSITE PS00455) of the AAE family, were identified and further analyzed (**Table 1**, **Supplementary Tables S1 and S2**).

Previous phylogenetic analysis has defined seven subfamilies of AMP-dependent synthetases and ligases in Arabidopsis (Shockey and Browse 2011). Phylogenetic analysis indicated that all the 31 *P. cablin* AAE unigenes group into the previously established six clades of the AAE superfamily (**Supplementary Fig. S2**). Several members of clade VI from *Arabidopsis thaliana* and *Humulus lupulus* have been previously functionally analyzed by *in vitro* biochemical assays. *Hl*CCL2 and *Hl*CCL4 from *H. lupulus* utilize various short-chain fatty acids and branched short-chain fatty acids as substrates (Xu et al. 2013). Both of

these CCLs prefer branched short-chain fatty acids as substrates, with isovaleric acid and isobutyric acid as the preferred substrates for *Hl*CCL2 and *Hl*CCL4, respectively (Xu et al. 2013). The Arabidopsis gene At2G17650 encodes an enzyme that also prefers branched short-chain fatty acids as the substrates, with isovaleric acid serving as the most preferred substrate in *in vitro* enzymatic assays (Xu et al. 2013). Recently, identified *Hc*AAE1 from *Hypericum calycinum* exhibits promiscuous substrate preference, with benzoic acid and several straight/branched short-chain fatty acids such as butyric acid and isobutyric acid as preferred substrates (Singh et al. 2020). These data suggest that there might also be one or more members among the *P. cablin* clade VI genes that encodes an enzyme with a preference for branched-chain fatty acids, including 4-methylvaleric acid. Therefore, we concentrated on the eight genes in the *P. cablin* clade VI, which we designated as *Pc*AAE1∼8, for further examination (**Table 1** and **Supplementary Table S1**).

Tissue-specific/developmental expression patterns of *Pc*AAE clade VI genes

Analysis of expression patterns of the eight genes in *P. cablin* RNAseq database indicated that *Pc*AAE2 expression shows the strongest positive correlation with the distribution pattern of pogostone in these tissues of all these eight genes (**Supplementary Table S1**; **Fig. 2B**). In particular, *Pc*AAE2 transcript reads are much lower in older tissues (8W-TLeaf and 8W-SLeaf) than in seedling and younger plants (5W-Stem and 5W-TLeaf). To more accurately test whether the transcript profiles of *Pc*AAE1∼8 in the clade VI of AAE superfamily are correlated with the distribution pattern of pogostone in different tissues of *P. cablin*, their expression patterns were further confirmed by quantitative reverse transcription polymerase chain reaction (RT-qPCR) in the same tissue samples used for distribution analysis of pogostone (**Fig. 3**). The expression patterns of *Pc*AAE1∼8 obtained by RT-qPCR analysis are mostly consistent with that in RNAseq database (**Supplementary Table S1**; **Fig. 3**). Again, the expression patterns of *Pc*AAE2 in the tested samples show a strong positive correlation with the distribution pattern of pogostone: (i) *Pc*AAE2 is highly expressed in the five
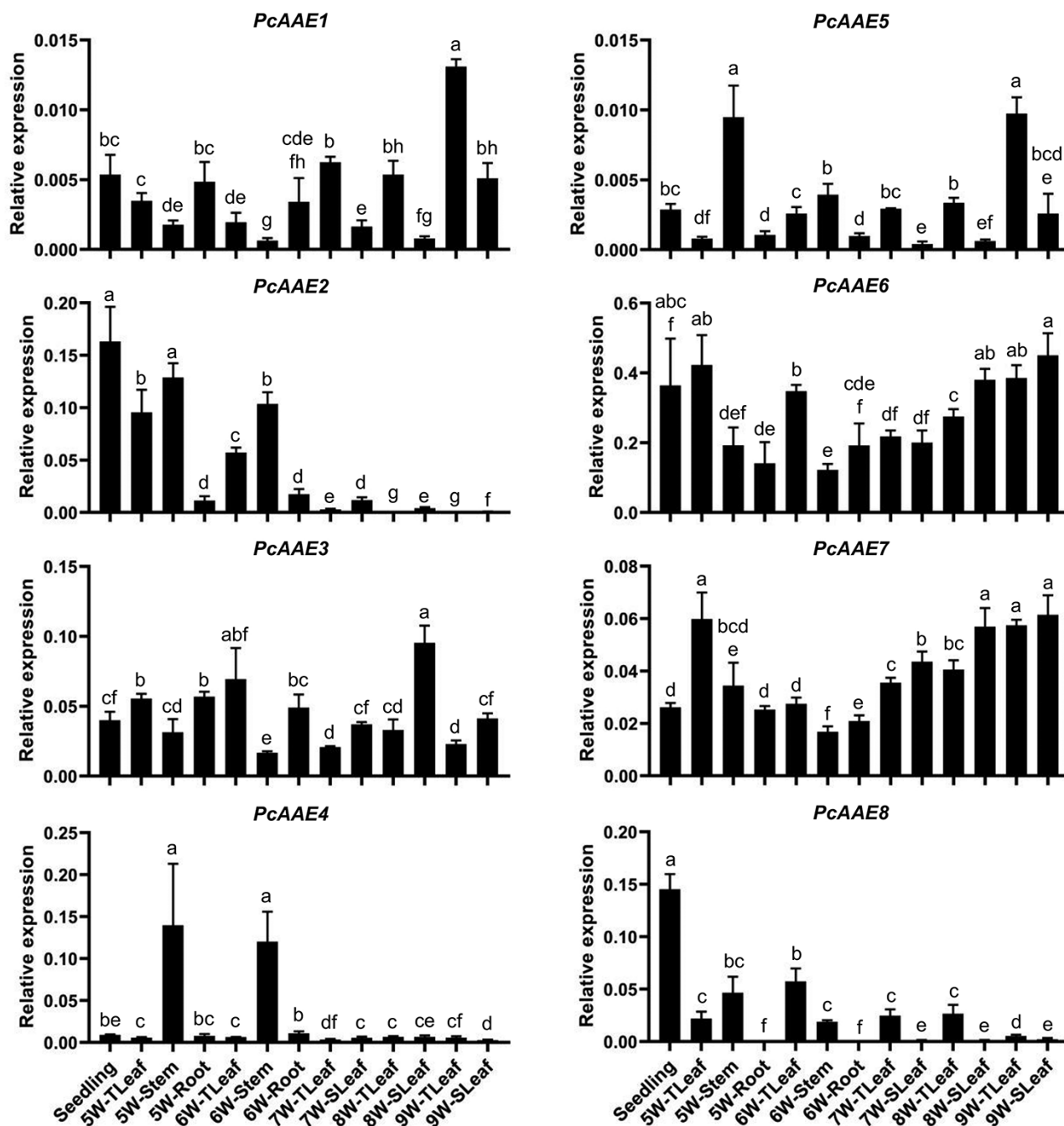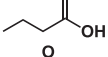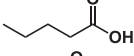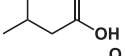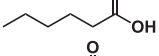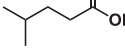
Fig. 3 RT-qPCR analysis of transcript levels of *PcAAE1∼8* in different tissues of *P. cablin* at different stages of development. The RT-qPCR data were expressed as relative expressions of *PcAAE1-8* relative to *P. cablin GAPDH* (glyceraldehyde-3-phosphate dehydrogenase), *Actin7* and *Tubulin3* in each tissue sample. Data are means ± SD of four independent biological replicates. Significant difference among the expression in different tissues of *P. cablin* at different stages of development was tested using independent *t*-tests $P < 0.05$ [Two-tailed distribution; Two-sample unequal variance (heteroscedastic)]. Different letters above the bars mark the statistically significant groups after *t*-tests.

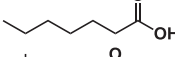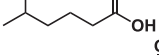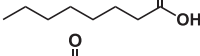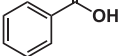tissue samples of seedling, 5W-TLeaf, 5W-Stem, 6W-TLeaf and 6W-Stem of *P. cablin* plants, which show large levels of pogostone; (ii) the expression levels of *PcAAE2* in top leaves from the main stems of 7, 8 and 9-week-old *P. cablin* plants, including the tissue samples of 7W-TLeaf, 8W-TLeaf and 9W-TLeaf, dramatically decrease, and these tissues show negligible levels of pogostone compared to those observed in tissue samples of earlier developmental stages.

PcAAE2 encode a 4MVCoA synthetase

Sequence alignment showed that of the eight *P. cablin* clade VI proteins, *Pc*AAE1 and *Pc*AAE2 share the highest amino acid sequence identities with *Hl*CCL2 (66.0% and 66.2%, respectively) and *Hc*AAE1 (74.6% and 68.9%, respectively) (**Supplementary Fig. S3**) and, thus, are the closest homologs to *Hl*CCL2 and *Hc*AAE1, suggesting that they may have the similar

Table 2 Substrate Specificity of Recombinant *Pc*AAE1, *Pc*AAE2 and *Pc*AAE3

Substrate	Structure	*Pc*AAE1	*Pc*AAE2	*Pc*AAE3
Acetic acid		N.D.	5 ± 1.2	47 ± 0.5
Propionic acid		5 ± 0.2	6 ± 0.6	$100^c \pm 2.2$
Butyric acid		69 ± 1.3	85 ± 5.4	74 ± 1.8
Isobutyric acid		N.D.	6 ± 0.7	42 ± 1.0
Valeric acid		36 ± 0.6	89 ± 2.4	12 ± 0.1
Isovaleric acid		$100^a \pm 4.9$	51 ± 2.4	N.D.
Hexanoic acid		11 ± 0.4	52 ± 3.5	N.D.
4-Methylvaleric acid		21 ± 0.6	$100^b \pm 8.6$	N.D.
Heptanoic acid		N.D.	54 ± 3.6	N.D.
5-Methylhexanoic acid		N.D.	51 ± 3.1	N.D.
Octanoic acid		9 ± 1.3	10 ± 0.4	N.D.
Benzoic acid		N.D.	13 ± 0.3	N.D.

These enzymatic assays were performed using the coupled enzymes assay (Koo et al. 2006, Chen et al. 2011). These activities were measured with substrate and CoA at concentration of 0.5 mM, respectively. Data are expressed as relative mean percentages of three independent biological replicates. N.D. represents 'not detected' or relative activity < 5%.
[a] 100% relative activity of *Pc*AAE1 corresponds to 4.02 μmol min^{-1} mg^{-1} on isovaleric acid.
[b] 100% relative activity of *Pc*AAE2 corresponds to 1.96 μmol min^{-1} mg^{-1} on 4-methylvaleric acid.
[c] 100% relative activity of *Pc*AAE3 corresponds to 6.75 μmol min^{-1} mg^{-1} on propionic acid.

substrate preference toward branched-chain fatty acids. The combination of the positive correlation of expression patterns of *PcAAE2* with the distribution pattern of pogostone and the predicted substrate preference of the protein encoded by *PcAAE2* toward branched-chain fatty acids made it the most likely candidate gene encoding the enzyme responsible for the formation of 4MVCoA precursor of pogostone compared with the other seven genes encoding *P. cablin* clade VI proteins.

To test this hypothesis, we performed *in vitro* biochemical assays with all eight AAE proteins in the *P. cablin* clade VI of the family (**Table 1**). All proteins were produced in *Escherichia coli* with a fused N-terminal His6 tag and purified as soluble proteins. All the eight purified recombinant proteins were initially tested in the coupled enzyme assay with a variety of fatty acid substrates including straight short/medium-chain fatty acids and branched short/medium-chain fatty acids and benzoic acid at a 0.5 mM concentration of the acid substrate and 0.5 mM of CoA. Except for the three recombinant proteins *Pc*AAE1~3, the other five recombinant proteins showed no detected enzymatic activities with any of these selected substrates.

Recombinant *Pc*AAE1 utilizes isovaleric acid as the best substrate among the selected substrates and shows lower activities toward butyric acid, valeric acid and 4-methylvaleric acid (**Table 2**). Recombinant *Pc*AAE2 shows high activities toward 4-methylvaleric acid, valeric acid and butyric acid and utilizes 4-methylvaleric acid as the most preferred substrate (**Table 2**). Recombinant *Pc*AAE2 also accepts 5-methylhexanoic acid, heptanoic acid, hexanoic acid and isovaleric acid as substrates but at a lower rate (**Table 2**). Purified *Pc*AAE3 accepts very short-chain fatty acids as substrates, including acetic acid, propanoic acid and butyric acid, with propionic acid as the preferred substrate (**Table 2**). Therefore, *Pc*AAE1 and *Pc*AAE2 are assigned as the probable isovaleryl-CoA synthetase and 4MVCoA synthetase, respectively, and may supply CoA ester precursors for pogostone and its analogue pogostone B biosynthesis. Thus, in-depth studies of the dependency of their enzymatic activities on incubation conditions and components were further performed. Highest activities of *Pc*AAE1 and *Pc*AAE2 with their respective preferred substrates were observed at pH 7.5 and 30°C. The activities of *Pc*AAE1 and *Pc*AAE2 were dependent on divalent cations with Mg^{2+} being preferred and increased by 1.7-fold and 0.8-fold, respectively, when the concentration of

Table 3 Kinetic Properties of Recombinant *Pc*AAE1 and *Pc*AAE2

Enzyme	Substrate	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m (M^{-1} s^{-1})
*Pc*AAE1	Butyric acid	161.0 ± 5.0	4.0 ± 0.2	24,741 ± 1276
	Isovaleric acid[a]	12.9 ± 0.8	4.1 ± 0.1	332,937 ± 7532
*Pc*AAE2	Butyric acid	570.1 ± 86.5	2.4 ± 0.1	4344 ± 90
	Valeric acid	53.9 ± 2.7	2.0 ± 0.02	37,288 ± 436
	Isovaleric acid[b]	216.9 ± 13.1	1.7 ± 0.08	7759 ± 371
	Hexanoic acid	41.5 ± 2.4	1.0 ± 0.03	22,904 ± 842
	4-Methylvaleric acid[c]	43.4 ± 5.1	2.3 ± 0.1	51,854 ± 2765
	Heptanoic acid	84.5 ± 7.1	1.1 ± 0.04	13,184 ± 433
	5-Methylhexanoic acid	175.5 ± 20.0	1.2 ± 0.05	6951 ± 312

These enzymatic assays were performed using the same coupled enzymes assay as **Table 2**. Data are presented as mean ± SD of three independent biological replicates.
[a]The K_m values of cosubstrates CoA and ATP were 28.2 and 89 μM, respectively.
[b]The K_m values of cosubstrates CoA and ATP were 150.8 and 540.2 μM, respectively.
[c]The K_m values of cosubstrates CoA and ATP were 40.2 and 160.5 μM, respectively.

Mg^{2+} rose from 1 to 10 mM. Additional supplementation with dithiothreitol or a univalent cation at 2.5 mM including K$^+$ does not obviously increase their activities. These two enzymes were not stable at 4, −20 and −80°C and lost more activities at 4, −20°C than at −80°C for 24 h (**Supplementary Fig. S4**).

Since we aim to identify the AAEs responsible for the formation of 4MVCoA and isovaleryl-CoA, the predicted CoA ester precursors of pogostone and its analogue pogostone B, respectively, the enzymatic products of *Pc*AAE1 with isovaleric acid as substrate and *Pc*AAE2 with 4-methylvaleric acid and isovaleric acid as substrates were further verified with LC-QTOF-MS system (**Fig. 4**) and the kinetic properties of *Pc*AAE1 with isovaleric acid as substrate and *Pc*AAE2 with 4-methylvaleric acid and isovaleric acid as substrates were further determined (**Table 3**). The kinetic analysis revealed that *Pc*AAE1 has a K_m value of 12.9 ± 0.8 μM and a catalytic efficiency of 3.33×10^5 s^{-1}M^{-1} for isovaleric acid, and *Pc*AAE2 has a K_m value of 43.4 ± 5.1 μM and catalytic efficiency of 5.19×10^4 s^{-1}M^{-1} for 4-methylvaleric acid, while the K_m value and catalytic efficiency for isovaleric acid are 216.9 ± 13.1 μM and 0.78×10^4 s^{-1}M^{-1}, respectively (**Table 3**).

Considering the previously reported substrate promiscuity of AAEs such as *Cs*AAE1, *Ph*4CL1 and *Hc*AAE1 (Stout et al. 2012, Klempien et al. 2012, Singh et al. 2020), we further determined the kinetic properties of *Pc*AAE1 and *Pc*AAE2 with substrates toward which *Pc*AAE1 and *Pc*AAE2 show relative activities higher than 40% in **Table 2**. The kinetic analysis further revealed that *Pc*AAE2 also has low K_m values with the substrates valeric acid, hexanoic acid and heptanoic acid besides 4-methylvaleric acid (**Table 3**). Although the K_m value of *Pc*AAE2 with hexanoic acid is about the same as that of *Pc*AAE2 with 4-methylvaleric acid, the catalytic efficiency of *Pc*AAE2 with hexanoic acid is only 44% of that of *Pc*AAE2 with 4-methylvaleric acid (**Table 3**).

Subcellular localization of *Pc*AAE1 and *Pc*AAE2

Subcellular localization prediction software TargetP 2.0 and PTS1 predictor were utilized for signal peptide prediction of *Pc*AAE1~8 (**Table 1**). *Pc*AAE1 is predicted to harbor a mitochondrial transit peptide with a high score, while no signal peptide was found for *Pc*AAE2, suggesting that *Pc*AAE1 may target

mitochondria while *Pc*AAE2 is probably localized in cytoplasm. Since 4-methylvaleric acid and isovaleric acid are the most preferred substrates of *Pc*AAE2 and *Pc*AAE1, and the corresponding acyl-CoA ester products are precursors of pogostone and pogostone B, respectively, the subcellular localizations of these two proteins were further analyzed experimentally. This was done by transiently expressing fluorescent protein fusions of the first 30 amino acids (predicted mitochondrial transit peptide) of *Pc*AAE1, whole *Pc*AAE1 and *Pc*AAE2, respectively, in Arabidopsis leaf protoplasts. When the first 30 amino acids of *Pc*AAE1 and whole *Pc*AAE2 were expressed as C-terminal green fluorescent protein (GFP) fusions (SP$_{PcAAE1}$-GFP and *Pc*AAE2-GFP, respectively), complete co-localization of SP$_{PcAAE1}$-GFP signal with MitoTracker Red, a marker dye specific for mitochondria, was observed, indicating that *Pc*AAE1 is a mitochondria-targeted protein (**Fig. 5A**). The GFP signal of *Pc*AAE2-GFP was detected exclusively in the cytoplasm and no GFP signal could be found to overlay with the red chloroplast autofluorescence (**Fig. 5B**). When *Pc*AAE1 and *Pc*AAE2 were expressed as N-terminal GFP fusions (GFP-*Pc*AAE1 and GFP-*Pc*AAE2, respectively), partial co-localization of GFP-*Pc*AAE1 signal with MitoTracker Red was observed (**Fig. 5A**), suggesting that N-terminal GFP partially blocks the mitochondria signal of *Pc*AAE1. The GFP signal of GFP-*Pc*AAE2 was the same as that observed for The GFP signal of *Pc*AAE2-GFP, indicating that *Pc*AAE2 is a cytosol protein (**Fig. 5B**).

Phylogenic relationships of *Pc*AAE2 homologs

To gain further insight into the evolution of *Pc*AAE2, we expanded the phylogenetic analysis of AAE clade VI to include 77 homologs from chlorophytes including *Chlamydomonas reinhardtii* and *Volvox carteri*, Embryophyte, Tracheophyte and angiosperms. The phylogenetic tree was constructed using a maximum likelihood algorithm (**Fig. 6**). The proteins in the expanded clade VI were mainly grouped into four distinct subclades (**Fig. 6**). We designated the group that contains *Pc*AAE1, *Pc*AAE2 and *Pc*AAE7 as subclade VIa, the group that contained *Pc*AAE3 as subclade VIb, the group that included *Pc*AAE5, *Pc*AAE6 and *Pc*AAE8 as subclade VIc, and the group that included *Pc*AAE4 as subclade VId (**Fig. 6**). Six homologs of the AAE clade VI
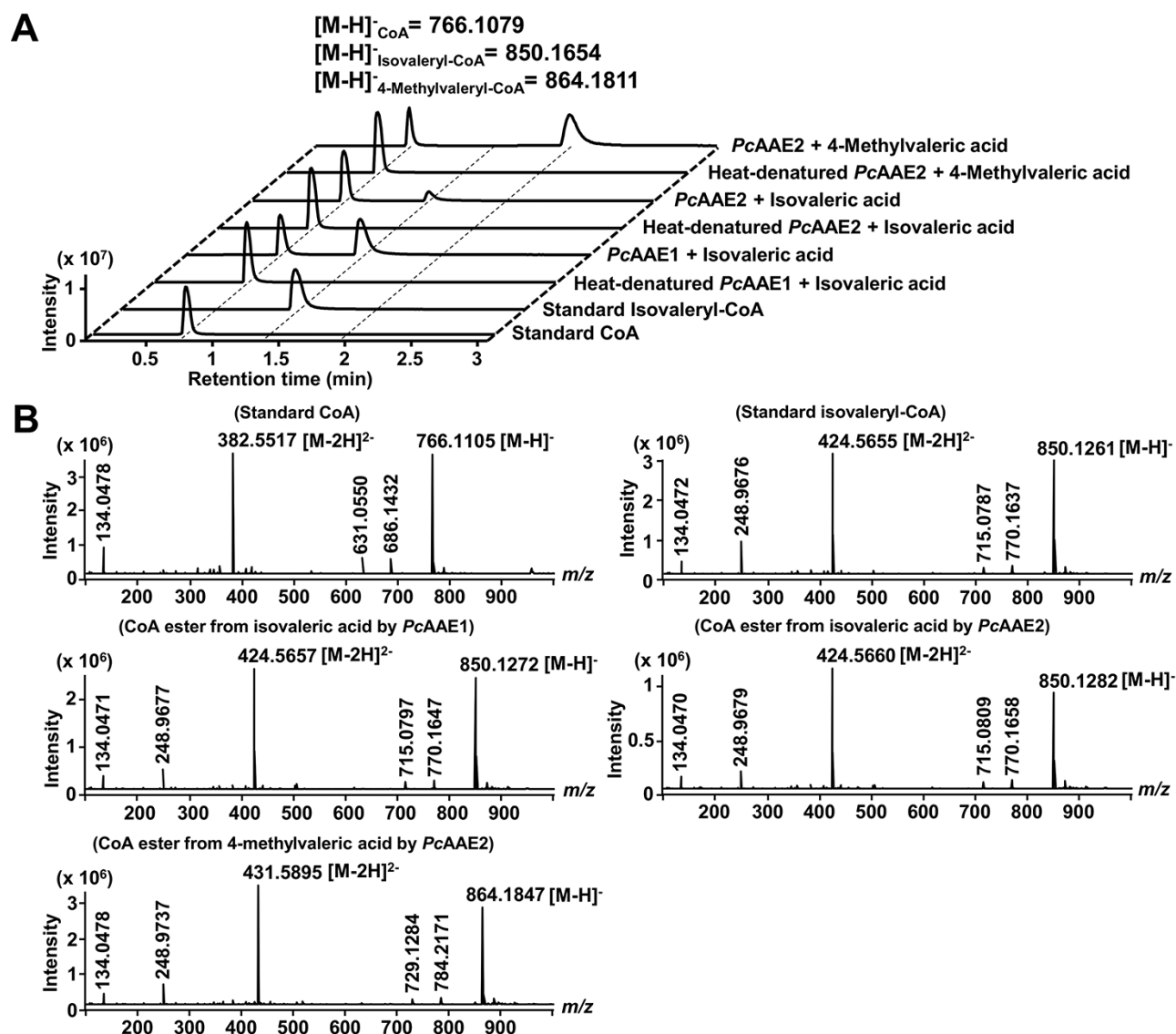
Fig. 4 *In vitro* analyses of *Pc*AAE1 and *Pc*AAE2 activities. A, LC-QTOF-MS analysis of products formed in reactions containing purified *Pc*AAE1 or *Pc*AAE2, free CoA, ATP, MgCl$_2$, isovaleric acid or 4-methylvaleric acid after 30 min of incubation. The extracted ion chromatograms in negative mode of *m/z* 766.1079, 850.1654 and 864.1811 are shown for free CoA, isovaleryl-CoA and 4MVCoA, respectively. B, The mass spectra in negative mode of standard CoA, standard isovaleryl-CoA and CoA esters (predicted isovaleryl-CoA and 4MVCoA) generated *in vitro* by the action of *Pc*AAE1 or *Pc*AAE2 with the substrates of isovaleric acid or 4-methylvaleric acid. The singly [M-H]⁻ and doubly [M-2H]²⁻ charged pseudo-molecular ion are shown horizontally. The CoA ester from isovaleric acid by the action of *Pc*AAE1 or *Pc*AAE2 was identified as isovaleryl-CoA through the comparison of its retention time and mass spectra with that of standard isovaleryl-CoA, while the CoA ester from 4-methylvaleric acid by the action of *Pc*AAE2 was identified as 4MVCoA based on the exact mass and fragment pattern.

from *Chlamydomonas reinhardtii*, *Volvox carteri*, *Physcomitrella patens*, *Selaginella moellendorffii* grouped out of the designated four subclades, suggesting an earlier evolutionary split.

It is interesting that most of clade VI AAEs possess canonical PTS1 signals and are predicted to be targeted to the peroxisome by PTS1 predictor (**Fig. 6**). We also further investigated subcellular localization prediction of other selected genes using Targetp 2.0. 3 out of the 77 clade VI AAEs contain mitochondrial transit peptides at the N-terminus and are predicted to be located in mitochondria, while 15 out of these 77 AAEs contain no obvious signal peptides and are predicted to be located in cytoplasm (**Fig. 6**). In addition, all the six outgroups contain prototypical PTS1 signals with high prediction score, suggesting likely peroxisomal localization. *Pc*AAE1 and *Pc*AAE2 are grouped closely with *Hc*AAE1, *Hl*CCL2 and At2G17650 in the subclade via (**Fig. 6**), which are consistent with the fact that they share very similar substrate preference toward branched/straight short-chain fatty acids.
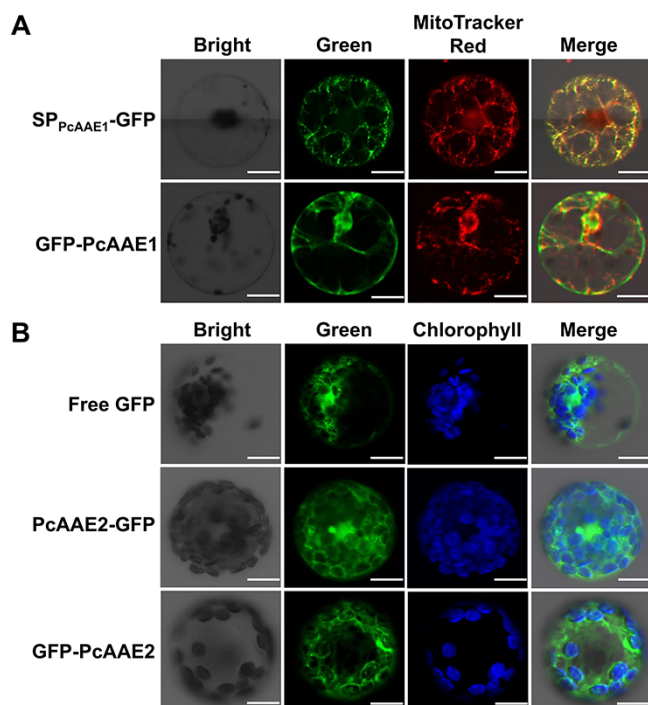
Fig. 5 Subcellular localization of *Pc*AAE1 and *Pc*AAE2 in Arabidopsis leaf mesophyll protoplasts. GFP fusion proteins were visualized by laser confocal microcopy. A, Subcellular localization of *Pc*AAE1 in Arabidopsis leaf mesophyll protoplasts. Both the first 30 amino acids (predicted mitochondrial transit peptide) of *Pc*AAE1 fused C-terminal GFP and whole *Pc*AAE1 fused N-terminal GFP were used for assay of subcellular localization of *Pc*AAE1. Mitochondria were revealed by MitoTracker ® Red-staining. B, Subcellular localization of *Pc*AAE2 in Arabidopsis leaf mesophyll protoplasts. Both *Pc*AAE2 fused N-terminal and C-terminal GFPwere used for assay of subcellular localization of *Pc*AAE2. The signal of free GFP was used as control and chloroplasts were revealed by chlorophyll autofluorescence in pseudo-color blue. Scale bar = 10 μm.

DISCUSSION

*Pc*AAE2 catalyzes the formation of 4MVCoA precursor of pogostone in *P. cablin*

The medicinal plant *P. cablin* is an important commercial source of patchouli oil (Yang et al. 2016, van Beek and Joulain 2018). Patchouli oil has the ability to blend with other essential oils and give a strong base, lasting character, fixative properties and aids in preventing evaporation, thus promoting tenacity; hence, it is widely used in the manufacturing of perfume, cosmetics, soaps, scents body lotions and detergents (Gubareva 2004, Kiso et al. 2004, Swamy and Sinniah 2015, Hu et al. 2017). Previous studies have indicated that pogostone, one of the major chemical constituent of patchouli oil, is largely responsible for its intense aromatic odor, and more recently, this compound has been proved to exert many pharmaceutical activities (Li et al. 2012, 2014, Yi et al. 2013, Su et al. 2015, 2017, Chen et al. 2016, Wang et al. 2016, Cao et al. 2017). At present, pogostone is mainly isolated from patchouli oil, which is usually produced by

steam distillation of the shade dried leaves and stems (Swamy and Sinniah 2015).

In patchouli oil, besides pogostone, a previous study also detected its different analogues including pogostone B in much smaller concentrations (Nakahara et al. 1975, Rijke et al. 1978, van Beek and Joulain 2018), which is similar to that was observed in the present study. Pogostone B, albeit at a much lower level, showed very similar accumulation pattern to that of pogostone in tested *P. cablin* tissues (Fig. 2B). The distribution of both pogostone and pogostone B in *P. cablin* top leaves newly generated from main stems shows a developmental pattern (Fig. 2). To begin elucidating the biosynthetic pathway of pogostone, we searched for the enzyme that catalyzes the formation of 4MVCoA, one of the precursors of pogostone. Transcriptome profiling and searches of RNAseq databases obtained from three different stages of developing *P. cablin* plants identified 31 AAE unigenes (Supplementary Table S1), with one unigene, named *PcAAE2*, showing the strongest positive correlation with the distribution pattern of pogostone (Supplementary Table S1; Figs. 2B and 3).

The purified recombinant *Pc*AAE2 protein is able to effectively catalyze the formation of 4MVCoA *in vitro* from 4-methylvaleric acid, with a K_m value of 43.4 μM and a catalytic efficiency of 5.19×10^4 s^{-1}M^{-1} for 4-methylvaleric acid (Table 3; Fig. 4). In addition to 4-methylvaleric acid, *Pc*AAE2 is also able to effectively catalyze the formation of CoA esters *in vitro* from valeric acid and hexanoic acid, with K_m values of 53.9 and 41.5 μM, respectively (Table 3), thus indicating substrate promiscuity, which has been commonly observed for AAEs (Stout et al. 2012, Klempien et al. 2012, Xu et al. 2013, Singh et al. 2020). Despite the observation of substrate promiscuity, *Pc*AAE2 does show the highest catalytic efficiency toward 4-methylvaleric acid among the tested fatty acid substrates (Table 3), suggesting it as the 4MVCoA synthetase *in vivo*. In addition, *Pc*AAE2 is also able to catalyze the formation of isovaleryl-CoA *in vitro* from isovaleric acid with a K_m value of 216.9 μM and a catalytic efficiency of 0.78×10^4 s^{-1}M^{-1}, which accounts for about 15% of that with 4-methylvaleric acid (Table 3). Therefore, given the much lower level of pogostone B compared with that of pogostone in tested *P. cablin* tissues (Fig. 2), *Pc*AAE2 may also supply isovaleryl-CoA precursor for pogostone B biosynthesis in cytoplasm, where the polyketide pogostone and pogostone B are believed to be synthesized.

Besides *Pc*AAE2, *Pc*AAE1, another member of clade VI AAEs from *P. cablin*, also shows substrate preference toward branched short-chain fatty acid (Table 2). The purified recombinant *Pc*AAE1 protein accepts isovaleric acid as the best-preferred substrate and is able to effectively catalyze the formation of isovaleryl-CoA *in vitro* from isovaleric acid, with a K_m value of 12.9 μM and a catalytic efficiency of 3.33×10^5 s^{-1}M^{-1} for isovaleric acid (Table 3; Fig. 4). Although the expression pattern of *Pc*AAE1 did not show an obvious positive correlation with the accumulation pattern of pogostone B in the tested *P. cablin* tissues (Supplementary Table S1; Figs. 2 and 3), the
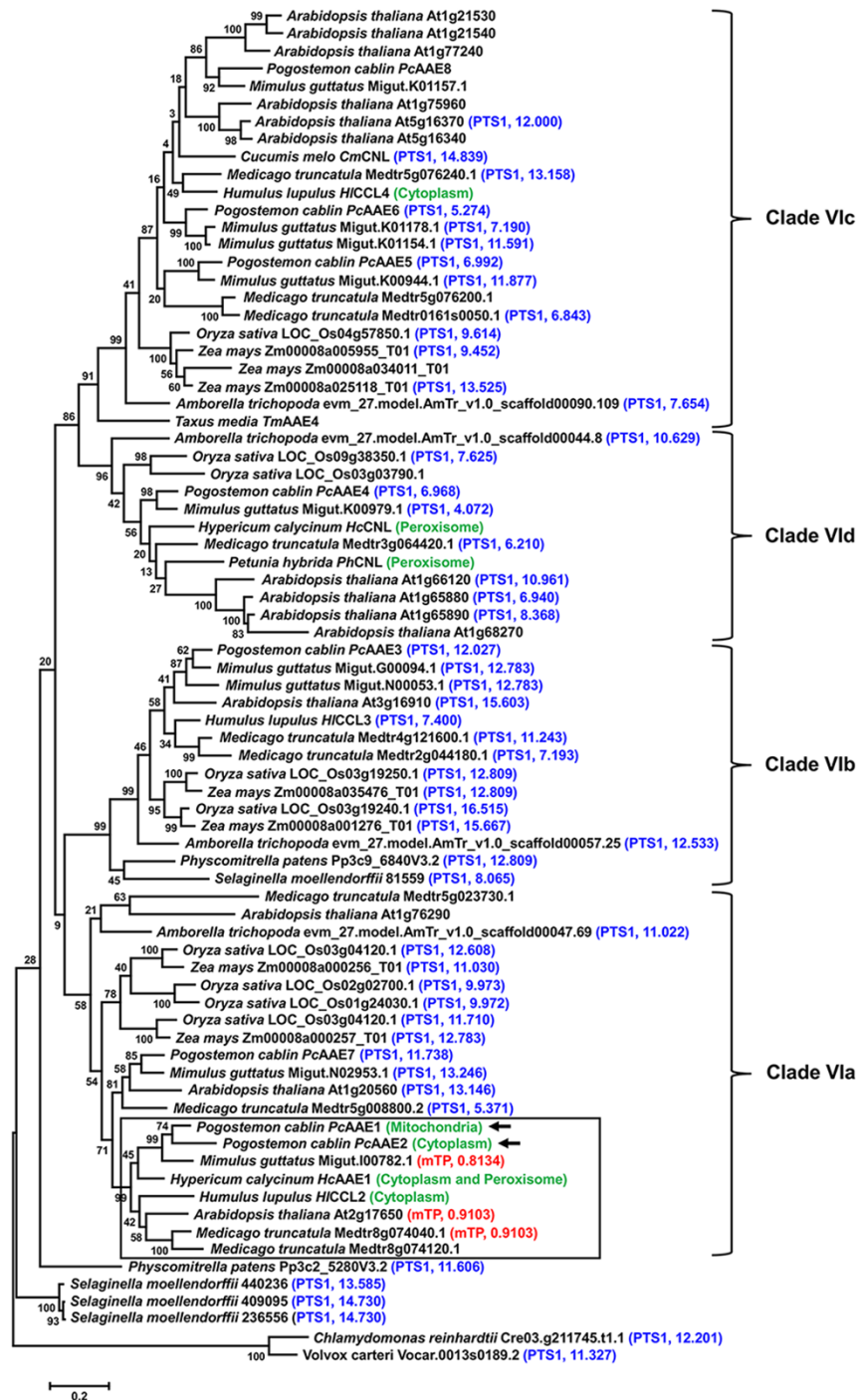
Fig. 6 Phylogenetic analyses of AAE clade VI homologs and *Pc*AAE2 homologs from the plant kingdom. All proteins used for phylogenetic tree construction except *Pc*AAE1 ∼ 8 were obtained from the PHYTOZONE v12.1 site (https://phytozome.jgi.doe.gov/pz/portal.html), NCBI website (https://www.ncbi.nlm.nih.gov/) and TAIR website (https://www.arabidopsis.org/). The protein sequences of *Pc*AAE1 ∼ 8 were directly obtained from *P. cablin* transcriptome database. Sequences were aligned using ClustalW and tree were generated by the maximum likelihood algorithm using MEGA7.0 (Kumar et al. 2016). Branch point bootstrap values were calculated with 1000 replicates. The subcellular localization of each protein was analyzed using signal prediction algorithm including Targetp 2.0, SignalP 5.0 and The PTS1 predictor (http://www.cbs.dtu.dk/services/TargetP/, http://www.cbs.dtu.dk/services/SignalP/ and http://mendel.imp.ac.at/pts1/). The predicted signal peptide and prediction value were listed along with protein name and species. Peroxisome targeting sequence 1 (PTS1) was marked in blue color, while mitochondrial transit peptide (mTP) was marked in red color. None of chloroplast transit peptide was detected in all selected AAEs. The verified subcellular localizations of AAEs were labeled in green color (**Supplementary Table S3**). Black arrows indicate proteins of which the enzymatic activities and subcellular localizations were characterized in this paper. The subclade that *Pc*AAE1, *Pc*AAE2, *Hc*AAE1, *Hl*CCL2 and At2g17650 were grouped into was indicated in black frame.

possibility of involvement of *PcAAE1* in pogostone B biosynthesis cannot be completely ruled out. Subcellular localization assays indicated that *PcAAE1* is a mitochondria-targeted protein (**Fig. 5A**). Therefore, it is mostly likely that *PcAAE1* maintains the isovaleryl-CoA pool in mitochondria under certain conditions and plays a role in a dynamic balance between leucine degradation and the electron transport chain of mitochondria (Araujo et al. 2010).

Cytosolic *PcAAE2* may have evolved from an ancestral peroxisomal AAE

The AAE gene family expanded and diversified as land plants evolved from their smaller, simpler algal ancestors (Shockey and Browse 2011). Each of the AAE gene family of *Chlamydomonas reinhardtii* and *Volvox carteri*, two unicellular green algae, contains one gene (Cr03g211745 for *Chlamydomonas reinhardtii* and Vc0013s0189 for *Volvox carteri*) only distally related to *PcAAE2*, and this gene was grouped out of the AAE clade VI subfamily in our phylogenetic tree (**Fig. 6**), suggesting that this clade VI is specific to multicellular land plants as previously reported (Shockey and Browse 2011).

Interestingly, most of clade VI AAEs in this phylogenetic analysis (46 out of 71) possess a predicted peroxisomal targeting signal 1 (PTS1) and are therefore predicted to be targeted to the peroxisomes (**Fig. 6**), while three clade VI AAEs possess predicted mitochondrial transit peptides and are predicted to be located in the mitochondria (**Fig. 6**). Fifteen clade VI AAEs possess no obvious predicted signal peptides and are predicted to be targeted to cytoplasm (**Fig. 6**). Each of *PcAAE2*'s closest homologs (Cr03g211745, Vc0013s0189, Pp3c25280, Pp3c96840, Sm409095, Sm440236, Sm236556 and Sm81559) from *Chlamydomonas reinhardtii* (a unicellular green algae), *Volvox carteri* (unicellular green algae), *Physcomitrella patens* (a moss) and *Selaginella moellendorffii* (an ancient vascular plant) contains a prototypical PTS1 sequence at the C-terminus with high prediction score (**Fig. 6**) and is therefore highly likely to be located in peroxisome where the catabolic process of β-oxidation occurs. These ancestor peroxisomal *PcAAE2* homologs may play important roles in β-oxidation. *PcAAE2* lacks a PTS1 and locates in cytoplasm where the polyketide pogostone is believed to be synthesized (**Fig. 5B**).

Since the possession of a PTS1 sequence is the ancestral trait in the AAE VI clade throughout the plant kingdom, and a PTS1 sequence is also present in the *P. cablin* VIa clade (in *PcAAE7*, a protein that has 65% amino acid identity to *PcAAE2*), the recruitment of *PcAAE2* for pogostone biosynthesis appears to have arisen through a duplication event of an ancestral peroxisomal protein and subsequent divergence in both subcellular localization and substrate specificity. Similar cases for a cytoplasm localized hexanoyl-CoA synthetase (*CsAAE1*) recruited to cannabinoid biosynthesis in *Cannabis sativa* trichomes and for two cytoplasm localized branched short-chain CoA ligases (*HlCCL2* and *HlCCL4*) recruited to bitter acid biosynthesis in *H. lupulus* trichomes and for recently

identified cytoplasm and peroxisome dually localized benzoyl-CoA ligase (*HcAAE1*) recruited to xanthone biosynthesis in *H. calycinum* have been already reported (Stout et al. 2012, Xu et al. 2013, Singh et al. 2020). A similar case may also have occurred for *PcAAE1*, which encodes a mitochondrial isovaleryl-CoA synthetase (**Table 3**; **Fig. 5A**), probably playing a role in a dynamic balance between leucine degradation and the electron transport chain of mitochondria through maintaining the pool of isovaleryl-CoA in mitochondria (Araujo et al. 2010). This conclusion could be confirmed by further experimental subcellular localization analysis of more clade VI genes from different plant species.

Assessment of *P. cablin PcAAE1* and *PcAAE2* sequence properties

AAEs belonging to the acyl-CoA synthetase subfamily are members of the so-called ANL superfamily, in which some recognized conserved signature motifs (A1-A10), common to all adenylation domains have been identified (Marahiel et al. 1997, Gulick 2009) and are also present in *P. cablin PcAAE1* and *PcAAE2* (**Supplementary Fig. S3**). AAEs exhibit similar overall folding patterns with large N-terminal and small C-terminal domains. Biochemical and structural studies demonstrate that AAE enzymes undergo large-scale domain movement to facilitate the two-step catalytic reaction, an adenylate-forming conformation during the first half-reaction and a thioester-forming conformation during the second step (Bar-Tana and Rose 1968, Gulick et al. 2003, Reger et al. 2008, Gulick 2009, Li and Nair 2015). The hinge residue that facilitates this movement is usually aspartate in the A8 motif (Gulick et al. 2003), which is conserved in *P. cablin PcAAE1* and *PcAAE2* (**Supplementary Fig. S3**). As in all adenylate forming enzyme, two conserved AMP-binding motifs (AMP1 and AMP2) (Weimar et al. 2002, Morgan-Kiss and Cronan 2004) and the catalytically relevant amino acid lysine (Reger et al. 2007, Hu et al. 2010, Li and Nair 2015) are also present in *P. cablin PcAAE1* and *PcAAE2* (**Supplementary Fig. S3**).

In addition, *PcAAE1*∼3 from *P. cablin* and some other medium/short-chain acyl-CoA synthetases from *A. thaliana*, *C. sativa*, *H. lupulus*, *H. calycinum*, *P. cablin*, *Taxus media* have been shown to be active with a range of medium/short-chain fatty acids (Stout et al. 2012, Xu et al. 2013, Srividya et al. 2020, Singh et al. 2020). Phylogenetic analysis indicated that they fall into two clades (**Fig. 7A**). Clade I, which contains acyl-CoA synthetases preferring branched/straight short-chain fatty acids, is clearly distinct from clade II, which contains *CsAAE3*, *TmAAE13*, *TmAAE15* and *TmAAE16*, all of which show high activities toward straight medium-chain fatty acids. Previous analysis of bacterial AAEs indicated an fatty acyl-CoA synthetase (FACS) signature motif appearing to contribute to the fatty acid-binding pocket and thus affects the preference for fatty acids of different chain lengths (Black et al. 1997, 2000). Sequence divergence across their FACS signature motifs indicated that they fall into two groups that are consistent with
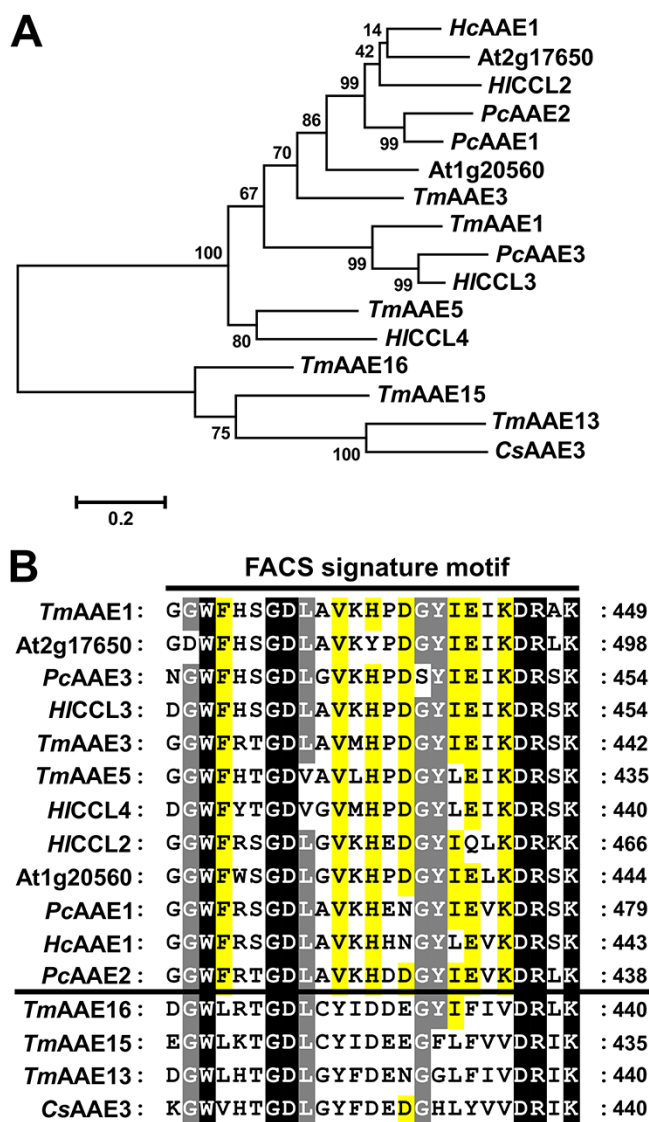
Fig. 7 Analysis of the characterized medium/short chain acyl- CoA synthetase from *A. thaliana*, *C. sativa*, *H. lupulus*, *H. calycinum*, *P. cablin*, *T. media*. A, Maximum likelihood inference of the phylogenetic relationship of the characterized medium/short chain acyl-CoA synthetase from *C. sativa*, *H. lupulus*, *H. calycinum*, *P. cablin*, *T. media* (**Supplementary Table S3**). Sequences were aligned using ClustalW and tree were generated by the Maximum likelihood algorithm using MEGA7.0 (Kumar et al. 2016). Branch point bootstrap values were calculated with 1000 replicates. B, Alignment of the amino acid residues of the FACS signature motifs of the acyl-CoA synthetases. Alignment was conducted by ClustalW (https://embnet.vital-it.ch/software/ClustalW.html) and edited using GeneDoc software (https://genedoc.software.informer.com/2.7/).

the result of phylogenetic analysis (**Fig. 7B**), suggesting that this motif is an important determining factor for their functional differentiation and evolutionary split. Phe[4], Val[11] and Lys[21] in the 25 amino acids of the FACS signature motif are specifically conserved in the first group, while Tyr[11], Asp[13] and Val[21] are specifically conserved in the latter group (**Fig. 7B**), suggesting

that these amino acids may play a vital role in fatty acid chain length specificity. To verify this hypothesis, *in silico* homolog model based on the available crystal structures of AAEs such as *Pt*4CL1 and *Nt*4CL2 (Hu et al. 2010, Li and Nair 2015) should be required. Further extensive comparisons on the structures and gain-of-function mutations in this motif are needed for a deeper understanding of the substrate preferences for branch/straight chain and chain length of these enzymes.

Metabolic engineering of pogostone biosynthesis

P. cablin (Lamiaceae) is a commercially important plant for its essential oil (patchouli oil). It is native to southeast Asia and is now cultivated widely in many tropical and subtropical regions (Wu et al. 2010, Xie et al. 2017), including China, Indonesia, the Philippines and Thailand for the extraction of essential oils from its leaves and stems, and especially as a natural source of patchouli alcohol and pogostone (Singh and Rao 2009, He et al. 2016, Swamy and Sinniah 2016). Different *P. cablin* cultivars exhibit significant differences in quality and bioactive components, as these factors are influenced by climate, soil nutrients and water in the different locations (Swamy et al. 2015). Based on the main components of essential oils, *P. cablin* in China can be divided into two chemotypes: pogostone type, with a high content of pogostone and low content of patchouli alcohol, and patchoulol type, which has a high content of patchouli alcohol but a low content of pogostone (Liu et al. 2002, Luo et al. 2003, Huang et al. 2016). Traditionally, the patchoulol-type cultivars are mainly used in the perfume industry, whereas the pogostone-type cultivars are considered as medicinal plants in China. However, the production of pogostone, likely the main effective compound in medicinal use, is extremely low due to the limitation of cultivation area (Hu et al. 2006, Wu et al. 2010). Due to the discovery of more and more pharmacological activities of pogostone, the pogostone-type cultivars are highly popular (Chen et al. 2021). However, their natural supplies cannot sufficiently satisfy the economic demand (Zhang et al. 2019).

Our discovery of *Pc*AAE2, the enzyme that catalyzes the biosynthesis of 4MVCoA, will serve as a step forward in the construction of high-yield pogostone *P. cablin* varieties. Increasing *Pc*AAE2 transcript levels or changing its developmental expression pattern could lead to further increase in the production of pogostone. Pogostone has received increased attention along with the discovery of its additional potential bioactivities and pharmacological properties such as antibacterial, anti-inflammatory, anticancer activities and others (Chen et al. 2015, 2016, Sun et al. 2016, Wang et al. 2016, Cao et al. 2017). The discovery of 4-MVCoA synthetase could also facilitate the functional characterization of other enzymes such as polyketide synthases involved in the pogostone biosynthesis and, thus, will allow for the reconstruction of this pathway in the heterologous system using a synthetic biology approach to produce pogostone in large amounts.

Materials and Methods

Plant materials and chemicals

P. cablin seeds were germinated in soil, and plants were grown in a greenhouse with a 16 h light/8 h darkness photoperiod at 30°C. The various tissues samples including seedling, roots, main stems and leaves were collected from *P. cablin* plants. For RT-qPCR analysis, the harvested tissues were flash frozen in liquid nitrogen and stored at −80°C until use. *Arabidopsis* plants were grown in soil in a growth chamber with the photoperiod (12 h light/12 h dark at 23°C) under 60% humidity. The well-expanded leaves from 4-week-old *Arabidopsis* plant were selected for protoplast isolation for subcellular localization.

Pogostone and pogostone B standard were synthesized as described previously (Yu et al. 2018). Since standard 4MVCoA is not available commercially, the CoA esters produced by *Pc*AAE2 from 4-methylvaleric acid were identified as 4MVCoA based on the exact mass and fragment pattern by LC-QTOF-MS. Mito-Tracker Red used for mitochondria staining was purchased from ThermoFisher, and other chemicals were purchased from Sigma-Aldrich.

GC-MS analysis of pogostone and pogostone B from *P. cablin* seedling, roots, main stems and leaves from the main stem at different stages of development

P. cablin seedlings, roots, main stems and leaf tissues at different stages of development were collected and cut into small pieces, 100 mg of which were transferred into a tube containing 500 μL MTBE with 0.05 ng/μl tetradecane as internal standard. The tube was vortexed for 3 min at maximum speed, incubated at room temperature for 1 h and then centrifuged for 3 min at the speed of 10,000 g/min. The MTBE phase was collected and analyzed by GC-MS. For GC-MS analysis, 1 μl aliquot of the sample was injected into Agilent GC-MSD (Agilent 7890B-5977B) system equipped with the DB-5MS column (30 m × 0.25 mm × 0.25 μm film thickness, Agilent, USA). The oven temperature was programmed as follows: initial temperature, 50°C for 3 min, 50 to 320°C at a rate of 10°C min^{-1} and hold for 5 min. The identification of compounds was assigned by comparison of their retention times and mass fragmentations with corresponding authentic standards. The measurement of pogostone and pogostone analogue was performed by comparison to the corresponding authentic standard curves.

RNAseq analysis

Total RNA was extracted from seedling, root, main stem and leaf parts of *P. cablin* plants at different developmental stages using Total RNA Isolation Kit from Omega. The RNAseq library construction and transcriptome assembling were performed as described previously (Xu et al. 2018). The expression levels of all transcripts were estimated using the Trinity software, which estimated transcript abundance using the RSEM function (Li and Dewey 2011).

Identifying AAE genes in the *P. cablin* transcriptome

To find candidate unigenes encoding AAEs in our *P. cablin* transcriptome database, we constructed local Nucleotide Database using the software (BioEdit Sequence Alignment Editor, https://bioedit.software.informer.com/7.2/) and queried it with Arabidopsis representative AAEs from each of seven clades of AAE superfamily using the TBLASTN function. We also screened the entire annotated database for the functional annotations 'AMP binding protein' and 'CoA ligase' and added transcripts identified in this way to our list of candidates.

Quantitative reverse transcription polymerase chain reaction analysis of *PcAAE1*∼8 transcript Levels

For quantitative real-time PCR analysis of transcripts in different tissues, total RNA was isolated using total RNA Isolation Kit from Omega with a DNA digestion step. RNA concentration and quality (260/280 and 260/230 ratios) were determined using the Thermo Scientific™ NanoDrop™ Lite Spectrophotometer. Reverse transcription polymerase chain reaction (RT-PCR) was prepared using High Capacity cDNA Reverse Transcription Kit from Thermo Fisher Scientific following the manufacturer's instructions with minor modification. 3 μg total RNA was used for making cDNAs in a 20 μl RT-PCR reaction system containing 5 μM oligo(dT), which was obtained from Thermo Fisher Scientific (cat number; #SO132). RT-qPCR analysis was performed using Power SYBR Green PCR master mix (Thermo Fisher) using Bio-Rad CFX Connect Real-Time PCR Detection System. Each reaction consisted of 10 μl Power SYBR Green PCR master mix, 2 μl cDNA and 1 μl of 5 μM forward and reverse primers in a total volume of 20 μl made up with sterile water. The thermal profile included, after a 2 min step at 95°C, 40 cycles of denaturation at 95°C for 5 s and annealing at 55°C for 30 s and amplification for 30 s. Melting curves were recorded by melting the amplified amplicon at 65–95°C in an increment increase of 0.5°C per cycle. *P. cablin* glyceraldehyde 3-phosphate dehydrogenase (*PcGAPDH*), *PcActin7*, and *PcTubulin3* were used as internal reference genes. For *PcAAE1*∼8 and the three reference genes, single melting curve peaks were obtained. Primer specificities were evaluated by melt curve analysis and agarose gel electrophoresis. cDNAs of four biological replicates were analyzed. No-template (NTC) and no-reverse transcriptase controls (NRC) were used to rule out primer dimer formation and genomic DNA contamination, respectively.

The RT-qPCR data were expressed as relative expressions of *PcAAE1*∼8 relative to the three reference genes in each tissue. The relative expressions of *PcAAE1*∼8 in each tissue were calculated using 2^dCq with the dCq calculated by subtracting Cq values of *PcAAE* genes from the geometric mean Cq values of the three reference genes in every tissue and the amplification efficiency (*E*) set as 2. The equation used in this process is defined by just one group of Cq values containing Cq for the gene of interest and geometric mean Cq for the three reference genes in tested tissues (Vandesompele et al. 2002). Stable expression of the three reference genes among all selected samples was confirmed before integrating their Cq values in normalization. Primers used in this study are designed using Primer 5.0 software and listed in Supplementary Table S4. Assay of unbiased RNA coverage of oligo (dT) and random primer for *PcAAE1*∼8 in the top leaf and main stem tissues of 6-week-old *P. cablin* plants was performed and the relevant results were supplied as **Supplementary Fig. S5**.

Isolation of His-tagged recombinant proteins of *Pc*AAE1∼8

The open reading frames of *PcAAE2*∼8 were obtained by RT-PCR from prepared cDNA of *P. cablin* seedlings and were introduced into the expression vector pET-28a(+) between the restriction sites of NdeI and BamHI using NEBuilder® HiFi DNA Assembly Cloning Kit (NEB, Catalog number: E5520S), in each case generating a fusion gene that encoded a 'tag' of His6 residues at the N-terminus for expression in *Escherichia coli*. To obtain soluble proteins for expression in *E. coli*, a truncated open reading frame of *PcAAE1*, missing the first 30 codons, was obtained by RT-PCR from prepared cDNA of *P. cablin* seedlings and was introduced into the expression vector pET-28a(+) to generate a fusion gene that encoded a 'tag' of His6 residues at the N-terminus for expression in *E. coli*. Recombinant proteins of *Pc*AAE1∼8 were expressed in *E. coli* strain BL21(BE3) and purified using Ni-NTA agarose chromatography (Qiagen) as previously described (Xu et al. 2018).

Enzymatic assays of recombinant *Pc*AAE1∼8

Acyl-CoA synthetase activities of recombinant *Pc*AAE1∼8 proteins were measured using the coupled enzymes assay (Koo et al. 2006, Chen et al. 2011). For substrate specificity assays, reaction mixtures (250 μl) contain 0.1 M Tris-HCl, pH 7.5, 2 mM DTT, 5 mM ATP, 10 mM MgCl$_2$, 0.5 mM CoA, 0.4 mM NADH, 0.5 mM fatty acid substrate, 1 mM phosphoenolpyruvate and four units each of myokinase, pyruvate kinase, and lactate dehydrogenase. The reaction was initiated by adding 1 μg of purified enzyme. Oxidation of NADH was monitored at 30°C for 15 mins at 340 nm ($\varepsilon_{340\,nm} = 6220\,M^{-1}\,cm^{-1}$) with a Beckman DU530 spectrophotometer. The enzyme concentration (1 μg) and incubation time (15 min) were selected to attain a linear reaction velocity during the assay period. To determine the kinetic parameters of recombinant *Pc*AAE2, a similar

protocol was followed. The K_m value for fatty acid substrate was determined by fixing the concentration of CoA at 1 mM and ATP at 5 mM, whereas the K_m value for CoA was determined by fixing the concentration of fatty acid substrate at 1 mM and ATP at 5 mM, and the K_m value for ATP was determined by fixing the concentration of fatty acid substrate at 1 mM and CoA at 1 mM. K_m and k_{cat} values were calculated from initial rate data by using the hyperbolic regression analysis method in Hyper32 software (version 1.0.0, http://hyper32.software.informer.com/).

In addition, Acyl-CoA synthetase activity of recombinant *Pc*AAE1 and *Pc*AAE2 proteins was further measured using an LC-QTOF-MS-based method. Briefly, the reaction mixtures (250 μl) contain 0.1 M Tris-HCl, pH 7.5, 2 mM DTT, 5 mM ATP, 10 mM $MgCl_2$, 0.5 mM CoA, 0.5 mM fatty acid substrate. The reaction was initiated by adding 10 μg of purified *Pc*AAE1 or *Pc*AAE2 protein. Assays were stopped after 30 min by the addition of 250 μl methanol. The solution was filtered through a 0.45 μm filter prior to analysis by LC-QTOF-MS system (Agilent, 6545 LC/QTOF-MS) coupled with a C_{18} column (ZORBAX RRHD Plus C_{18}, Φ2.1 × 50 mm, 1.8 μm) at 35°C. The gradient (solvent A, water + 0.1% formic acid; solvent B, methanol) program was set as follows at a flow of 0.5 ml min^{-1}: 0–0.5 min, a linear gradient from 5% to 40% of B; 0.5–2.5 min, 40% of B; 2.5–3.5 min, a linear gradient to 100% of B; 3.5–4.5 min, 100% of B; 4.5–4.6 min, a linear gradient to 100% of A and equilibration for 2 min before starting the next sample. The operating parameters were set as follows: capillary voltage, 3,500 V; nebulizer pressure, 35 pounds per square inch gauge; drying gas flow rate, 8 L min^{-1}; gas temperature, 320°C; sheath gas temperature, 350°C; sheath gas flow, 11 L min^{-1} and the fragmentor, 175 V. Mass spectrum were acquired in negative mode, and lock mass correction was performed using standard hexakis (2,2,3,3-tetrafluoropropoxy) phosphazine and purine. The compounds were analyzed with the parameters as follows: 766.1079 (*m/z*) for CoA, 850.1654 (*m/z*) for isovaleryl-CoA and 864.1811(*m/z*) for 4MVCoA.

Subcellular Localization of *Pc*AAE1 and *Pc*AAE2

The vector used for subcellular localization of analysis was pTF486. For the construction of these vectors expressing fusion proteins with GFP fused with C-terminus of the first 30 amino acids of *Pc*AAE1 and whole *Pc*AAE2, the open reading frames of the first 30 amino acids of *Pc*AAE1 and whole *Pc*AAE2 were amplified from *P. cablin* seedling cDNA, respectively, and inserted into pTF486 vector between restriction sites of SalI and BamHI using In-Fusion® HD Cloning Kit (TaKaRa, Catalog number: 638910) following the protocol. For the construction of these vectors expressing fusion proteins with GFP fused with N-terminus of *Pc*AAE1 and *Pc*AAE2, respectively, the open reading frames of *PcAAE1* and *PcAAE2* and gene encoding GFP were amplified from *P. cablin* seedling cDNA and pTF486 plasmid, respectively, the purified PCR product of either *PcAAE1*or *PcAAE2* and gene encoding GFP and purified double enzyme digestion product of pTF486 plasmid between restriction sites of SalI and BamHI were assembled together using In-Fusion® HD Cloning Kit (TaKaRa) following the protocol. Protoplasts were prepared from Arabidopsis leaves, and transformation and confocal microscopy were performed as described previously (Yoo et al. 2007). Primers used in the construction of these vectors are listed in **Supplementary Table S4,** and the overlaps (15–30 bp) in the primer sequences that were used for DNA fragment assembly were underlined.

Supplementary Data

Supplementary data are available at PCP online.

Data Availability

Data supporting the finding of this study are available within the article and its supplementary files. The sequences of the genes used in this study are available under GenBank accessions as follows: *Pc*AAE1, MW413950; *Pc*AAE2, MW413951; *Pc*AAE3, MW413952; *Pc*AAE4, MW413953; *Pc*AAE5, MW413954; *Pc*AAE6, MW413955; *Pc*AAE7, MW413956; *Pc*AAE8, MW413957; *Pc*GAPDH, MW735976; *Pc*Actin7, MW735977; *Pc*Tubulin3, MW735978. The bioproject accession number for the RNAseq data is PRJNA713906.

Author Contributions

H.X. designed the experiments; J.C., L.L., Y.W. and Z.L. conducted the experiments, analyzed data or provided material; J.C., G.W., E.P. and H.X wrote the article; all authors edited the article.

Disclosures

The authors have no conflicts of interest to declare.

References

Araujo, W.L., Ishizaki, K., Nunes-Nesi, A., Larson, T.R., Tohge, T., Krahnert, I., et al. (2010) Identification of the 2-Hydroxyglutarate and Isovaleryl-CoA dehydrogenases as alternative electron donors linking lysine catabolism to the electron transport chain of arabidopsis mitochondria. *Plant Cell* 22: 1549–1563.

Bar-Tana, J. and Rose, G. (1968) Studies on medium-chain fatty acyl-coenzyme a synthetase. Enzyme fraction I: mechanism of reaction and allosteric properties. *Biochem. J.* 109: 275–282.

Black, P.N., DiRusso, C.C., Sherin, D., MacColl, R., Knudsen, J. and Weimar, J.D. (2000) Affinity labeling fatty acyl-CoA synthetase with 9-*p*-azidophenoxy nonanoic acid and the identification of the fatty acid-binding site. *J. Biol. Chem.* 275: 38547–38553.

Black, P.N., Zhang, Q., Weimar, J.D. and DiRusso, C.C. (1997) Mutational analysis of a fatty acyl-coenzyme A synthetase signature motif identifies seven amino acid residues that modulate fatty acid substrate specificity. *J. Biol. Chem.* 272: 4896–4903.

Cao, Z.X., Yang, Y.T., Yu, S., Li, Y.Z., Wang, W.W., Huang, J., et al. (2017) Pogostone induces autophagy and apoptosis involving PI3K/Akt/mTOR axis in human colorectal carcinoma HCT116 cells. *J. Ethnopharmacol.* 202: 20–27.

Chen, H., Kim, H.U., Weng, H. and Browse, J. (2011) Malonyl-CoA synthetase, encoded by ACYL ACTIVATING ENZYME13, is essential for growth and development of arabidopsis. *Plant Cell* 23: 2247–2262.

Chen, H.M., Liao, H.J., Liu, Y.H., Zheng, Y.F., Wu, X.L., Su, Z.Q., et al. (2015) Protective effects of pogostone from *Pogostemonis Herba* against ethanol-induced gastric ulcer in rats. *Fitoterapia* 100: 110–117.

Chen, J.R., Xie, X.F., Li, M.T., Xiong, Q.Y., Li, G.M., Zhang, H.Q., et al. (2021) Pharmacological activities and mechanisms of action of *Pogostemon cablin* Benth: a review. *Chin Med.* 16: 5.

Chen, X.Y., Chen, H.M., Liu, Y.H., Zhang, Z.B., Zheng, Y.F., Su, Z.Q., et al. (2016) The gastroprotective effect of pogostone from *Pogostemonis Herba* against indomethacin-induced gastric ulcer in rats. *Exp. Biol. Med.* 241: 193–204.

Gubareva, L.V. (2004) Molecular mechanisms of influenza virus resistance to neuraminidase inhibitors. *Virus Res.* 103: 199–203.

Gulick, A.M. (2009) Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* 4: 811–827.

Gulick, A.M., Starai, V.J., Horswill, A.R., Homick, K.M. and Escalante-Semerena, J.C. (2003) The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5′-propylphosphate and coenzyme A. *Biochemistry* 42: 2866–2873.

He, J.J., Chen, H.M., Li, C.W., Wu, D.W., Wu, X.L., Shi, S.J., et al. (2013) Experimental study on antinociceptive and anti-allergy effects of patchouli oil. *J. Essent. Oil Res.* 25: 488–496.

He, Y., Deng, C., Xiong, L., Qin, S.S. and Peng, C. (2016) Transcriptome sequencing provides insights into the metabolic pathways of patchouli alcohol and pogostone in *Pogostemon cablin* (Blanco) Benth. *Genes Genom.* 38: 1031–1039.

Hu, G., Peng, C., Xie, X., Zhang, S. and Cao, X. (2017) Availability, pharmaceutics, security, pharmacokinetics, and pharmacological activities of patchouli alcohol. *Evid. Based Complement. Alternat. Med.* 2017: 4850612.

Hu, L.F., Li, S.P., Cao, H., Liu, J.J., Gao, J.L., Yang, F.Q., et al. (2006) GC-MS fingerprint of *Pogostemon cablin* in China. *J. Pharm. Biomed.* 42: 200–206.

Hu, Y., Gai, Y., Yin, L., Wang, X., Feng, C., Feng, L., et al. (2010) Crystal structures of a *Populus tomentosa* 4-coumarate: CoAligase shed light on its enzymatic mechanisms. *Plant Cell* 22: 3093–3104.

Huang, H.R., Wu, W., Zhang, J.X., Wang, L.J., Yuan, Y.M. and Ge, X.J. (2016) A genetic delineation of Patchouli (*Pogostemon cablin*) revealed by specific-locus amplified fragment sequencing. *J. Syst. Evol.* 54: 491–501.

Huang, S.H., Xian, J.D., Kong, S.Z., Li, Y.C., Xie, J.H., Lin, J., et al. (2014) Insecticidal activity of pogostone against *Spodoptera litura* and *Spodoptera exigua* (Lepidoptera: Noctuidae). *Pest Manag. Sci.* 70: 510–516.

Kiso, M., Mitamura, K., Sakai-Tagawa, Y., Shiraishi, K., Kawakami, C., Kimura, K., et al. (2004) Resistant influenza A viruses in children treated with oseltamivir: descriptive study. *Lancet* 364: 759–765.

Klempien, A., Kaminaga, Y., Qualley, A., Nagegowda, D.A., Widhalm, J.R., Orlova, I., et al. (2012) Contribution of CoA ligases to benzenoid biosynthesis in petunia flowers. *Plant Cell* 24: 2015–2030.

Koo, A.J.K., Chung, H.S., Kobayashi, Y. and Howe, G.A. (2006) Identification of a peroxisomal acyl-activating enzyme involved in the biosynthesis of jasmonic acid in Arabidopsis. *J. Biol. Chem.* 281: 33511–33520.

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

Li, B. and Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 12: 323.

Li, C.W., Wu, X.L., Zhao, X.N., Su, Z.Q., Chen, H.M., Wang, X.F., et al. (2013) Anti-inflammatory property of the ethanol extract of the root and rhizome of *Pogostemon cablin* (Blanco) Benth. *Sci. World J.* 2013: 434151.

Li, W., Wei, G., Pan, C.M., Liu, X.X., Huang, S. and Xiu, H.H. (2004) Investigation on the influential factors of the volatile oil and main constituent content in *Pogostemon cablin*. *Zhongguo Zhong Yao Za Zhi = Zhongguo Zhongyao Zazhi = China Journal of Chinese Materia Medica* 29: 28–31.

Li, Y.C., Liang, H.C., Chen, H.M., Tan, L.R., Yi, Y.Y., Qin, Z., et al. (2012) Anti-Candida albicans activity and pharmacokinetics of pogostone isolated from *Pogostemonis Herba*. *Phytomedicine* 20: 77–83.

Li, Y.C., Xian, Y.F., Su, Z.R., Ip, S.P., Xie, J.H., Liao, J.B., et al. (2014) Pogostone suppresses proinflammatory mediator production and protects against endotoxic shock in mice. *J. Ethnopharmacol.* 157: 212–221.

Li, Z. and Nair, S.K. (2015) Structural basis for specificity and flexibility in a plant 4-coumarate: CoALigase. *Structure* 23: 2032–2042.

Lin, R.F., Feng, X.X., Li, C.W., Zhang, X.J., Yu, X.T., Zhou, J.Y., et al. (2014) Prevention of UV radiation-induced cutaneous photoaging in mice by topical administration of patchouli oil. *J. Ethnopharmacol.* 154: 408–418.

Liu, Y.P., Luo, J.P., Feng, Y.F., Guo, X.L. and Cao, H. (2002) DNA profiling of *Pogostemon cablin* chemotypes differing in essential oil composition. *Yao Xue Xue Bao = Acta Pharmaceutica Sinica* 37: 304–308.

Lu, T.C., Liao, J.C., Huang, T.H., Lin, Y.C., Liu, C.Y., Chiu, Y.J., et al. (2011) Analgesic and anti-inflammatory activities of the methanol extract from *Pogostemon cablin*. *Evid-Based Compl. Alt.* 2011: 1–9.

Luchesi, L.A., Paulus, D., Busso, C., Frata, M.T. and Oliveira, J.B. (2020) Chemical composition, antifungal and antioxidant activity of essential oils from *Baccharis dracunculifolia* and *Pogostemon cablin* against *Fusarium graminearum*. *Nat. Prod. Res.* 34: 1–4.

Luo, J.P., Liu, Y.P., Feng, Y.F., Guo, X.L. and Cao, H. (2003) Two chemotypes of *Pogostemon cablin* and influence of region of cultivation and harvesting time on volatile oil composition. *Yao Xue Xue Bao = Acta Pharmaceutica Sinica* 38: 307–310.

Marahiel, M.A., Stachelhaus, T. and Mootz, H.D. (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* 97: 2651–2674.

Morgan-Kiss, R.M. and Cronan, J.E. (2004) The *Escherichia coli* fadK (ydiD) gene encodes an anerobically regulated short chain acyl-CoA synthetase. *J. Biol. Chem.* 279: 37324–37333.

Nakahara, S., Kumatani, K. and Kameoka, H. (1975) Acidic compounds in patchouli oil. *Phytochemistry* 14: 2712–2713.

Nobuta, K., Okrent, R.A., Stoutemyer, M., Rodibaugh, N., Kempema, L., Wildermuth, M.C., et al. (2007) The GH3 acyl adenylase family member PBS3 regulates salicylic acid-dependent defense responses in Arabidopsis. *Plant Physiol.* 144: 1144–1156.

Okrent, R.A., Brooks, M.D. and Wildermuth, M.C. (2009) Arabidopsis GH3.12 (PBS3) conjugates amino acids to 4-substituted benzoates and is inhibited by salicylate. *J. Biol. Chem.* 284: 9742–9754.

Pattnaik, S., Subramanyam, V.R. and Kole, C. (1996) Antibacterial and antifungal activity of ten essential oils *in vitro*. *Microbios* 86: 237–246.

Peng, F., Wan, F., Xiong, L., Peng, C., Dai, M. and Chen, J.P. (2014) *In vitro* and *in vivo* antibacterial activity of Pogostone. *Chin. Med. J-Peking* 127: 4001–4005.

Reger, A.S., Carney, J.M. and Gulick, A.M. (2007) Biochemical and crystallographic analysis of substrate binding and conformational changes in acetyl-CoA synthetase. *Biochemistry* 46: 6536–6546.

Reger, A.S., Wu, R., Dunaway-Mariano, D. and Gulick, A.M. (2008) Structural characterization of a 140 degrees domain movement in the two-step reaction catalyzed by 4-Chlorobenzoate : CoA ligase. *Biochemistry* 47: 8016–8025.

Rijke, D.D., Traas, P.C., Heide, R.T., Boelens, H. and Takken, H.J. (1978) Acidic components in essential oils of costus root, patchouli and olibanum. *Phytochemistry* 17: 1664–1666.

Shockey, J. and Browse, J. (2011) Genome-level and biochemical diversity of the acyl-activating enzyme superfamily in plants. *Plant J.* 66: 143–160.

Shockey, J.M., Fulda, M.S. and Browse, J. (2003) Arabidopsis contains a large superfamily of acyl-activating enzymes. Phylogenetic and biochemical analysis reveals a new class of acyl-coenzyme A synthetases. *Plant Physiol.* 132: 1065–1076.

Singh, M. and Rao, R.S.G. (2009) Influence of sources and doses of N and K on herbage, oil yield and nutrient uptake of patchouli [*Pogostemon cablin* (Blanco) Benth.] in semi-arid tropics. *Ind. Crop Prod.* 29: 229–234.

Singh, P., Preu, L., Beuerle, T., Kaufholdt, D., Hansch, R., Beerhues, L., et al. (2020) A promiscuous coenzyme A ligase provides benzoyl-coenzyme A for xanthone biosynthesis in *Hypericum*. *Plant J.* 104: 1472–1490.

Srividya, N., Lange, I., Hartmann, M., Li, Q., Mirzaei, M. and Lange, B.M. (2020) Biochemical characterization of acyl activating enzymes for side chain moieties of Taxol and its analogs. *J. Biol. Chem.* 295: 4963–4973.

Stout, J.M., Boubakir, Z., Ambrose, S.J., Purves, R.W. and Page, J.E. (2012) The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in *Cannabis sativa* trichomes. *Plant J.* 71: 353–365.

Su, J., Li, C., Yu, X., Yang, G., Deng, J., Su, Z., et al. (2017) Protective effect of pogostone on 2,4,6-Trinitrobenzenesulfonic acid-induced experimental colitis via inhibition of T helper cell. *Front Pharmacol.* 8: 829.

Su, J.Y., Luo, X., Zhang, X.J., Deng, X.L., Su, Z.R., Zhou, L., et al. (2015) Immunosuppressive activity of pogostone on T cells: blocking proliferation via S phase arrest. *Int. Immunopharmacol.* 26: 328–337.

Sun, C.Y., Xu, L.Q., Zhang, Z.B., Chen, C.H., Huang, Y.Z., Su, Z.Q., et al. (2016) Protective effects of pogostone against LPS-induced acute lung injury in mice via regulation of Keap1-Nr12/NF-kappa B signaling pathways. *Int. Immunopharmacol.* 32: 55–61.

Swamy, M.K., Mohanty, S.K., Sinniah, U.R. and Maniyam, A. (2015) Evaluation of patchouli (*Pogostemon cablin* Benth.) cultivars for growth, yield and quality parameters. *J. Essent. Oil-Bear. Plants* 18: 826–832.

Swamy, M.K. and Sinniah, U.R. (2015) A comprehensive review on the phytochemical constituents and pharmacological activities of *Pogostemon cablin* Benth.: an aromatic medicinal plant of industrial importance. *Molecules* 20: 8521–8547.

Swamy, M.K. and Sinniah, U.R. (2016) Patchouli (*Pogostemon cablin* Benth.): botany, agrotechnology and biotechnological aspects. *Ind. Crop Prod.* 87: 161–176.

van Beek, T.A. and Joulain, D. (2018) The essential oil of patchouli, *Pogostemon cablin*: A review. *Flavour Frag. J.* 33: 6–51.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3: RESEARCH0034.

Wang, X.F., Huang, Y.F., Wang, L., Xu, L.Q., Yu, X.T., Liu, Y.H., et al. (2016) Photo-protective activity of pogostone against UV-induced skin premature aging in mice. *Exp. Gerontol.* 77: 76–86.

Weimar, J.D., DiRusso, C.C., Delio, R. and Black, P.N. (2002) Functional role of fatty acyl-coenzyme A synthetase in the transmembrane movement and activation of exogenous long-chain fatty acids. Amino acid residues within the ATP/AMP signature motif of *Escherichia coli* FadD are required for enzyme activity and fatty acid transport. *J. Biol. Chem.* 277: 29369–29376.

Wu, Y.G., Guo, Q.S., He, J.C., Lin, Y.F., Luo, L.J. and Liu, G.D. (2010) Genetic diversity analysis among and within populations of *Pogostemon cablin* from China with ISSR and SRAP markers. *Biochem. Syst. Ecol.* 38: 63–72.

Xie, J.H., Chai, T.T., Xu, R., Liu, D., Yang, Y.X., Deng, Z.C., et al. (2017) Induction of defense-related enzymes in patchouli inoculated with virulent *Ralstonia solanacearum*. *Electron. J. Biotechn.* 27: 63–69.

Xu, H., Li, W., Schilmiller, A.L., van Eekelen, H., de Vos, R.C.H., Jongsma, M.A., et al. (2019) Pyrethric acid of natural pyrethrin insecticide: complete pathway elucidation and reconstitution in *Nicotiana benthamiana*. *New Phytol.* 223: 751–765.

Xu, H., Zhang, F., Liu, B., Huhman, D.V., Sumner, L.W., Dixon, R.A., et al. (2013) Characterization of the formation of branched short-chain fatty acid: CoAsfor bitter acid biosynthesis in hop glandular trichomes. *Mol. Plant* 6: 1301–1317.

Xu, H.Y., Moghe, G.D., Wiegert-Rininger, K., Schilmiller, A.L., Barry, C.S., Last, R.L., et al. (2018) Coexpression analysis identifies two oxidoreductases involved in the biosynthesis of the monoterpene acid moiety of natural pyrethrin insecticides in *Tanacetum cinerariifolium*. *Plant Physiol.* 176: 524–537.

Yang, Y.H., Kong, W.J., Feng, H.H., Dou, X.W., Zhao, L.H., Xiao, Q., et al. (2016) Quantitative and fingerprinting analysis of *Pogostemon cablin* based on GC-FID combined with chemometrics. *J. Pharm. Biomed.* 121: 84–90.

Yi, Y.Y., He, J.J., Su, J.Q., Kong, S.Z., Su, J.Y., Li, Y.C., et al. (2013) Synthesis and antimicrobial evaluation of pogostone and its analogues. *Fitoterapia* 84: 135–139.

Yoo, S.D., Cho, Y.H. and Sheen, J. (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2: 1565–1572.

Yu, J., Landberg, J., Shavarebi, F., Bilanchone, V., Okerlund, A., Wanninayake, U., et al. (2018) Bioengineering triacetic acid lactone production in *Yarrowia lipolytica* for pogostone synthesis. *Biotechnol. Bioeng.* 115: 2383–2388.

Zhang, C.Y., Liu, T.J., Yuan, X., Huang, H.R., Yao, G., Mo, X.L., et al. (2019) The plastid genome and its implications in barcoding specific-chemotypes of the medicinal herb *Pogostemon cablin* in China. *PLoS One* 14: e0215512.

Zhao, Z.Z., Lu, J., Leung, K., Chan, C.L. and Jiang, Z.H. (2005) Determination of patchoulic alcohol in herba pogostemonis by GC-MS-MS. *Chem. Pharm. Bull.* 53: 856–860.