PCP
PLANT & CELL PHYSIOLOGY

Evolution and Function of the *Populus* SABATH Family Reveal That a Single Amino Acid Change Results in a Substrate Switch

Xue-Min Han[1], Qi Yang[2], Yan-Jing Liu[3], Zhi-Ling Yang[1], Xiao-Ru Wang[1], Qing-Yin Zeng[2] and Hai-Ling Yang[1,*]

[1]College of Biological Sciences and Biotechnology, Beijing Forestry University, Beijing 100083, China
[2]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[3]State Key Laboratory of Tree Genetics and Breeding, Chinese Academy of Forestry, Beijing 100091, China
*Corresponding author: E-mail, yhailing77@163.com; Fax, +86-10-62336164.
(Received July 13, 2017; Accepted December 8, 2017)

Regular Paper

Evolutionary mechanisms of substrate specificities of enzyme families remain poorly understood. Plant SABATH methyltransferases catalyze methylation of the carboxyl group of various low molecular weight metabolites. Investigation of the functional diversification of the SABATH family in plants could shed light on the evolution of substrate specificities in this enzyme family. Previous studies identified 28 SABATH genes from the *Populus trichocarpa* genome. In this study, we re-annotated the *Populus* SABATH gene family, and performed molecular evolution, gene expression and biochemical analyses of this large gene family. Twenty-eight *Populus* SABATH genes were divided into three classes with distinct divergences in their gene structure, expression responses to abiotic stressors and enzymatic properties of encoded proteins. *Populus* class I SABATH proteins converted IAA to methyl-IAA, class II SABATH proteins converted benzoic acid (BA) and salicylic acid (SA) to methyl-BA and methyl-SA, while class III SABATH proteins converted farnesoic acid (FA) to methyl-FA. For *Populus* class II SABATH proteins, both forward and reverse mutagenesis studies showed that a single amino acid switch between PtSABATH4 and PtSABATH24 resulted in substrate switch. Our findings provide new insights into the evolution of substrate specificities of enzyme families.

Keywords: Enzyme activity • Enzyme family • Expression divergence • Molecular evolution • *Populus trichocarpa* • Substrate switch.

Abbreviations: BA, benzoic acid; CA, coumalic acid; FA, farnesoic acid; JH III, juvenile hormone III; MBP, maltose-binding protein; MeBA, methyl benzoate; MeFA, methyl farnesoate; MeIAA, methyl-IAA; MeJA, methyl jasmonate; MeSA, methyl salicylate; ML, maximum likelihood; MT, methyltransferase; NA, nicotinic acid; SA, salicylic acid; SAM, *S*-adenosyl-l-methionine; VA, vanillic acid.

Introduction

Understanding evolution at the molecular level is a central aim in modern biology because the ultimate source of biological diversity can be found in the functional diversity of macromolecules (Camps et al. 2007). Enzymatic proteins as macromolecules are essential components of every biological system. As biological catalysts, enzymes play a major role in metabolic pathways by lowering the activation energy necessary for chemical reactions to occur. An enzyme family is defined as a group of enzymes with similar amino acid sequences and protein folds encoded by a group of genes that are created by gene duplication (Lan et al. 2013). Many studies have shown that different members of an enzyme family have distinct substrate specificities or catalytic activities (Barkman et al. 2007, Dixon et al. 2009, Falara et al. 2011, Liu et al. 2013). However, the molecular mechanism of this functional divergence within enzyme families remains poorly understood.

Plants can produce many small-molecule metabolites, such as IAA, salicylic acid (SA) and benzoic acid (BA). These small molecules are involved in many essential biological processes, including plant development and environmental responses (Rivas-San Vicente and Plasencia 2011, Sauer et al. 2013). These small-molecule compounds can be modified by *S*-adenosyl-l-methionine (SAM)-dependent methylation. For example, methyltransferases (MTs) can convert IAA, SA and BA into methyl-IAA (MeIAA), methyl salicylate (MeSA), and methyl benzoate (MeBA) (Chen et al. 2003, Zhao et al. 2009). Methylation of these small molecules might regulate the concentrations of their free acids in plant tissues. In addition, the methylated products may have biological functions different from that of their precursors. In plants, SAM-dependent methylation of small-molecule metabolites is catalyzed by SABATH MTs, which are named after the first three functionally characterized members, salicylic acid methyltransferase (SAMT), benzoic acid methyltransferase (BAMT) and theobromine synthase (D'Auria et al. 2003). Plant SABATH MTs are multifunctional proteins encoded by a large gene family. The *Arabidopsis thaliana* and rice (*Oryza sativa*) genomes contain 24 and 41 members, respectively (D'Auria et al. 2003, Zhao et al. 2008). Some SABATH proteins can also catalyze nitrogen methylation, and are involved in caffeine biosynthesis (Ogawa et al. 2001). Recently, a study of moss SABATH proteins found that PpSABATH1 did not show any activity towards carboxylic acids or nitrogen-containing compounds, but had MT activity towards a number of thiols (Zhao et al. 2012). Because SABATH proteins have a wide substrate spectrum, this family is an ideal model to investigate the molecular mechanisms of functional divergence within an enzyme family.

Plant Cell Physiol. 59(2): 392–403 (2018) doi:10.1093/pcp/pcx198, Advance Access publication on 11 December 2017, available online at www.pcp.oxfordjournals.org

To date, genome-wide analyses of the SABATH gene family have mainly focused on herbaceous annual plants, such as Arabidopsis and rice. The expansion and functional divergence of this large gene family in other land plants have not been well investigated. *Populus* is the most important tree model system in plant genomics. The relatively close phylogenetic relationship between *Populus* and Arabidopsis facilitates our investigation of gene family evolution in eudicots (Jansson and Douglas 2007). In addition, genome sequences of many land plants are currently available, including the bryophyte *Physcomitrella patens* (Rensing et al. 2008) and the lycophyte *Selaginella moellendorffii* (Banks et al. 2011). A systematic analysis of SABATH genes in bryophytes, lycophytes, gymnosperms and angiosperms will improve our understanding of the evolutionary history of this large gene family in land plants.

Populus trichocarpa was the first tree species with a sequenced genome (Tuskan et al. 2006), which allowed the investigation of large gene families, such as the SABATH MTs. Previous studies identified 28 SABATH genes (*PtSABATH1*–*PtSABATH28*) in the *P. trichocarpa* genome (Zhao et al. 2013). PtSABATH1, also named PtIAMT1, can catalyze the formation of MeIAA using SAM as a methyl donor and IAA as a methyl acceptor (Zhao et al. 2007). PtSABATH3 showed higher activity towards jasmonic acid (JA) than BA (Zhao et al. 2013). PtSABATH4 displayed the highest level of catalytic activity towards SA and a relatively low level of activity towards BA (Zhao et al. 2013). In this study, we re-annotated the *Populus* SABATH MT gene family, and systematically examined the evolutionary and biochemical characteristics of this family of enzymes. Functional divergence was observed among *Populus* SABATH genes by examining multiple characteristics, such as gene expression patterns and biochemical characteristics. Previous studies have shown that a single amino acid can play a critical role in substrate variation (Barkman et al. 2007, Huang et al. 2012). Thus, we performed site-directed mutagenesis of the *Populus* SABATH proteins in this study, which revealed that a single amino acid change contributed to substrate switch.

Results

Identification of the SABATH gene family in the Populus genome

Previous studies identified 28 SABATH genes in the *P. trichocarpa* genome (Zhao et al. 2013). In this study, re-identification and manual re-annotation were performed to confirm the SABATH genes of *P. trichocarpa* further. We found that *Potri*.017G122*900* contained two complete SABATH genes. The proteins encoded by these two genes showed 88% sequence identity. We also obtained full-length coding sequences of both genes by reverse transcription–PCR (RT–PCR). Thus, in this study, *Potri*.017G122*900* was re-annotated as two distinct SABATH genes (*PtSABATH18* and *19*; Supplementary Table S1). We also found that *Potri*.010G104*700* only contained a partial SABATH fragment, and thus was considered to be a partial pseudogene and was not included in the subsequent phylogenetic and gene expression analyses. In addition, this study considered three SABATH genes (*Potri*.017G121*600*, *Potri*.019G022*200* and *Potri*.019G022*600*) to be full-length pseudogenes based on the presence of frameshifts disrupting the coding region or premature stop codons, resulting in a truncated protein. After removing the stop codons or revising the frameshifts by deletion of one nucleotide, these sequences were included in the phylogenetic and gene expression analyses. Altogether, this study identified 28 full-length SABATH genes in the *Populus* genome. The names, GenBank accession numbers and protein sequences of the 28 full-length SABATH genes are listed in Supplementary Table S1.

Chromosomal distribution of the *Populus* SABATH gene family

We examined the chromosomal distribution of the SABATH gene family in the *Populus* genome. The physical locations of 25 of the 28 full-length SABATH genes were assigned to 10 of the 19 *Populus* chromosomes (**Fig. 1**), while the other three were assigned to two as yet unattributed scaffold fragments (Supplementary Table S1). The SABATH genes were unevenly distributed on the 10 *Populus* chromosomes. Chromosomes 1, 10, 14 and 15 each harbored only one SABATH gene. Relatively high densities of SABATH genes were discovered on chromosomes 17 and 19, each of which harbored six SABATH genes. In this study, if three or more SABATH genes were clustered together, they were considered a SABATH gene cluster. Based on this criterion, two gene clusters containing 11 SABATH genes were observed on chromosomes 17 and 19 (**Fig. 1**). Based on the phylogenetic tree and the positions of these 11 SABATH genes, we reconstructed the expansion history of the two clusters. The most parsimonious scenario for gene duplication and rearrangement was predicted (**Fig. 2**). For cluster I, five SABATH genes were created by four rounds of tandem duplication and one rearrangement event. In cluster II, six SABATH genes were created by three rounds of tandem duplication and one rearrangement event.

Sequence and structural characteristics of the *Populus* SABATH gene family

We reconstructed a maximum likelihood (ML) tree of the 28 *Populus* SABATH genes using two SABATH genes from *Aspergillus clavatus* and *Coccomyxa subellipsoidea* as outgroups. The phylogenetic tree showed that *Populus* SABATH genes were divided into three distinct clades (red, green and blue in **Fig. 3A**). In this study, the SABATH genes in the red, green and blue clades were considered class I, II and III SABATH genes, respectively. Class I, II and III contained one, 10 and 17 members, respectively. Conserved gene structures were found within each class among the 27 class II and III *Populus* SABATH genes (**Fig. 3C**). All class III SABATH genes had a three-exon/two-intron structure. The class I and II SABATH genes contained a four-exon/three-intron structure. This class-specific gene structure supports the subfamily designations among the 28 full-length *Populus* SABATH genes.

Pairwise comparisons of the 28 full-length SABATH protein sequences revealed some notable features (**Fig. 4**). Ten class II

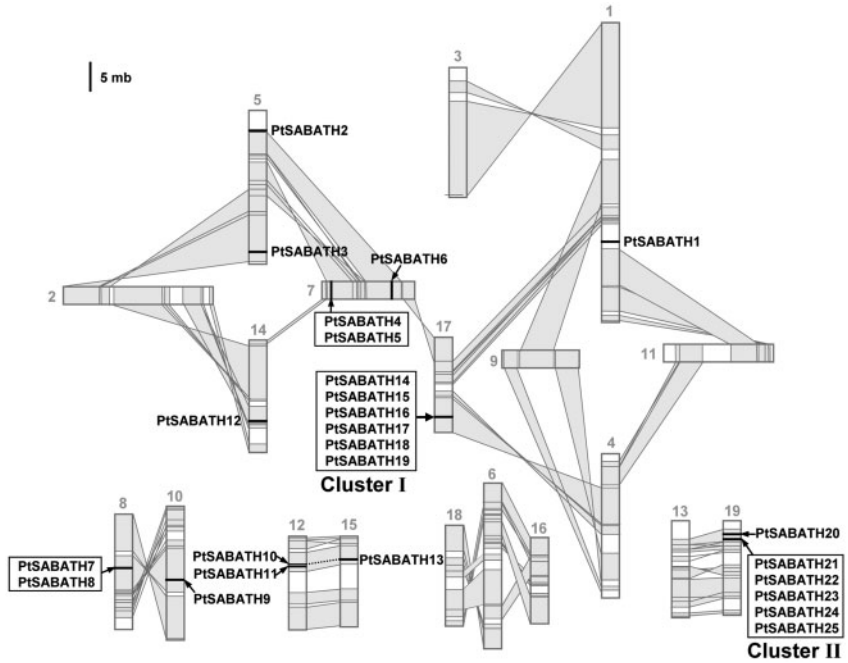


Fig. 1 Genomic localization of *Populus* SABATH genes. Schematic view of chromosome reorganization by the most recent whole-genome duplication in *Populus* (adapted from Tuskan et al. 2006). Regions that are assumed to correspond to homologous genome blocks are shaded gray and connected by lines.

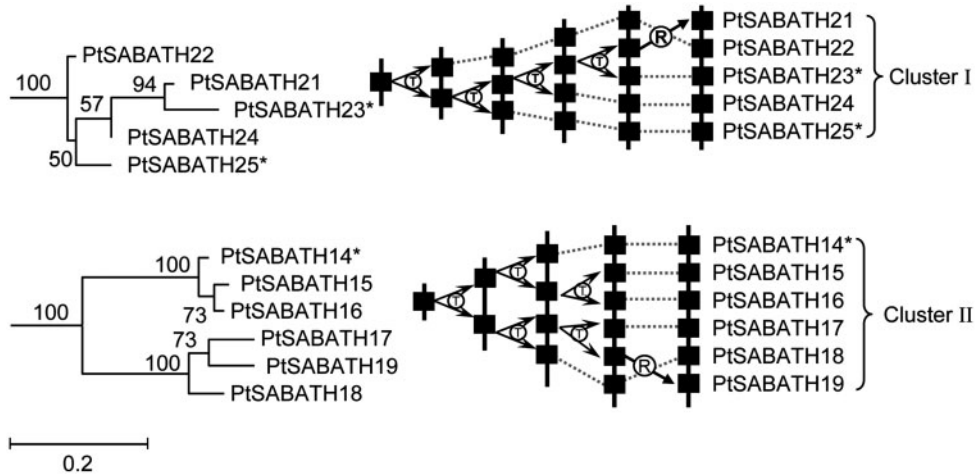


Fig. 2 Hypothetical evolutionary histories of the *Populus* SABATH genes in clusters I and II. The phylogenetic tree corresponds to the tree shown in **Fig.3A**. Numbers at each node in the phylogenetic tree are bootstrap values. The letters T and R within a circle indicate putative tandem duplication and rearrangement, respectively. Full-length pseudogenes are indicated with asterisks.

SABATH proteins showed 35.2–96.1% pairwise sequence identities. Seventeen class III SABATH proteins had 39.0–100% pairwise sequence identities. The protein sequences in these two SABATH classes were significantly different (independent sample *t*-test, $P < 0.001$) and only had 30.4–35.6% pairwise sequence identities. *Populus* SABATH class I contained only one gene. The sequence identities between class I and II and class I and III SABATH proteins were <32.7% and <33.5%, respectively. An independent sample *t*-test showed that the protein sequence identities within each class were higher than that between classes ($P < 0.001$). Sequence comparison further supports the class designations of the 28 full-length *Populus* SABATH genes.

SABATH gene family in land plants

To understand further the evolutionary history of this large gene family in land plants, this study identified 65 SABATH genes from five other land plant species, which represent the four major land plant lineages (bryophytes, lycophytes, gymnosperms and angiosperms) (Supplementary Table S2). We reconstructed the phylogenetic relationships among 93 SABATH genes from six land plant species using *Aspergillus* and

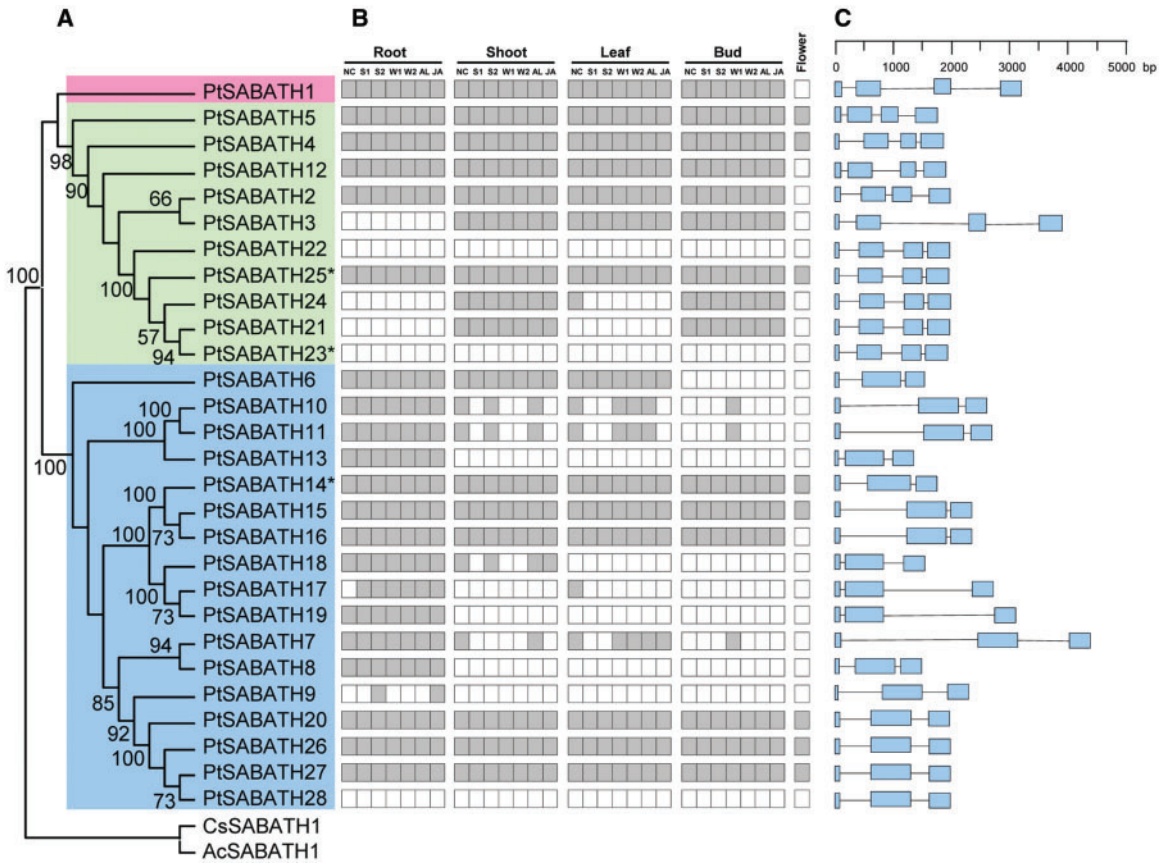


Fig. 3 Phylogenetic tree of *Populus* SABATH genes (A), their expression patterns (B) and gene structures (C). Numbers at each node in the phylogenetic tree are bootstrap values, and only values >50% are shown. Class I, II and III SABATH genes are shaded red, green and blue, respectively. In (B), the gray box indicates positive detection of gene expression in the corresponding tissues under normal growth conditions (NC) and the following stress treatments: SA at 2 h (S1), SA at 24 h (S2), alamethicin (AL), methyl jasmonate (JA), wounding at 2 h (W1) and wounding at 24 h (W2). In (C), exons and introns are shown as blue boxes and lines, respectively.

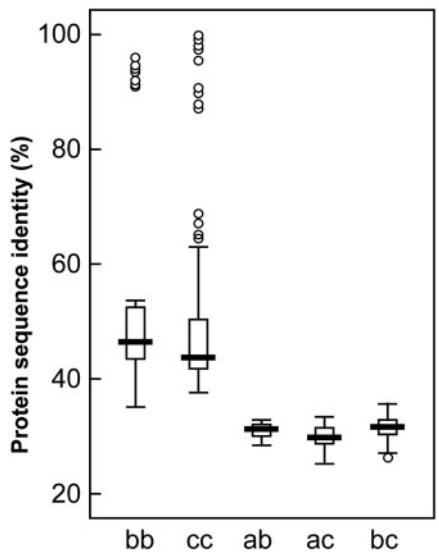


Fig. 4 Pairwise sequence identity of *Populus* SABATH proteins. bb and cc represent pairwise sequence identities within *Populus* class II and III SABATH proteins, respectively. ab, ac and bc represent pairwise sequence identities between *Populus* class I and II, class I and III, and class II and III SABATH proteins, respectively.

Coccomyxa SABATH genes as outgroups. Although the phylogenetic tree did not provide clear phylogenetic relationships of the SABATH genes of land plants, we observed that *Populus* SABATH genes fell into three distinct clades (clades I, II and III in **Fig. 5A**). The three clades contained 81.7% of SABATH genes from these six land plant species. Both clades I and II contained gymnosperm and angiosperm SABATH genes, while clade III only contained angiosperm SABATH genes. One *Populus* class I SABATH gene (PtSABATH1) was grouped into clade I (**Fig. 5B**), 10 *Populus* class II SABATH genes were grouped into clade II (**Fig. 5C**) and 17 *Populus* class III SABATH genes were grouped into clade III (**Fig. 5D**).

Difference in selective pressure between *Populus* class II and III SABATH members

Populus SABATH class II and III contained 10 and 17 members, respectively. We investigated if there was a significant change in the selective pressure between the SABATH classes using Phylogenetic Analysis by Maximum Likelihood (PAML) software. Two assumptions were tested: a one-ratio model that assumed the same ω (= d_N/d_S) ratio for two SABATH classes, and a two-ratio model that assumed the two SABATH classes

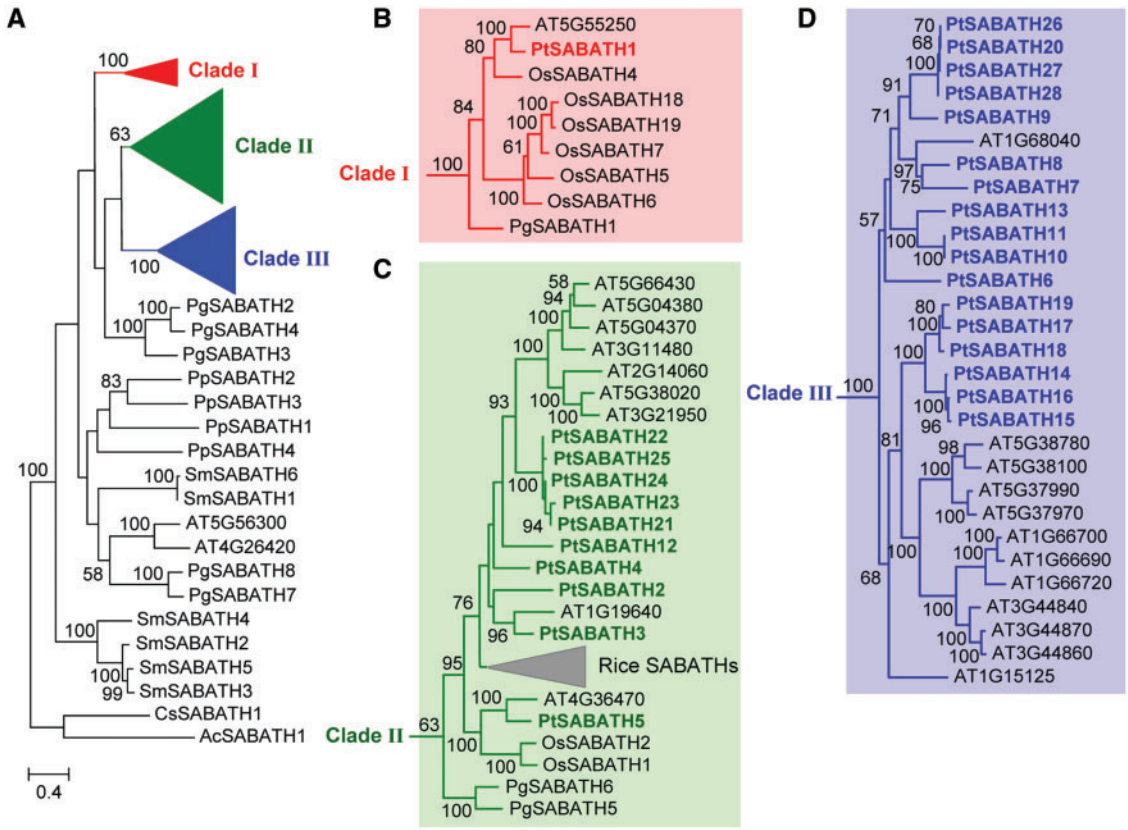


Fig. 5 Phylogenetic tree of 93 SABATH genes from six land plant species (A) and detailed phylogenetic relationships of SABATH genes in clade I (B), clade II (C) and clade III (D). Numbers at each node in the phylogenetic tree are bootstrap values, and only values >50% are shown. *Aspergillus clavatus* and *Coccomyxa subellipsoidea* SABATH genes were used as outgroups. *Populus* SABATH genes in clade I, II and III are colored red, green and blue, respectively.

had different ω ratios (**Table 1**; Supplementary Fig. S1). The log-likelihood values under the one-ratio and two-ratio models were $\log_e L = -17{,}462.4804$ and $-17{,}450.6965$, respectively. The likelihood ratio test (LRT) showed that the two-ratio model rejected the one-ratio model, indicating that the selective pressure differed significantly between the two classes ($P < 0.001$). Under the two-ratio model, the ω values for class II and III were 0.2047 and 0.3211, respectively, indicating that the *Populus* class III SABATH genes were under more relaxed selection constraints.

Expression patterns of *Populus* SABATH genes under normal growth conditions and abiotic stress

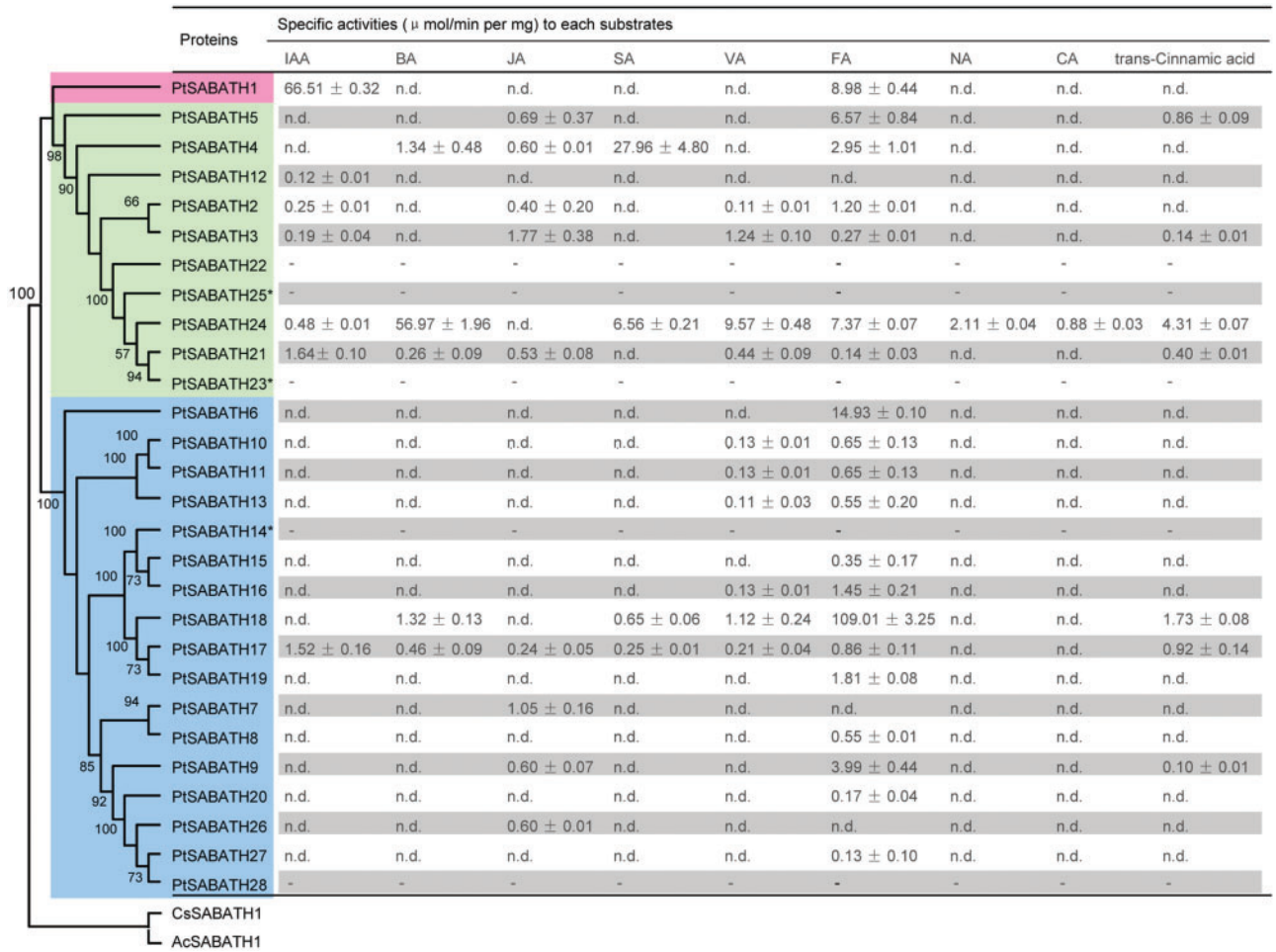
We investigated the expression patterns of *Populus* SABATH genes by RT–PCR. Tissues under normal growth conditions and in response to abiotic stress treatments [SA, alamethicin, methyl jasmonate (MeJA) and wounding treatments] were examined (**Fig. 3B**). Among the 28 *Populus* SABATH genes, eight genes (*PtSABATH4*, *5*, *14*, *15*, *20*, *25*, *26* and *27*) were expressed in all tissues under all growth conditions, while three genes (*PtSABATH22*, *23* and *28*) were not expressed in any tissue or in response to any treatment applied in this study. Seventeen genes were selectively expressed in a specific tissue and/or in response to a specific treatment. The pseudogenes *PtSABATH14* and *25* were expressed in all tissues examined, indicating that the regulation of these two pseudogenes is still functional.

Populus contained one class I SABATH gene (*PtSABATH1*). This gene was expressed in the root, shoot, leaf and bud tissues, but not expressed in the *Populus* flower. Class II contained 10 *Populus* SABATH genes; five (*PtSABATH21*, *22*, *23*, *24* and *25*) were arranged in a tandem (cluster I) on chromosome 19. In this cluster, *PtSABATH25* was expressed in all tissues under all growth conditions, *PtSABATH22* and *23* were not expressed in any tissues examined, and *PtSABATH21* and *24* were expressed in shoot and bud tissues. Expression data indicated that members in this cluster exhibited functional divergence.

Populus contained 17 class III SABATH genes. Substantial variation in expression patterns was found among the class III members (**Fig. 3B**). Among the class III SABATHs, five (*PtSABATH14*, *15*, *20*, *26* and *27*) were expressed in all tissues under all growth conditions. *PtSABATH28* was not expressed in any tissue examined in this study, and *PtSABATH9* was only expressed in the root following treatment with SA and MeJA. *PtSABATH17* was not expressed in root tissues, but was expressed following SA, alamethicin, MeJA and wounding treatments. Similar to the five class II SABATH genes in cluster I, six class III genes in cluster II showed marked divergence in expression.

Table 1 Results of the CODEML analyses of the selective pattern for SABATH genes

Tree	Branch model	Estimates of parameters	ln *L*	χ^2	*P*
Tree in **Supplementary Fig. S1**	One ratio	$\omega = 0.2755$ for class II and III SABATHs	−17,462.4804		
	Two ratios	$\omega_1 = 0.3211$ for class III SABATHs $\omega_0 = 0.2047$ for class II SABATHs	−17,450.6965	23.5677	<0.001



Proteins	Specific activities (μ mol/min per mg) to each substrates								
	IAA	BA	JA	SA	VA	FA	NA	CA	trans-Cinnamic acid
PtSABATH1	66.51 ± 0.32	n.d.	n.d.	n.d.	n.d.	8.98 ± 0.44	n.d.	n.d.	n.d.
PtSABATH5	n.d.	n.d.	0.69 ± 0.37	n.d.	n.d.	6.57 ± 0.84	n.d.	n.d.	0.86 ± 0.09
PtSABATH4	n.d.	1.34 ± 0.48	0.60 ± 0.01	27.96 ± 4.80	n.d.	2.95 ± 1.01	n.d.	n.d.	n.d.
PtSABATH12	0.12 ± 0.01	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
PtSABATH2	0.25 ± 0.01	n.d.	0.40 ± 0.20	n.d.	0.11 ± 0.01	1.20 ± 0.01	n.d.	n.d.	n.d.
PtSABATH3	0.19 ± 0.04	n.d.	1.77 ± 0.38	n.d.	1.24 ± 0.10	0.27 ± 0.01	n.d.	n.d.	0.14 ± 0.01
PtSABATH22	-	-	-	-	-	-	-	-	-
PtSABATH25*	-	-	-	-	-	-	-	-	-
PtSABATH24	0.48 ± 0.01	56.97 ± 1.96	n.d.	6.56 ± 0.21	9.57 ± 0.48	7.37 ± 0.07	2.11 ± 0.04	0.88 ± 0.03	4.31 ± 0.07
PtSABATH21	1.64± 0.10	0.26 ± 0.09	0.53 ± 0.08	n.d.	0.44 ± 0.09	0.14 ± 0.03	n.d.	n.d.	0.40 ± 0.01
PtSABATH23*	-	-	-	-	-	-	-	-	-
PtSABATH6	n.d.	n.d.	n.d.	n.d.	n.d.	14.93 ± 0.10	n.d.	n.d.	n.d.
PtSABATH10	n.d.	n.d.	n.d.	n.d.	0.13 ± 0.01	0.65 ± 0.13	n.d.	n.d.	n.d.
PtSABATH11	n.d.	n.d.	n.d.	n.d.	0.13 ± 0.01	0.65 ± 0.13	n.d.	n.d.	n.d.
PtSABATH13	n.d.	n.d.	n.d.	n.d.	0.11 ± 0.03	0.55 ± 0.20	n.d.	n.d.	n.d.
PtSABATH14*	-	-	-	-	-	-	-	-	-
PtSABATH15	n.d.	n.d.	n.d.	n.d.	n.d.	0.35 ± 0.17	n.d.	n.d.	n.d.
PtSABATH16	n.d.	n.d.	n.d.	n.d.	0.13 ± 0.01	1.45 ± 0.21	n.d.	n.d.	n.d.
PtSABATH18	n.d.	1.32 ± 0.13	n.d.	0.65 ± 0.06	1.12 ± 0.24	109.01 ± 3.25	n.d.	n.d.	1.73 ± 0.08
PtSABATH17	1.52 ± 0.16	0.46 ± 0.09	0.24 ± 0.05	0.25 ± 0.01	0.21 ± 0.04	0.86 ± 0.11	n.d.	n.d.	0.92 ± 0.14
PtSABATH19	n.d.	n.d.	n.d.	n.d.	n.d.	1.81 ± 0.08	n.d.	n.d.	n.d.
PtSABATH7	n.d.	n.d.	1.05 ± 0.16	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
PtSABATH8	n.d.	n.d.	n.d.	n.d.	n.d.	0.55 ± 0.01	n.d.	n.d.	n.d.
PtSABATH9	n.d.	n.d.	0.60 ± 0.07	n.d.	n.d.	3.99 ± 0.44	n.d.	n.d.	0.10 ± 0.01
PtSABATH20	n.d.	n.d.	n.d.	n.d.	n.d.	0.17 ± 0.04	n.d.	n.d.	n.d.
PtSABATH26	n.d.	n.d.	0.60 ± 0.01	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
PtSABATH27	n.d.	n.d.	n.d.	n.d.	n.d.	0.13 ± 0.10	n.d.	n.d.	n.d.
PtSABATH28	-	-	-	-	-	-	-	-	-
CsSABATH1									
AcSABATH1									

Fig. 6 Phylogenetic tree and enzymatic activities of *Populus* SABATH proteins. *Populus* SABATH class I, II and III members are shaded red, green and blue, respectively. Full-length pseudogenes are indicated by asterisks. Values shown are means ± SD, as calculated from three replicates. n.d., no activity detected. –, not detected.

Substrate specificities and activities of the *Populus* SABATH proteins

This study investigated the catalytic characteristics of *Populus* SABATH proteins, which may be related to their biological functions. Among the 28 *Populus* SABATH genes, three (*PtSABATH14*, *23* and *25*) were pseudogenes and two (*PtSABATH22* and *28*) were not expressed in any of the tissues examined. Thus, these five genes were not selected for protein expression and purification. As the proteins PtSABATH10 and 11 share an identical sequence, we selected only PtSABATH11 to investigate its biochemical function. In total, 22 SABATH proteins were selected to be expressed and purified. Ten recombinant PtSABATH proteins with a 6×His tag (PtSABATH1, 3, 5, 6, 7, 8, 15, 18, 20 and 27) were expressed as soluble proteins in *Escherichia coli*. However, 12 other recombinant 6×His-tagged PtSABATH proteins (PtSABATH2, 4, 9, 11, 12, 13, 16, 17, 19, 21, 24 and 26) were expressed as inclusion bodies. To increase their solubility, these 12 PtSABATH proteins were fused to the maltose-binding protein (MBP) tag and subsequently expressed as soluble proteins in *E. coli*. Thus, the activity of 22 purified SABATH proteins, including one, seven and 14 class I, II and III proteins, respectively, were examined in this study (**Fig. 6**). The substrate specificities of the purified *Populus* SABATH proteins were investigated with nine substrates: IAA, BA, JA, SA, vanillic acid (VA), farnesoic acid (FA), nicotinic acid (NA) coumalic acid (CA) and *trans*-cinnamic acid.

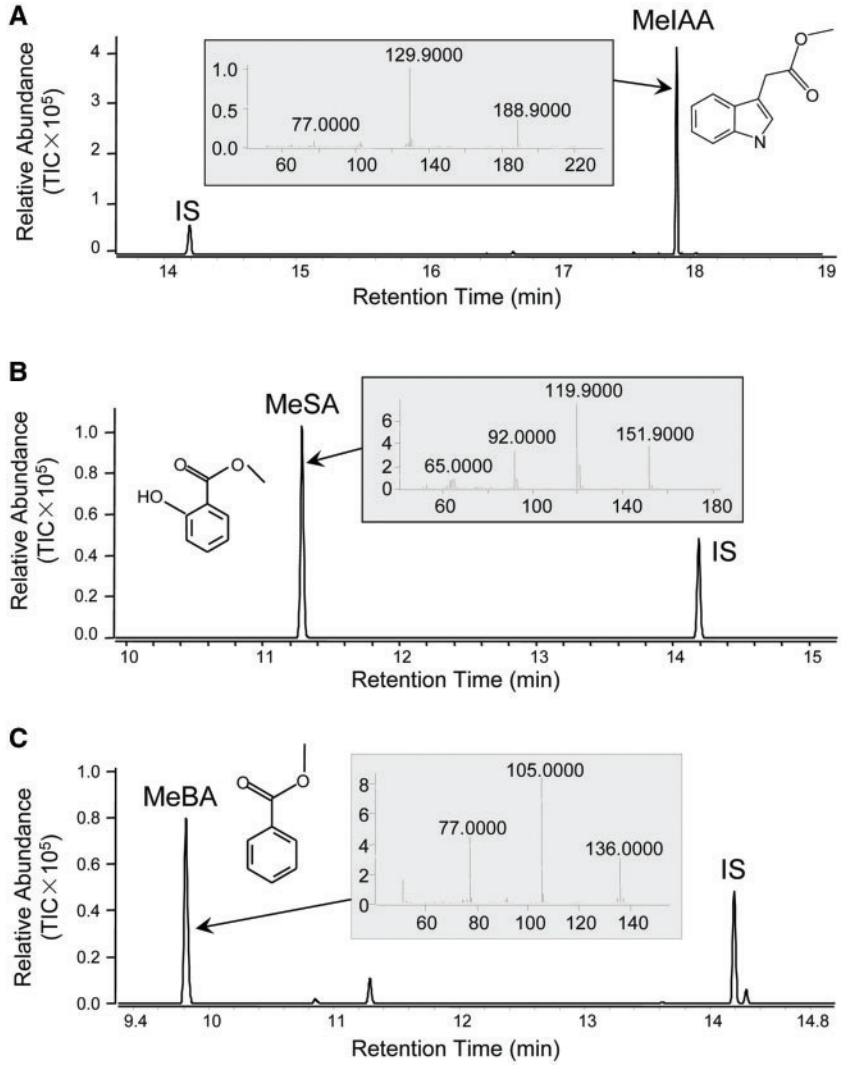


Fig. 7 Gas chromatography–mass spectrometry analysis of the products catalyzed by PtSABATH1 (A), PtSABATH4 (B) and PtSABATH24 (C). Insets show mass spectra of the assay products. Each experiment was repeated three times, and quantitative data were obtained by integration of peaks and comparison with standards. Menthyl anthranilate was used as the internal standard (IS).

Populus SABATH class I only contained one member (PtSABATH1). The PtSABATH1 protein did not show any activity towards substrates BA, JA, SA, VA, NA, CA or *trans*-cinnamic acid. However, PtSABATH1 exhibited high enzymatic activity (66.51 μmol min^{-1} mg^{-1}) towards IAA and weak activity (8.98 μmol min^{-1} mg^{-1}) towards FA. Previous studies have also shown that PtSABATH1 could catalyze formation of MeIAA using SAM as a methyl donor and IAA as a methyl acceptor (Zhao et al. 2007). IAA is the most abundant endogenous auxin, which is able to fulfill most of the actions of auxin involved in plant development and environmental responses. Among the 28 *Populus* SABATH proteins, PtSABATH2, 3, 12, 17, 21 and 24 had very weak activities against IAA (<1.64 μmol min^{-1} mg^{-1}). Only PtSABATH1 had high activity towards IAA, and can convert IAA to MeIAA (**Fig. 7A**), indicating that PtSABATH1 might play an important role in auxin homeostasis.

Populus SABATH class II contained 10 genes. Among these 10 genes, *PtSABATH22* and *23* were not expressed in any of the tissues examined, and *PtSABATH25* was classified as a pseudogene. The seven remaining genes were expressed as soluble proteins in *E. coli*. Comparative analyses of substrate specificities and activities of the *Populus* class II SABATH proteins revealed some notable features (**Fig. 6**). PtSABATH4 exhibited high enzymatic activity (27.96 μmol min^{-1} mg^{-1}) towards the substrate SA and weak activity (<2.95 μmol min^{-1} mg^{-1}) towards BA, JA and FA. However, PtSABATH4 did not show any activity towards IAA, VA, NA, CA or *trans*-cinnamic acid (**Fig. 7B**). Compared with other *Populus* SABATH proteins, PtSABATH4 showed at least 4.3-fold higher enzymatic activity towards the substrate SA. Previous studies also found that PtSABATH4 displayed the highest level of catalytic activity towards SA and a relatively low level of activity towards BA (Zhao et al. 2013). We found that PtSABATH24 had a wide substrate spectrum, exhibiting enzymatic activity towards eight of the substrates tested (IAA, BA, SA, VA, FA, NA, CA and *trans*-cinnamic acid), with the highest enzymatic activity (56.97 μmol

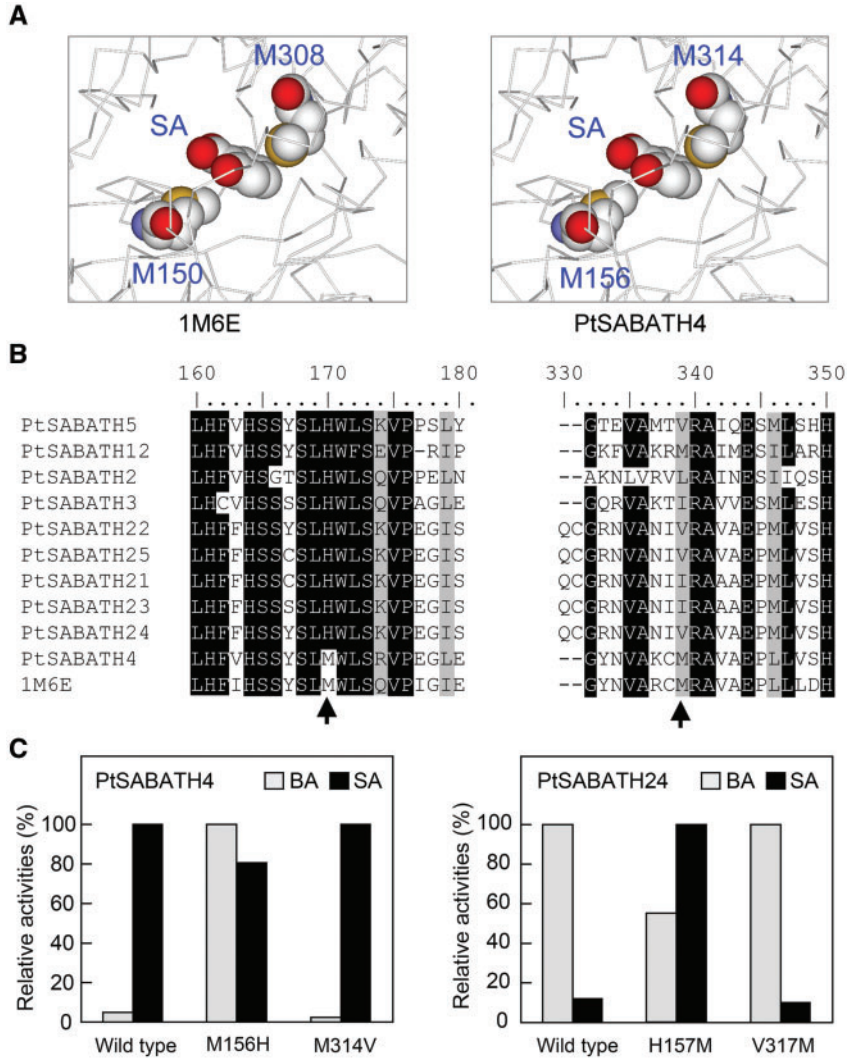


Fig. 8 Site-directed mutagenesis analysis of PtSABATH4 and PtSABATH24. (A) Structural comparison of the active sites of PtSABATH4 and the 1M6E protein. (B) Sequence alignment of 1M6E and *Populus* class II SABATH proteins. (C) Changes in enzymatic activity induced by mutations at the methionine residues (marked with arrows) relative to the wild-type proteins.

min^{-1} mg^{-1}) towards BA (**Fig. 7C**). Compared with other *Populus* SABATH proteins, PtSABATH24 showed at least 42.5-fold higher enzymatic activity towards BA.

Among *Populus* class III SABATH proteins, all proteins, with the exception of PtSABATH7 and 26, showed enzymatic activity towards the substrate FA. PtSABATH18 had the highest enzymatic activity (109.01 μmol min^{-1} mg^{-1}) towards FA, and showed at least 7.3-fold higher enzymatic activity towards FA compared with the other *Populus* SABATH proteins.

Site-directed mutagenesis analysis of the *Populus* SABATH proteins

The crystal structure of the *Clarkia breweri* SABATH protein 1M6E has been determined (Zubieta et al. 2003). The thioether moieties of Met150 and Met308 of 1M6E create a molecular clamp that encompasses both faces of the benzyl ring of SA (**Fig. 8A**) (Zubieta et al. 2003). Previous studies found that two amino acid sites determined the substrate discrimination (Barkman et al. 2007, Huang et al. 2012). To examine the functions of the two amino acid sites in substrate discrimination, we constructed the phylogenetic tree of the 28 *Populus* SABATH proteins and 1M6E. The phylogenetic tree showed that 1M6E and PtSABATH4 (*Populus* SABATH class II) were grouped together (Supplementary Fig. S2). We modeled the three-dimensional structure of PtSABATH4 based on the 1M6E crystal structure using the SWISS-MODEL Workspace (https://swissmodel.expasy.org/interactive). Similar to 1M6E, Met156 and Met314 of PtSABATH4 also created a molecular clamp that encompassed the benzyl ring of SA (**Fig. 8A**). A previous study showed that 1M6E exhibited high activity towards SA (Zubieta et al. 2003). Similar to 1M6E, PtSABATH4 also showed high activity towards SA. At the corresponding Met150 site of 1M6E, only 1M6E and PtSABATH4 shared an identical methionine residue, while the other *Populus* SABATH class II members contained a histidine residue (alignment position 170 in **Fig. 8B**). To investigate the functional characteristics

of this methionine residue, two sets of mutants were generated for biochemical assays. First, we mutated Met156 of PtSABATH4 to the histidine residue present in other *Populus* SABATH class II members. This mutation resulted in a switch from a preference for SA over BA in wild-type PtSABATH4 to a preference for BA over SA in the M156H mutant (**Fig. 8C;** Supplementary Table S3). Secondly, because PtSABATH24 exhibited the highest activity towards BA among *Populus* SABATH class II members, we mutated His157 of PtSABATH24 to a methionine residue. Interestingly, this mutation also resulted in a switch from a preference for BA over SA in wild-type PtSABATH24 to a preference for SA over BA in the H157M mutant (**Fig. 8C;** Supplementary Table S3). Taken together, the single amino acid changes (Met156 and His157) between PtSABATH4 and 24 resulted in a substrate switch.

At the Met308 site of the 1M6E protein, PtSABATH4 and 12 share an identical methionine residue, while the other *Populus* SABATH class II members contain a valine or isoleucine residue (alignment position 339 in **Fig. 8B**). PtSABATH12 did not exhibit activity towards the substrates SA or BA. Of the *Populus* SABATH class II proteins, PtSABATH4 and 24 showed the highest activity towards SA and BA, respectively. Thus, to investigate the functional characteristics of this methionine residue, PtSABATH4 and 24 were selected for site-directed mutagenesis analysis. We mutated Met314 of PtSABATH4 to the valine residue present in PtSABATH24, and mutated Val317 of PtSABATH24 to the methionine residue present in PtSABATH4. Both mutations resulted in decreased enzymatic activities towards the substrates SA and BA (Supplementary Table S3). However, these two mutations did not result in a substrate switch (**Fig. 8C**).

Discussion

Plant SABATH MTs are a multifunctional protein family involved in a broad range of biological processes that play critical roles throughout the life of the plant (Chen et al. 2003, Koo et al. 2007, Liu et al. 2011, Zhao et al. 2012). The *Populus* genome contains 28 SABATH genes. Tandem duplication was the major factor responsible for the rapid expansion of SABATH genes in *Populus*. In this study, we observed two SABATH gene clusters, cluster I and II. Cluster I contained five PtSABATH genes (*PtSABATH21*– *PtSABATH25*). *PtSABATH23* and *25* were pseudogenes, and *PtSABATH22* was not expressed in any of the examined tissues, suggesting that the function of these three genes may have been lost. PtSABATH24 had a wide substrate spectrum, and showed enzymatic activities towards eight of the substrates tested (IAA, BA, SA, VA, FA, NA, CA and *trans*-cinnamic acid). PtSABATH21 exhibited enzymatic activities towards six substrates (IAA, BA, JA, VA, FA and *trans*-cinnamic acid). These two proteins showed a partially overlapping substrate spectrum. Although both PtSABATH24 and PtSABATH21 showed enzymatic activity towards BA, PtSABATH24 had 219-fold higher enzymatic activity than PtSABATH21. These data indicated that functional divergences occurred among these five genes in cluster I. The functional divergence pattern of this gene cluster was also observed in the PtSABATH genes of cluster II.

In this study, *Populus* SABATH genes were divided into three distinct classes (**Fig. 3**). The protein sequence identities within each class were significantly higher than those between classes (**Fig. 4**). The class III SABATH genes had a two-intron/three-exon structure, while the class I and II SABATH genes contained a three-intron/four-exon structure (**Fig. 3C**). In addition, we found that *Populus* SABATH class I protein had specific activities towards IAA and class II proteins towards BA and SA, while class III proteins exhibited activity towards FA (**Fig. 6**). Thus, phylogenetic analysis, gene structures and biochemical characteristics of SABATHs supported three class designations. Although we did not obtain a clear phylogenetic relationship of the SABATH genes of land plants, we found that land plant SABATHs contained three distinct clades (clades I, II and III in **Fig. 5A**). *Populus* class I SABATH members fell into clade I, class II into clade II, and class III into clade III (**Fig. 5B–D**). The phylogenetic tree of land plant SABATH genes also supported three class designations of *Populus* SABATH genes.

SA is a key signaling molecule that mediates plant defense against a variety of pathogens in a number of plant species. SA can be methylated to form the volatile ester MeSA by salicylic acid carboxyl methyltransferase, a member of the SABATH MT family (Zubieta et al. 2003). MeSA is normally absent in plants but is dramatically induced upon pathogen infection (Seskar et al. 1998). MeSA is an important long-distance signal in systemic acquired resistance in tobacco and Arabidopsis (Park et al. 2007, Vlot et al. 2008). One Arabidopsis SABATH gene (At3g11480) can convert SA to MeSA (Chen et al. 2003). Expression of this gene correlated with induced MeSA emission in response to a variety of treatments, including moth herbivory and alamethicin treatment (Chen et al. 2003). Among the 28 *Populus* SABATH proteins, PtSABATH4, 17, 18 and 24 showed enzymatic activities towards SA. Compared with PtSABATH17, 18 and 24, PtSABATH4 showed at least 4.3-fold higher enzymatic activity towards SA, indicating that probably only PtSABATH4 exhibited salicylic acid carboxyl methyltransferase activity.

Among the 28 *Populus* SABATH proteins, the class II proteins PtSABATH4 and 24 showed the highest enzymatic activities towards the substrates SA and BA, respectively. The replacement of Met156 with a histidine residue in PtSABATH4 resulted in a switch in preference for SA over BA in the wild type to a preference for BA over SA in the M156H mutant. In contrast, the replacement of His157 in PtSABATH24 with a methionine residue resulted in an opposite switch from preference for BA over SA in the wild type to a preference for BA over SA in the H156M mutant. Our results indicated that a single amino acid switch in PtSABATH4 and 24 resulted in a substrate switch. Similar results were also observed in the *Datura wrightii*, *Hoya carnosa* and *Nicotiana suaveolens* SABATH families (Barkman et al. 2007, Huang et al. 2012). The kinetic analysis of *H. carnosa* and *N. suaveolens* SABATH proteins indicated that these substrate switches were a result of a change in the catalytic efficiency for the two competing substrates, SA and BA (Huang et al. 2012). The change, which improved the preference for one substrate, came at the expense of a reduced relative preference for the other.

Methyl farnesoate (MeFA), an analog of the insect juvenile hormone III (JH III), is believed to play important roles in enhancing reproductive maturation, maintaining juvenile morphology and influencing male sex determination in crustaceans (Duan et al. 2014). MeFA and JH III have been isolated from the plants *Cyperus iria* and *Cananga latifolia* (Toong et al. 1988, Yang et al. 2013). It has been suggested that the high content of JH III plays a role in plant defense against insects through its ability to interfere with insect development (Yang et al. 2013, Yang et al. 2006). SABATH proteins can convert FA to MeFA. The Arabidopsis SABATH protein At3g44860 showed high catalytic specificity towards FA (Yang et al. 2006). The mRNA levels of At3g44860 increase in response to the exogenous addition of several compounds previously shown to induce plant defense responses at the transcriptional level (Yang et al. 2006). In this study, we found that At3g44860 was grouped into clade III in the phylogenetic tree of land plant SABATH genes (**Fig. 5D**). Among the 28 *Populus* SABATH genes, all of the 17 members of class III were grouped into this clade (clade III in **Fig. 5C**). Compared with *Populus* SABATH class I and II, class III proteins showed high activity towards FA. For example, PtSABATH18 had the highest enzymatic activity (109.01 μmol min^{-1} mg^{-1}) towards the substrate FA. In particular, six *Populus* SABATH class III proteins (PtSABATH6, 8, 15, 19, 20 and 27) only showed specific enzymatic activity towards FA. Interestingly, molecular evolution analyses found that, compared with class II, *Populus* SABATH class III genes were under more relaxed selection constraints, indicating that *Populus* SABATH class III members might evolve new functions more freely. These data indicated that *Populus* SABATH class III members, even the members of clade III in **Fig. 5**, might have a specific function towards FA, and may play important roles in defense against insect herbivory.

IAA, the primary auxin in plants, is involved in many important physiological processes, including cell enlargement and division, tissue differentiation, and responses to light and gravity (Leveau and Lindow 2005). Free IAA, which makes up approximately 1% of IAA derivatives, accounts for IAA activity, whereas most IAA molecules exist as IAA-ester or IAA-amide conjugates for storage or degradation (Ljung et al. 2002). A previous study showed that IAA could be converted into non-polar MeIAA by IAA-methyltransferase-1 (IAMT1, At5g55250) in Arabidopsis. Overexpression of IAMT1 in Arabidopsis resulted in decreased IAA responsiveness and agravitropic growth, whereas *IAMT1* RNA interference transgenic plants showed leaf epinasty, decreased stature and decreased fertility (Qin et al. 2005). The Arabidopsis IAMT1 (At5g55250) gene was grouped into clade I in the phylogenetic tree of land plant SABATH genes (**Fig. 5B**). Among the 28 *Populus* SABATH genes, only PtSABATH1, which belonged to *Populus* SABATH class I, was grouped into clade I. Compared with other *Populus* SABATH proteins (PtSABATH2, 3, 12, 17, 21 and 24 had very weak activities towards IAA), PtSABATH1 showed at least 40.5-fold higher activity towards IAA. We identified eight SABATH genes from the *Picea glauca* genome. Among the eight *P. glauca* SABATH genes, only PgSABATH1 was grouped into clade I in the phylogenetic tree of land plants (**Fig. 5B**). A previous study showed that PgSABATH1 had a high level of catalytic activity towards IAA (Zhao et al. 2009). Among 23 rice SABATH genes, six (OsSABATH4, 5, 6, 7, 18 and 19) were grouped into clade I in the phylogenetic tree of land plants (**Fig. 5B**). OsSABATH4 displayed the highest level of catalytic activity towards IAA (Zhao et al. 2008). These data indicated that the SABATH genes in clade I had a specific function to convert IAA to MeIAA, and might play an important role in auxin homeostasis.

Gene duplication is considered a key process that generated the raw material for evolutionary innovation, although the mechanism of duplicate gene retention is still not well understood. This study re-annotated the members of the *Populus* SABATH MT gene family, and systematically examined their evolutionary and biochemical characteristics. *Populus* SABATH genes showed divergence in gene structure, gene expression patterns and biochemical characteristics. Additionally, site-directed mutagenesis revealed that a single amino acid change resulted in a substrate switch. Our findings provide new insights into the retention and subsequent functional divergence of duplicated genes.

Materials and Methods

Identification of SABATH genes

Previous studies identified 28 SABATH genes from the *P. trichocarpa* genome (Zhao et al. 2013). The re-identification of *Populus* SABATH genes was performed by TBLASTN searches of the *P. trichocarpa* genome database version 3.0 (http://phytozome.jgi.doe.gov/pz/portal.html) using 41 full-length rice SABATH proteins (Zhao et al. 2008) with default algorithm parameters. The candidate sequences were confirmed by protein structure analysis using a National Center for Biotechnology Information (NCBI) conserved domain search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). All candidate proteins have the Methyltransf_7 domain. To verify the sequence and gene structure, *Populus* SABATH genes were amplified from *Populus* cDNA, cloned into the *pEASY*-T3 vector (TransGen) and sequenced in both directions.

Phylogenetic analysis

The full-length SABATH protein sequences were aligned using MUSCLE software (http://www.drive5.com/muscle/) and further adjusted manually using BioEdit software (Hall 1999). The phylogenetic trees were reconstructed using an ML procedure in PhyML software with the Jones, Taylor and Thornton amino acid substitution model (Guindon and Gascuel 2003). One hundred bootstrap replicates were performed in each analysis to obtain the confidence support. *Aspergillus clavatus* SABATH (*AcSABATH1*, GenBank accession No. XP_001270502) and *C. subellipsoidea* SABATH (*CsSABATH1*, GenBank accession No. XP_005646815) were used as outgroups. To evaluate divergence of selective pressure between different classes of SABATH genes, the branch models of CODEML in PAML software (Yang 2007) were used to estimate ω (= d_N/d_S) under two assumptions: one-ratio model and two-ratio model. To verify which of the models best fit the data, LRTs were applied by comparing twice the difference in log-likelihood values between pairs of the models using a χ^2 distribution, with the degrees of freedom equal to the differences in the number of parameters between the models.

Expression of SABATH genes under normal conditions and abiotic stress

The cuttings and flowers of approximately 15-year-old *P. trichocarpa* trees were collected from the garden of Umeå University in Sweden. To investigate the expression patterns of the *Populus* SABATH genes under normal and abiotic stress conditions, the seedlings of *P. trichocarpa* were raised from the cuttings, and cultivated in potting soil for 2 months before treatment. For physical injury treatment, a sterile razor blade was used to produce three lateral incisions on

each side of the leaf midvein, and then the samples were collected at 2 and 24 h after wounding. For SA treatment, 5 mM SA solution was irrigated and sprayed. Samples (root, shoot, leaf and bud) were separately collected at 2 and 24 h after initiation of the treatment. For MeJA treatment, 1 µl of MeJA was dissolved in 200 µl of ethanol and a cotton tip was applied with the mixed MeJA solution. Subsequently, the poplar seedlings and cotton tip were placed in a 5 liter glass jar. Samples were collected 4 h after initiation of the treatment. For alamethicin treatment, the seedlings were irrigated and sprayed with 5 µg ml^{-1} alamethicin solution and placed in a glass beaker sealed with Saran wrap. Samples were collected 2 h after treatment. Three replicates were performed for each treatment. Total RNA was isolated using an RNAprep Pure Plant Kit (Tiangen) and reverse transcribed into cDNA using an RNA PCR Kit (AMV) version 3.0 (TAKARA BIO INC.). The *Populus actin* gene (GenBank accession No. XM_002316253.1) was used as an internal control. Specific primers were designed for PCR analysis (Supplementary Table S4) based on the multiple sequence alignment of *Populus* SABATH genes. PCR conditions were optimized to consist of an initial denaturation of 3 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 40 s at 65 °C and 1 min at 72 °C, and a final extension of 3 min at 72 °C. The PCR products from each sample were analyzed with 1% agarose gel electrophoresis and validated by DNA sequencing.

Expression and purification of recombinant *Populus* SABATH proteins

To investigate the enzymatic functions of *Populus* SABATH proteins, 22 proteins were selected for expression and purification. Ten *Populus* SABATH genes (*PtSABATH1*, *3*, *5*, *6*, *7*, *8*, *15*, *18*, *20* and *27*) were subcloned into modified ΔpET-30a expression vectors (Yang et al. 2009) to obtain an N-terminal 6 × His tag. The remaining 12 genes (*PtSABATH2*, *4*, *9*, *11*, *12*, *13*, *16*, *17*, *19*, *21*, *24* and *26*) were subcloned into pMAL-c5X expression vectors to obtain an MBP tag engineered for tighter binding to amylose. Site-directed mutagenesis of PtSABATH4 and 24 was performed as described previously (Zeng and Wang 2005). Using the plasmids of target genes as templates, the mutants were subcloned into the pMAL-c5X expression vectors. The primers used to construct the wild-type and mutant SABATH expression vectors are listed in Supplementary Table S5. The resultant plasmids were transformed into *E. coli* BL21 (DE3) and verified by sequencing. *Escherichia coli* BL21 (DE3) cells harboring plasmids with PtSABATH genes were cultured overnight, diluted 1 : 100 and grown until the optical density (A_{600}) reached 0.5. A final concentration of 0.1 mM isopropyl-β-d-thiogalactopyranoside was added to induce synthesis of the SABATH proteins. Then, 10 h after induction, cells were harvested by centrifugation (10,000 × *g*, 3 min, 4 °C), resuspended in binding buffer (20 mM sodium phosphate, 0.5 M NaCl and 20 mM imidazole, pH 7.4) for proteins with an N-terminal 6 × His tag or column buffer (20 mM Tris–HCl, pH 7.4, 200 mM KCl and 1 mM EDTA) for proteins with an MBP tag, and disrupted by cold sonication. In each case, the homogenate was then subjected to centrifugation (10,000 × *g*, 10 min, 4 °C). The resulting particulate material and a small portion of the supernatant were analyzed by SDS–PAGE. The supernatant containing the target proteins was loaded onto a nickel–Sepharose high performance column (GE Healthcare Life Sciences) or an amylose resin column (New England Biolabs) pre-equilibrated with binding or column buffer, respectively. The SABATH proteins bound to the nickel–Sepharose high performance column were eluted with a nickel–Sepharose elution buffer (20 mM sodium phosphate, 0.5 M sodium chloride and 500 mM imidazole, pH 7.4) and the SABATH proteins bound to the amylose resin column were eluted with an amylose resin elution buffer (20 mM Tris–HCl, pH 7.4, 200 mM KCl, 1 mM EDTA and 10 mM maltose). The buffers of the eluted recombinant proteins were replaced with the reaction buffer (50 mM Tris–HCl, pH 7.4 and 100 mM KCl) using a PD-10 column (Amersham Pharmacia Biotech).

Enzymatic assays of SABATH proteins

To determine the SAM-dependent MT activity towards nine carboxyl-containing substrates, 920 µl of reaction buffer, 20 µl of substrate (50 mM, dissolved in ethanol), 20 µl of SAM (50 mM, dissolved in ethanol and sulfuric acid) and 40 µl of purified protein (2.5–13 µM) were mixed and incubated for 30 min at 25 °C. The reaction was stopped by extraction with 1 ml of ethyl acetate or hexane, and an additional 5 µg of methyl-2-aminobenzoate was added as an internal standard. The upper layer was collected and concentrated under nitrogen gas, analyzed, and quantified with an Agilent 7890A gas chromatography (GC)–mass spectrometry system (Agilent). A DB-17ms column was used with helium as the carrier gas at a flow rate of 1.2 ml min^{-1}. The GC program was as follows: 50 °C for 2 min, ramp to 300 °C at 70 °C min^{-1}, followed by a 4 min hold at 300 °C. Eluted compounds were determined by comparison of GC retention time and mass fragmentation patterns with those of authentic standards. For MT activity analyses, peaks were quantified by area integration in conjunction with a standard curve. Assays were performed in triplicate for each substrate.

Supplementary Data

Supplementary data are available at PCP online.

Funding

This study was supported by grants from the National Natural Science Foundation of China [NSFC 31570664 and 31425006].

Disclosure

The authors have no conflicts of interest to declare.

References

Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., dePamphilis, C., et al. (2011) The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332: 960–963.

Barkman, T.J., Martins, T.R., Sutton, E. and Stout, J.T. (2007) Positive selection for single amino acid change promotes substrate discrimination of a plant volatile-producing enzyme. *Mol. Biol. Evol.* 24: 1320–1329.

Camps, M., Herman, A., Loh, E. and Loeb, L.A. (2007) Genetic constraints on protein evolution. *Crit. Rev. Biochem. Mol. Biol.* 42: 313–326.

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

D'Auria, J.C., Chen, F. and Pichersky, E. (2003) The SABATH family of MTs in *Arabidopsis* thaliana and other plant species. In Recent Advances in Phytochemistry. Edited by Romeo, J.T. pp. 253–283. Elsevier Science Ltd., Oxford.

Dixon, D.P., Hawkins, T., Hussey, P.J. and Edwards, R. (2009) Enzyme activities and subcellular localization of members of the *Arabidopsis* glutathione transferase superfamily. *J. Exp. Bot.* 60: 1207–1518.

Duan, Y., Liu, P., Li, J., Wang, Y., Li, J. and Chen, P. (2014) A farnesoic acid O-methyltransferase (FAMeT) from *Exopalaemon carinicauda* is responsive to *Vibrio anguillarum* and WSSV challenge. *Cell Stress Chaperones* 19: 267–377.

Falara, V., Akhtar, T.A., Nguyen, T.T., Spyropoulou, E.A., Bleeker, P.M., Schauvinhold, I., et al. (2011) The tomato terpene synthase gene family. *Plant Physiol.* 157: 770–789.

Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52: 696–704.

Hall, T.A. (1999) BioEdit: a user-friendly biological alignment sequence editor and analysis program for Windows 95/98/nT. *Nucleic Acids Symp. Ser.* 41: 95

Huang, R., Hippauf, F., Rohrbeck, D., Haustein, M., Wenke, K., Feike, J., et al. (2012) Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates. *Proc. Natl. Acad. Sci. USA* 109: 2966–2971.

Jansson, S. and Douglas, C.J. (2007) *Populus*: a model system for plant biology. *Annu. Rev. Plant Biol.* 58: 435–458.

Koo, Y.J., Kim, M.A., Kim, E.H., Song, J.T., Jung, C., Moon, J.K., et al. (2007) Overexpression of salicylic acid carboxyl methyltransferase reduces salicylic acid-mediated pathogen resistance in *Arabidopsis thaliana*. *Plant Mol. Biol.* 64: 1–15.

Lan, T., Wang, X.R. and Zeng, Q.Y. (2013) Structural and functional evolution of positively selected sites in pine glutathione *S*-transferase enzyme family. *J. Biol. Chem.* 288: 24441–24451.

Leveau, J.H. and Lindow, S.E. (2005) Utilization of the plant hormone indole-3-acetic acid for growth by *Pseudomonas putida* strain1290. *Appl. Environ. Microbiol.* 71: 2365–2371.

Liu, P.P., von Dahl, C.C. and Klessig, D.F. (2011) The extent to which methyl salicylate is required for signaling systemic acquired resistance is dependent on exposure to light after infection. *Plant Physiol.* 157: 2216–2226.

Liu, Y.J., Han, X.M., Ren, L.L., Yang, H.L. and Zeng, Q.Y. (2013) Functional divergence of the GST supergene family in *Physcomitrella patens* reveals complex patterns of large gene family evolution in land plants. *Plant Physiol.* 161: 773–786.

Ljung, K., Hull, A.K., Kowalczyk, M., Marchant, A., Celenza, J., Cohen, J.D., et al. (2002) Biosynthesis, conjugation, catabolism and homeostasis of indole-3-acetic acid in *Arabidopsis thaliana*. *Plant Mol. Biol.* 49: 249–72.

Ogawa, M., Herai, Y., Koizumi, N., Kusano, T. and Sano, H. (2001) 7-Methylxanthine methyltransferase of coffee plants. Gene isolation and enzymatic properties. *J. Biol. Chem.* 276: 8213–8218.

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

Qin, G., Gu, H., Zhao, Y., Ma, Z., Shi, G., Yang, Y., et al. (2005) An indole-3-acetic acid carboxyl methyltransferase regulates Arabidopsis leaf development. *Plant Cell* 17: 2693–2704.

Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., et al. (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319: 64–69.

Rivas-San Vicente, M. and Plasencia, J. (2011) Salicylic acid beyond defence: its role in plant growth and development. *J. Exp. Bot.* 62: 3321–3338.

Sauer, M., Robert, S. and Kleine-Vehn, J. (2013) Auxin: simply complicated. *J. Exp. Bot.* 64: 2565–2577.

Seskar, M., Shulaev, V. and Raskin, I. (1998) Endogenous methyl salicylate in pathogen-inoculated tobacco plants. *Plant Physiol.* 116: 387–392.

Toong, Y.C., Schooley, D.A. and Baker, F.C. (1988) Isolation of insect juvenile hormone III from a plant. *Nature* 333: 170–171.

Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596–1604.

Vlot, A.C., Liu, P.P., Cameron, R.K., Park, S.W., Yang, Y., Kumar, D., et al. (2008) Identification of likely orthologs of tobacco salicylic acid-binding protein 2 and their role in systemic acquired resistance in *Arabidopsis thaliana*. *Plant J.* 56: 445–456.

Yang, H., Kim, H.S., Jeong, E.J., Khiev, P., Chin, Y.W. and Sung, S.H. (2013) Plant-derived juvenile hormone III analogues and other sesquiterpenes from the stem bark of *Cananga latifolia*. *Phytochemistry* 94: 277–283.

Yang, X., Sun, W., Liu, J.P., Liu, Y.J. and Zeng, Q.Y. (2009) Biochemical and physiological characterization of a tau class glutathione transferase from rice (*Oryza sativa*). *Plant Physiol. Biochem.* 47: 1061–1068.

Yang, Y., Yuan, J.S., Ross, J., Noel, J.P., Pichersky, E. and Chen, F. (2006) An *Arabidopsis thaliana* methyltransferase capable of methylating farnesoic acid. *Arch. Biochem. Biophys.* 448: 123–132.

Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24: 1586–1591.

Zeng, Q.Y. and Wang, X.R. (2005) Catalytic properties of glutathione-binding residues in a tau class glutathione transferase (PtGSTU1) from *Pinus tabulaeformis*. *FEBS Lett.* 579: 2657–2662.

Zhao, N., Boyle, B., Duval, I., Ferrer, J.L., Lin, H., Seguin, A., et al. (2009) SABATH methyltransferases from white spruce (*Picea glauca*): gene cloning, functional characterization and structural analysis. *Tree Physiol.* 29: 947–957.

Zhao, N., Ferrer, J.L., Moon, H.S., Kapteyn, J., Zhuang, X., Hasebe, M., et al. (2012) A SABATH methyltransferase from the moss *Physcomitrella patens* catalyzes S-methylation of thiols and has a role in detoxification. *Phytochemistry* 81: 31–41.

Zhao, N., Ferrer, J.L., Ross, J., Guan, J., Yang, Y., Pichersky, E., et al. (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. and Chen, F. (2007) Molecular cloning and biochemical characterization of indole-3-acetic acid methyltransferase from poplar. *Phytochemistry* 68: 1537–1544.

Zhao, N., Yao, J., Chaiprasongsuk, M., Li, G., Guan, J., Tschaplinski, T.J., et al. (2013) Molecular and biochemical characterization of the jasmonic acid methyltransferase gene from black cottonwood (*Populus trichocarpa*). *Phytochemistry* 94: 74–81.

Zubieta, C., Ross, J.R., Koscheski, P., Yang, Y., Pichersky, E. and Noel, J.P. (2003) Structural basis for substrate recognition in the salicylic acid carboxyl methyltransferase family. *Plant Cell* 15: 1704–1716.