Large-Scale Biology Article

Evolutionarily Distinct BAHD N-acyltransferases are Responsible for Natural Variation of Aromatic Amine Conjugates in Rice

Meng Peng a,1, Yanqiang Gao a,1, Wei Chen a,1, Wensheng Wang a, Shuangqian Shen a, Jian Shi a, Cheng Wang a, Yu Zhang a, Li Zou a, Shouchuang Wang a, Wan Jian a, Xianqing Liu b, Liang Gong a, Jie Luo a,2

aNational Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China
bCollege of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
1These authors contributed equally to this work
2To whom correspondence should be addressed. E-mail: jie.luo@mail.hzau.edu.cn

Synopsis: Eight BAHD N-acyltransferases with diverse specificities, including 4 tryptamine/tyramine bifunctional N-acyltransferases, are present in the Gramineae.

Abstract

Phenolamides (PAs) are specialized (secondary) metabolites mainly synthesized by BAHD N-acyltransferases. Here, we report metabolic profiling coupled with association and linkage mapping of 11 PAs in rice (Oryza sativa). We identified 22 loci affecting PAs in leaves and 16 loci affecting PAs in seeds. We have identified 8 BAHD N-acyltransferases located on 5 chromosomes with diverse specificities, including 4 aromatic amine N-acyltransferases. We show that genetic variation in PAs is determined, at least in part, by allelic variation in the tissue specificity of expression of the BAHD genes responsible for their biosynthesis. Tryptamine hydroxycinnamoyl transferase 1/2 (Os-THT1/2) and tryptamine benzoyl transferase 1/2 (Os-TBT1/2) were found to be bifunctional tryptamine/tyramine N-acyltransferases. The specificity
of Os-THT1 and Os-TBT1 for agmatine involved four tandem arginine residues, which have not been identified as specificity determinants for other plant BAHD transferases, illustrating the versatility of plant BAHD transferases in acquiring new acyl acceptor specificities. With phylogenetic analysis, we identified both divergent and convergent evolution of N-acyltransferases in plants and we suggest that the BAHD family of tryptamine/tyramine N-acyltransferases evolved conservatively in monocots, especially in Gramineae. Our work demonstrates that omics-assisted gene-to-metabolite analysis provides a useful tool for bulk gene identification and crop genetic improvement.

Key words:

*Oryza sativa*, phenolamides, association mapping, linkage mapping, BAHD acyltransferase

INTRODUCTION

Plants produce a large number of specialized (secondary) metabolites that are crucial for their interactions with the ever-changing environment (Keurentjes, 2009; Saito and Matsuda, 2010). Much of the enormous diversity of specialized metabolites in plants comes from modifications or decorations of core structures by different tailing reactions such as acylation (Bontpart et al., 2015), glycosylation (Bowles et al., 2005), methylation (Lam et al., 2007), and so on (Schwab, 2003). In addition, many studies have revealed that plant specialized metabolites accumulate with tissue specificity (Schilmiller et al., 2010; McDowell et al., 2011; Kim et al., 2012; Toubiana et al., 2012; Watanabe et al., 2013; Dong et al., 2015; Wen et al., 2015).

Phenolamides (PAs), also referred to as hydroxycinnamic acid amides or phenylamides, are the hydroxycinnamoyl acylated products of the mono-, di-, or
tri-phenolic acid (coumaric, cafféic, or ferulic acid) substitutions of polyamines
(putrescine, spermidine, agmatine, tryptamine, and serotonin) or arylmonoamines
such as tyramine, octopamine, and anthranilate (Figure 1; Bienz et al., 2005; Edreva
et al., 2007; Bassard et al., 2010). PAs are a diverse group of specialized metabolites
found in many plants, including wheat (Triticum aestivum), barley (Hordeum vulgare),
rice (Oryza sativa), and maize (Zea mays) (Martin-Tanguy et al., 1978; Martin et al.,
1985; Facchini et al., 2002; Edreva et al., 2007; Bassard et al., 2010). Accumulating
evidence shows that PAs play an important role in plant defense responses against
pathogens and insect herbivores (Newman et al., 2001; Tanaka et al., 2003; Kaur et al.,
2010; Park et al., 2014). They are also suggested to play roles in sulphur starvation,
heat shock and protection against UV irradiation (Klapheck et al., 1983; Edreva et al.,
1998; Kaur et al., 2010), although clear evidence to support these roles is yet to be
reported (Fellenberg and Vogt, 2015).

Condensation of the hydroxycinnamoyl and amine moieties is the key step for the
biosynthesis of PAs, and is catalyzed by a diversity of hydroxycinnamoyl transferases
(Petersen, 2015). One of the first genes encoding this type of enzyme to be
characterized was the gene encoding tyramine: $N$-hydroxycinnamoyl transferase (THT;
Hohlfeld et al., 1995; Schmidt et al., 1999; Yu and Facchini, 1999; Back et al., 2001;
Von Roepenack-Lahaye et al., 2003; Kang et al., 2006). Analysis of their amino acid
sequences revealed that all these enzymes belong to the GCN5-related
$N$-acetyltransferases (GNAT) protein family, which also includes mammalian
spermidine/spermine acetyltransferase (SSAT) and several microbial antibiotic
$N$-acetyltransferases (Lu et al., 1996; Farmer et al., 1999; Kang et al., 2006). In
contrast to the situation where PAs of arylamine moieties are mainly biosynthesized
by the GNAT family of enzymes, the biosynthesis of the aliphatic type of
phenolamides is exclusively the domain of BAHD acyltransferases (D’Auria, 2006;
Bassard et al., 2010). The BAHD acyltransferase family was named according to the
first letter of each of the first four biochemically characterized enzymes of this family
BEAT (benzoyl alcohol $O$-acetyltransferase), AHCT (anthocyanin
$O$-hydroxycinnamoyl transferase), HCBT (anthranilate $N$-hydroxycinnamoyl/benzoyl
transferase), and DAT (deacetyl vindoline 4-O-acetyltransferase) (St-Pierre and De Luca, 2000). So far, enzymes involved in the biosynthesis of PAs have been identified from three (I, IV, and V) of the five clades of BAHD acyltransferases (D’Auria, 2006; Yu et al., 2009; Tuominen et al., 2011). ACTs (agmatine coumaroyl transferases) were characterized in Hordeum vulgare and in Arabidopsis thaliana (Burhenne et al., 2003; Muroi et al., 2009). Spermidine hydroxycinnamoyl transferases responsible for the production of hydroxycinnamoyl spermidines have been identified in Arabidopsis thaliana, Nicotiana attenuata, and Oryza sativa (Grienenberger et al., 2009; Luo et al., 2009; Onkokesung et al., 2012; Dong et al., 2015). More recently, PHTs (putrescine hydroxycinnamoyl transferases) have been identified in both dicots (Nicotiana) and monocots such as Zea mays, and O. sativa (Onkokesung et al., 2012; Wen et al., 2014; Chen et al., 2014). Interestingly, the BAHD enzymes, HCBT (anthranilate N-hydroxycinnamoyl/ benzoyltransferase) from Dianthus caryophyllus and HHT (hydroxycinnamoyl-CoA: hydroxyanthranilate N-hydroxycinnamoyltransferase) from Avena sativa, have also been reported to catalyze the acylation of arylmonoamines, anthranilate and 5-hydroxyanthranilate, respectively (Yang et al., 1997; Ishihara et al., 1998). Despite their diverse specificities, it remains to be shown whether members of this family are active on arylamine acyl acceptors such as tryptamine and serotonin.

Considering the large number and diverse activities of the BAHD proteins, a high-throughput strategy for BAHD identification needed to be developed. To date, the identification of hydroxycinnamoyl transferases has been achieved mostly by reverse genetics (Suzuki et al., 2002; Lepelley et al., 2007; Grienenberger et al., 2009; Luo et al., 2009; Elejalde-Palmett et al., 2015) and the role of these enzymes in determining the natural variation of PAs remains to be investigated. The combination of high-throughput metabolic profiling with large-scale genome re-sequencing by combined metabolite-based linkage and association studies have provided powerful forward-genetic strategies for dissecting the genetic and biochemical bases of metabolism in plants (Riedelsheimer et al., 2012; Gong et al., 2013; Chen et al., 2014; Wen et al., 2014; Alseekh et al., 2015; Wen et al., 2015; Matsuda et al., 2015). As for the BAHD family of enzymes in rice, OsAT1 has been reported to be involved in
responses to rice blight and blast (Mori et al., 2007). OsAT10 has proved to be a
*p*-coumaroyl:coenzyme A transferase involved in glucuronoarabinoxylan modification
(Bartley et al., 2013). We have already characterized two genes that contribute to
biosynthesis of *N*-feruloyl agmatine and *N*-feruloyl putrescine, respectively (Chen et
al., 2014), and identified a pair of tandem BAHD acyltransferases responsible for the
natural variation of spermidine-based PAs in rice (Dong et al., 2015).

Here, we describe metabolic profiling and subsequent association and linkage
mapping of non-spermidine-derived PAs in rice leaf and grain tissues to establish a
better understanding of the variation in PA types and contents that occurs in rice. Our
analyses disclosed tissue-specific differences in non-spermidine PAs, under genetic
control between leaf and grain, and defined 5 major loci associated with this variation
in PA levels in rice. Further molecular and biochemical studies resulted in the
identification of 8 distinct BAHD *N*-acyltransferases, which include 4
newly-identified bifunctional tryptamine/tyramine *N*-transferases. We show that four
tandem arginine residues play a role in determining the specificity of the
tryptamine/tyramine *N*-transferases for agmatine as the acyl acceptor. Our study
shows the evolution of tryptamine/serotonin acyltransferases in Gramineae and
demonstrates both divergent and convergent evolution of genes involved in
*N*-acylation of amines.

RESULTS

Natural Variation of PAs in Rice Leaf and Seed

To evaluate the extent of variation of PA types and levels in rice leaf and seed, we
measured the levels of non-spermidine-based PAs with different amine moieties, such
as agmatine, putrescine, and tryptamine, by LC-MS-based targeted metabolic
profiling using samples from a diverse worldwide collection of 480 *O. sativa*
accessions. Samples were collected from leaves at the five-leaf stage (Chen et al.,
2014; termed ‘leaf’ hereafter) and from mature seeds (termed ‘grain’ hereafter) for
profiling. A total of eleven PAs, eight in leaf and five in grain, were identified and their levels are shown in Table 1 and Supplemental Table 1. Levels of PAs varied substantially in both tissues between different accessions. Using leaves as an example, evaluation of the PA contents in *japonica* and *indica* by analysis of variance (ANOVA) showed significant variation in all PAs detected (*P* < 0.01) (Figure 2B), with the coefficients of variation (CV) ranging from 59.5 for feruloyl putrescine to 296.2 for cinnamoyl tryptamine (Supplemental Table 1). Hierarchical cluster analysis (HCA) grouped these PAs clearly into two clusters based on their relative abundance in different accessions (Figure 2A). PAs in cluster I were mainly putrescine- and agmatine-derived conjugates and could be further divided into three subclusters. In contrast, cluster II comprised exclusively tryptamine-derived PAs, suggesting likely specific control of this group of PAs. Similar accumulation patterns were observed in grains where putrescine- and tryptamine/serotonin-derived PAs grouped into separate clusters (Supplemental Figure 1). Broad-sense heritabilities (H²) ranged from 0.54 to 0.83 for all traits (Supplemental Table 1), suggesting that phenotypic variation is controlled primarily by QTLs.

**Association and Linkage Mapping of PA Traits in Rice**

To investigate the genetic control of natural variation in the major PAs in rice, a metabolite-based genome-wide association study (mGWAS) was performed using a world-wide collection of 480 rice accessions that has been genotyped by low-coverage re-sequencing (Zhao et al., 2015). The single nucleotide polymorphisms (SNPs) and the imputed genotypes of the association panel after removal of low-quality SNPs and those with minor allele frequency (MAF) < 0.05 resulted in a total of 4,342,652 SNPs, and a unified mixed linear model that controls for population structure and familial relatedness was used for the analysis. The genome-wide significant and suggestive thresholds were set to 6.15×10⁻⁸ and 1.23×10⁻⁶ after Bonferroni correction (see Methods; Supplemental Table 2) and loci with signals over the suggestive threshold were used for further analysis. GWASs of the PAs were
visualized in Manhattan plots (Figure 3; Supplemental Figures 2 and 3) with genomic coordinates (chromosomes 1-12) displayed along the x axis, and the negative logarithm of the associated $P$ value for each single nucleotide polymorphism displayed on the y axis (Gibson, 2010).

We identified 56 and 37 lead SNPs, corresponding to 22 and 16 loci in leaf and grain, respectively (Supplemental Table 2 and Supplemental Data set 1). Significant associations were detected for 9 out of the 11 PAs in at least one of the tissues tested, with an average of 2.8 and 3.2 associated loci per metabolite in leaf and grain, respectively. These loci, in general, showed large effects, up to 43.1%, with an average of 12.5% and 12.6% in leaf and grain, respectively (Supplemental Table 2 and Supplemental Data set 1), indicating that the levels of most PAs are controlled by a few major loci with large effects. Using leaves as an example, natural variation of agmatine-derived PAs was controlled by one major locus, L3 on chromosome 4, suggesting common genetic control of all these phenolamides (Supplemental Figure 2; Supplemental Data set 1). Similar results were obtained from the content of variation of putrescine-derived PAs, in which two major loci, L2 (on chromosome 9) and L5 (on chromosome 1) were observed (Supplemental Figure 2; Supplemental Data set 1).

In addition, we detected one common locus, L21 on chromosome 1, for coumaroyl putrescine and coumaroyl agmatine, suggesting common genetic control of these two PAs (Supplemental Figure 2; Supplemental Data set 1). This result is consistent with the close grouping between agmatine and putrescine-derived PAs, as revealed in the HCA. Similar results were observed for tryptamine-derived PAs that share a common locus on chromosome 11 (Supplemental Figure 2).

Biparental QTL analyses were performed using a recombinant inbred line (RIL) population generated from Zhenshan97 (ZS97) and Minghui63 (MH63) (Gong et al., 2013) to dissect independently the genetic basis of PA variation in rice. In total, 7 significant chromosome regions for loci with LOD values over 6.0 were identified (Figure 3; Supplemental Figure 4; Supplemental Data set 2), explaining 10.1 to 71.1% of the total variation in the population, and three of them were shared with the loci detected by GWAS. These loci showed, overall, relatively high resolution, ranging
from 0.12 to 1.31 Mb, with an average of 0.54 Mb, possibly due to the high resolution of the map generated by second-generation sequencing of the population.

To find out whether there are interactions between the significant loci (including suggestive loci), the pairwise epistatic interactions between the significant loci were calculated against the average accumulation of the individual PA. Eight significant interactions ($P < 0.05$) for four of the PAs in leaf were detected, ranging from one to three epistatic interactions for a single PA (Supplemental Table 3). For example, we observed a significant interaction between the two major loci L3 and L12 in determining the level of coumaroyl agmatine, in which the effect of the G and T alleles at L12 was dependent on the T allele at L3 (Supplemental Figure 5A). This result indicated that these loci may act sequentially in the accumulation of coumaroyl agmatine (Supplemental Figure 5B). Similar results were observed for cinnamoyl tryptamine (Supplemental Figures 5C and 5D). We also observed interactions between three loci (L2, L5 and L14) in controlling coumaroyl putrescine content (Supplemental Figures 5E-G). Based on these interactions and the accumulation pattern (Supplemental Figure 5H), a sequential model was suggested for the genetic control of the level of coumaroyl putrescine (Supplemental Figure 5I).

### Assignment of Candidate Genes in the BAHD Aciyltransferase Family

As an initial step to understand the genetic determinants of PAs, we focused on the possible role of acylation, the final and key step in PA biosynthesis, by looking for BAHD acyltransferases within the confidence intervals of the identified loci (Grienenberger et al., 2009; Luo et al., 2009; Onkokesung et al., 2012; Wen et al., 2014). Eleven candidate genes that were made up of 5 tandem pairs were assigned to the loci identified, with two of them having been partially characterized in a previous study (Figures 3 and 4B; Chen et al., 2014). We found a strong association ($P = 4.33 \times 10^{-17}$) between benzoyl tryptamine content and SNP sf1125035584 located on chromosome 11 (Supplemental Data set 1). Os11g42370 and Os11g42290, two highly similar (91% identity at amino acid level) genes that lie in tandem located 0 kb and 34
kb downstream from this SNP, were assigned to underlie this metabolic locus (Supplemental Table 4). This association was supported by linkage mapping for benzoyl tryptamine (Peak at 25.2 Mb and LOD = 47.9, Figure 3; Supplemental Table 4). Similarly, the significant association ($P = 1.23 \times 10^{-6}$) between SNP sf1012171661 lying 60 kb away from Os10g23310 and 26 kb away from Os10g23820, and the levels of coumaroyl serotonin suggested that this pair of genes are candidate BAHD transferases for this GWAS hit (Supplemental Table 4). In addition to tryptamine-derived PAs, candidates for major loci controlling PAs of other amines were identified. Significant association with three agmatine derivatives suggested that Os04g56900 controls the contents of coumaroyl agmatine and caffeoyl agmatine. Os04g56910 in this locus has already been reported to be a putative feruloyl agmatine transferase (Chen et al., 2014; Supplemental Table 4). Furthermore, Os09g37180 was assigned as a putative putrescine N-acyltransferase for its high similarity to Os09g37200, a reported feruloyl putrescine N-transferase in rice (Chen et al., 2014; Supplemental Table 4). A significant peak at 21.53 Mb on chromosome 9 in the QTL mapping of coumaroyl putrescine confirmed the GWAS result (Figure 3, Supplemental Table 4). We identified 3 more candidates located in tandem, Os06g08580, Os06g08610 and Os06g08640, from a significant locus on chromosome 6 (LOD = 6.9), controlling coumaroyl putrescine levels, not previously identified (Figure 3, Supplemental Table 4).

Phylogenetic relationships among the BAHD acyltransferases were then investigated by reconstructing the phylogenetic tree (Figure 4A), using a neighbor-joining algorithm, based on the amino acid sequences of our candidates and a collection of 27 reported plant BAHD acyltransferases (D’Auria, 2006; Grienenberger et al., 2009; Luo et al., 2009; Onkokesung et al., 2012; Chen et al., 2014; Dong et al., 2015). The resulting phylogenetic tree is consistent with that of Tuominen et al., an updated version of that from D’Auria, who sorted biochemically-characterized BAHD acyltransferases into eight major clades (D’Auria, 2006; Yu et al., 2009; Tuominen et al., 2011). Clade IV, which had only 5 members in the tree of Tuominen et al., was further expanded into IVa and IVb in following
analysis (Figure 4A). Candidates for 4 out of the 5 loci identified grouped in clade IV, which is enriched in monocot genes, while candidates for the remaining locus were grouped into clade Vb, which is mainly made up of enzymes closely related to hydroxycinnamoyl transferases (HCTs) responsible for the biosynthesis of chlorogenic acid and monolignols (D’Auria, 2006). Closer examination of the phylogeny in clade IVa revealed that Os04g56900 lies next to HvACT and Os04g56910, two agmatine coumaroyl transferase from barley and rice, respectively, consistent with the strong association between the Os04g56900 locus and the content of agmatine derivatives (Burhenne et al., 2003; Chen et al., 2014). Our analysis also placed Os09g37180 close to Na-AT1 and Os09g37200, putrescine acyltransferases from tobacco and rice, within clade IVa, strongly supporting the annotation of this candidate as a putrescine acyltransferase in rice (Onkokesung et al., 2012; Chen et al., 2014). Four newly identified BAHD candidates for tryptamine acyltransferases were clustered in clade IVb, where no characterized protein has been reported, and they were further divided into two groups according to their potential specificity on different acyl donors (Figure 4A). Interestingly, besides putative putrescine acyltransferases in clade IVa (Os09g37180), we also found three candidates for production of this type of PA in clade Vb (Os06g08580, Os06g08610 and Os06g08640), which contained acyltransferases specific for spermidine as the acyl acceptor and enzymes specific for other phenylpropanoids such as hydroxycinnamoyl-CoA:shikimate/quinate hydroxycinnamoyltransferase (HCT) and hydroxycinnamoyl-CoA: quinate hydroxycinnamoyltransferase (HQT) that have previously been characterized (Figure 4A). This phylogenetic analysis suggested likely convergent evolution of BAHD N-acyltransferases for the biosynthesis of putrescine-derived PAs in rice, which was confirmed by further characterization of these enzymes in vitro and/or in vivo in below.

**In vitro Acyltransferase Activities and Polymorphisms Underlying PA Contents**

To characterize the enzymatic properties of the 9 candidate acyltransferases in vitro
(Supplemental Table 2), recombinant proteins were expressed in *E. coli* strain BL21 (DE3) and purified by affinity chromatography. Acyltransferase activities were assayed for each enzyme using various CoA-thioesters as acyl donors and a range of possible amines as acyl acceptors. The identity of the reaction products was confirmed by retention times, UV spectra and the fragmentation patterns obtained from liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis, and relative catalytic activities and kinetic parameters with the different substrates were estimated.

As predicted from the association and linkage mapping (Figure 3, Supplemental Table 4), Os10g23310, Os10g23820, Os11g42290 and Os11g42370 displayed strong acyl transfer activities with tryptamine and its 5-hydroxylation derivative serotonin (Figures 5A and 5B; Tables 2 and 3; Supplemental Figures 6A-6C; Supplemental Table 5). Interestingly, all of them displayed similar $k_{cat}$ values for tryptamine, serotonin and tyramine, indicating that these four proteins are bifunctional tryptamine/tyramine $N$-transferases (Tables 2 and 3). Additionally, Os10g23310 also showed mild activity for an aliphatic amine (agmatine) (Figures 5A and 5B). In terms of acyl donors, both Os10g23310 and Os10g23820 exhibited high selectivity towards hydroxycinnamoyl-CoAs such as $p$-coumaroyl- and caffeoyl-CoA, and were therefore named tryptamine hydroxycinnamoyl transferase 1 (Os-THT1) and Os-THT2, respectively (Figure 5C; Table 2; Supplemental Table 5). By contrast, Os11g42370 was almost entirely specific for benzoyl-CoA, while Os11g42290 exhibited high activities with both benzoyl- and $p$-coumaroyl-CoA (Figure 5C; Table 3; Supplemental Table 5). Based on these results, Os11g42290 and Os11g42370 were termed tryptamine benzoyl transferase 1 (Os-TBT1) and tryptamine benzoyl transferase 2 (Os-TBT2), respectively. In searching for possible functional polymorphism(s) within these genes, we observed strong associations between sixteen SNPs in the coding region of Os-TBT2 and the levels of benzoyl tryptamine ($P < 10^{-30}$; Figure 6B; Supplemental Data set 3). Of these, nine allelic mutations caused significant changes in amino acids, resulting in 11-fold differences in levels of benzoyl tryptamine between the two groups (alleles I and II; Supplemental Data set 3).
Similarly, we found that four non-synonymous SNPs (sf1125000703, sf1125001414, sf1125001516, and sf1125001660) in Os-TBT1 were significantly associated with the content of cinnamoyl tryptamine ($P < 10^{-14}$; Figure 6A; Supplemental Data set 3). These SNPs were in strong LD and resulted in over 4-fold differences in cinnamoyl tryptamine content, suggesting these could be functional polymorphisms underlying this GWAS hit. In addition, the multi-sequences of Os-TBT1 and Os-TBT2 were aligned between ZS97 and MH63 accessions. We found three synonymous SNPs, an insertion of 4 consecutive amino acids, together with a 7-bp deletion that resulted in a frame shift in Os-TBT1 in MH63 compared with the same gene in ZS97 (Figure 6C; Supplemental Data set 4). In addition, sequence comparison disclosed 19 non-synonymous SNPs that led to changes in amino acid with different polarity and an insertion of one amino acid in Os-TBT2 in the MH63 accession compared to the same gene in the ZS97 accession (Figure 6D; Supplemental Data set 4). These results indicated that the truncation of Os-TBT1 and the large number of differences in Os-TBT2 in MH63 could be responsible for the almost complete absence of benzoyl tryptamine and benzoyl serotonin in this accession.

To evaluate the expression of the candidate genes in leaves and in seeds of the two parents, transcript levels of these genes were quantified by qRT-PCR (Supplemental Data set 5). We found that Os-THT1 and Os-THT2 were expressed specifically in seed, consistent with the specific accumulation of coumaroyl serotonin in grain. Similarly, Os-TBT1 and Os-TBT2 were highly expressed in seed and showed different transcription abundances in ZS97 and MH63, in line with the specific accumulation of benzoyl serotonin in the grain of these two accessions. These two loci were specifically mapped in grain by association and linkage mapping (Figure 3).

In addition, enzyme assays of the remaining candidate proteins showed that Os06g08580 and Os06g08610 catalyze the biosynthesis of coumaroyl putrescine (termed Os-PHT1 and Os-PHT2, respectively; Supplemental Figures 6D and 6E). However, no activities were observed for Os04g56900 and Os09g37180 in vitro.
Confirmation of Four Genes as PA Transferases \textit{in vivo}

We went on to investigate the functions of the above genes \textit{in vivo} by testing one candidate for each type of BAHD acyltransferase. To this end, \textit{Os04g56900} for agmatine acyltransferase, \textit{Os09g37180} for putrescine acyltransferase, and \textit{Os-TBT1} (wide substrate specificity) and \textit{Os-TBT2} (narrow substrate specificity) for tryptamine/serotonin acyltransferase were selected. Considering the redundancy and close linkage between the two members of each tandem pair, we adopted the gain-of-function strategy to establish their \textit{in vivo} functions, by overexpressing each of them individually in ZH11, which has been widely used as a host in rice transformation and in previous work on PAs in rice (Chen et al., 2014; Wen et al., 2014; Dong et al., 2015), under the control of the maize \textit{ubiquitin} promoter. PA-targeted profiling was performed by LC-MS to determine the differences in the quantity and profiles of PAs between wild type and the transgenic lines for each individual construct.

Despite inactivity \textit{in vitro}, over-accumulation of similar amounts of coumaroyl, caffeoyl, and feruloyl agmatine in the \textit{Os04g56900}-overexpressing plants suggested that this gene encodes an agmatine hydroxycinnamoyl transferase (designated as \textit{Os-AHT1}) that shows low selectivity towards the different hydroxycinnamoyl-CoAs (Figure 7B). Similarly, we observed substantially enhanced accumulation of hydroxycinnamoyl putrescine, especially coumaroyl putrescine in the overexpression lines of \textit{Os09g37180} (designated as \textit{Os-PHT3}), which is consistent with the mapping results (Figure 7A). Metabolic profiling showed an over 20-fold increase of benzoyl tryptamine and serotonin and much lower enhancement of cinnamoyl tryptamine in \textit{Os-TBT2}-overexpressing plants than in the wild type, which mirrored its activity \textit{in vitro} (Figure 7D). However, although \textit{Os-TBT1} displayed higher activity towards benzoyl-CoA than towards cinnamoyl-CoA \textit{in vitro}, overexpression of this gene \textit{in planta} resulted in 10-fold more cinnamoyl tryptamine compared to the wild type counterpart, with no big change in the benzoyleted tryptamine content compared to control plants, which could be in part due to the differences in CoA ester pools \textit{in vivo}.
Combining prior knowledge of pathway architecture and the activities of the eight BAHD genes identified in this study, an updated phenolamide biosynthetic pathway was constructed, showing the important role of the N-acyltransferase gene network in the PA accumulation process in rice (Figure 8).

**Characterization of the Amine-Binding Specificity Domain**

The functional assignment of eight N-acyltransferases in clade IV presented us with an opportunity to examine the characteristics of this barely characterized subgroup of BAHD acyltransferases (D’Auria, 2006; Yu et al., 2009; Tuominen et al., 2011). When protein sequences of this clade were subjected to motif analysis using MEME (Bailey et al., 2015), we found the two BAHD family-conserved functional domains, HXXXD and DFGWG (Ma et al., 2005; D’Auria, 2006), and two clade IV-specific motifs (Supplemental Figures 7A and 7B). Considering the distinct specificity between clade IVa and IVb (exclusively using aliphatic or aromatic amines as acceptors, respectively; Tables 2 and 3), further motif searching coupled with visual inspection revealed a number of subclade-specific motifs. Three motifs were observed in clade IVa, and two of them (VDGRAR and FMPSYLP) were located near the GN and DFGWG domain that compose a solvent channel for substrate binding (Ma et al., 2005; Supplemental Figures 7C and 7D). Similarly, two conserved motifs were highlighted in clade IVb (Supplemental Figures 7E and 7F). One of them (VRVAVNCE) was near the GN domain, whereas the remaining one (RRRR) has not been reported for BAHD family enzymes, suggesting it might play a previously uncharacterized role in determining substrate specificity of BAHD acyltransferases.

To investigate the function of this newly identified motif (RRRR), an alignment of amino acid sequences of these domains was generated (Figure 9A). From this, it is clear that Os-TBT1/2, with no activity on agmatine, contains four tandem arginine residues in their sequences, whereas proteins, Os-THT1, Os-AHT1, 04g56900 and Hv-ACT, in which the RRRR motifs are absent, showed agmatine-binding activities.
This observation suggested that the four tandem arginine motif contributes to acceptor
specificity. Therefore, several mutants were generated including insertion and deletion
of the RRRR motif from Os-THT1 and Os-TBT1 in corresponding positions and the
kinetic constants of these mutant proteins were determined. As shown in Figure 9B,
wild-type Os-THT1 and Os-THT1::RRRR (RRRR insertion mutant) exhibited similar
$K_m$ values for serotonin but distinct affinities for agmatine, with Os-THT1::RRRR
showing a much higher $K_m$ (too large to be determined precisely) for agmatine than
that of the unmutated form of Os-THT1. Os-THT1 with four alanine amino acid
residues replacing the four arginine amino acid residues was also expressed, and
coumaroyl agmatine was detected in the reaction product of this enzyme, suggesting
that the RRRR motif is specific for preventing activity with agmatine. By contrast,
Os-TBT1ΔRRRR (RRRR deletion) restored activity on agmatine (similar $K_m$ value to
Os-THT1) whereas wild-type Os-TBT1 could not use agmatine as its acyl acceptor.
These results indicated that deletion of RRRR residues in $N$-terminal portions is
critical for enabling wild-type Os-TBT1 to accept agmatine.

**DISCUSSION**

The variation of almost all PAs in this study was controlled by a few loci with large
effects, which is in accordance with a handful of previous studies on specialized
metabolites and distinct from those described for QTLs of primary traits (Fridman et
al., 2004; Rowe et al., 2008; Schauer et al., 2008; Chan et al., 2010; Joseph et al.,
2013; Alseekh et al., 2015; Wen et al., 2015), for which a large number of minor loci
are usually detected. Tissue-specific accumulation of metabolites, representing one of
the main contributions to metabolite diversity, is of great interest to plant scientists
(Schilmiller et al., 2010; Chan et al., 2011; Watanabe et al., 2013). Metabolic diversity
between different tissues might arise from differences in spatial-temporal expression
of genes (Kim et al., 2012; Thatcher et al., 2014) and in some cases by duplicate
genomes that display spatially or temporally distinct expression patterns (Toubiana et al.,
2012; Matsuba et al., 2013). By comparing genetically controlled natural variation of
PAs in different tissues at high resolution, we have shown that although coordinated genetic control across various tissues can be observed for some PAs, the majority of the loci were obtained in a tissue-specific manner (Supplemental Table 4), suggesting distinct regulation underlying the metabolic readout between tissues. Furthermore, the distinct, genetically-controlled natural variation of PAs in different tissues must be determined, at least in part, by the allelic variation of the tissue-specifically expressed genes responsible for their biosynthesis (Supplemental Data set 5). The same likely holds true for the mechanism underlying developmental stage- and treatment-specific regulation of metabolism in Arabidopsis thaliana (Chan et al., 2011).

Unlike complex quantitative traits (such as flowering time) for which little epistasis has been detected (Buckler et al., 2009; Morris et al., 2013; Zuo and Li, 2014), variation in levels of metabolites, especially specialized metabolites, is predominately controlled by epistasis in both RILs (Kliebenstein et al., 2002; Loudet et al., 2003; Calenge et al., 2006; Lisec et al., 2008; Rowe et al., 2008; Gong et al., 2013) and natural populations (Chan et al., 2011; Chen et al., 2014; Sauvage et al., 2014; Dong et al., 2015; Matsuda et al., 2015). Similar results were observed between some of the loci identified in our study (Supplemental Figure 5; Supplemental Table 3).

Association mapping by GWAS is suitable for screening a large number of accessions for common variants within the population at relative high resolution (Huang et al., 2010; Huang and Han, 2012), while linkage mapping using artificial populations such as RILs and ILs is likely to be more powerful in identifying alleles with low frequency or small effects in the population (Fridman et al., 2004; Keurentjes et al., 2006; Schauer et al., 2006; Matsuda et al., 2012). Joint linkage and association mapping as illustrated in the study proved to be useful not just in cross-validating results from one another, but also complimenting each other in identifying new candidate loci (Luo, 2015).

The N-acyltransferases identified so far are mainly members from the BAHD and GNAT families of proteins. While BAHD proteins are responsible for the acylation of aliphatic and arylmonoamine acceptors, GNAT enzymes are active on arylamines (Bassard et al., 2010). Although they are from two distinct protein families that show
low sequence identity, the BAHD $N$-acyltransferases, Os-THTs and Os-TBTs, identified in our study demonstrate that proteins from both families can act on arylamines (Tables 2 and 3; Jang et al., 2004; Kang et al., 2006). Comparing the specificity of the enzymes between the two families showed that whereas GNAT members are highly specific for tyramine, the BAHD enzymes may be regarded as bifunctional tryptamine/tyramine $N$-acyltransferases. Examination of the specificity within the BAHD enzymes revealed further variation in acyl donor selectivity and identified the enzymes responsible for the biosynthesis and natural variation of benzoylated tryptamine/serotonin. Previous crystal studies have identified several amino acid residues in the $N$- and $C$- terminal regions of BAHD acyltransferases important for differential acyl-acceptor specificity (Ma et al., 2005; Unno et al., 2007; Lallemand et al., 2012; Walker et al., 2013). Here, comparative sequence analysis with in vitro assays using mutant enzymes led to the identification of a RRRR motif that is sufficient to transform agmatine utilization (Figures 9A and 9B), which is similar to the conserved arginine-rich motifs in RNA-binding proteins (Lazinski et al., 1989; Zapp et al., 1991; Yuryev et al., 1996). This integrated approach has also been applied to identify key residues of enzymes involved in coumaroyl serotonin, terpene, acylsucrose biosynthesis (Kang et al., 2006; Kang et al., 2014; Fan et al., 2016) and may be regarded as a general strategy in the structure-function study of enzymes.

Gene (and genome) duplication is a major driving force for the acquisition of new functions in plant secondary metabolism (PSM) (Pichersky and Gang, 2000; Ober, 2005). Since the selectivity on hydroxycinnamoyl-CoA is present in BAHD enzymes of all plant species (D’Auria, 2006; Bassard et al., 2010), from moss to legumes, we assume that Os-THTs were derived from one ancient BAHD protein that preserved its selectivity for acyl donors and evolved new catalytic capabilities in selecting tryptamine/serotonin as acyl acceptors. Protein BLAST searches in Phytozome and NCBI databases did not find orthologs among the 29 core Eudicotyledons and three Gymnosperms (even at 40% identity), but did detected them in Amborella trichopoda, Nelumbo nucifera, and in one primitive Eudicotyledon, Aquilegia coerulea (Figure 10B). Interestingly, at least two THT-like candidates were found in each of the 6
monocotyledons (Figure 10B), with 6 orthologs from *Zea mays*, *Brachypodium distachyon*, *Setaria italica* and *Hordeum vulgare* exhibiting THT/TBT activities (Figure 10A). This result is in line with the data obtained from phylogenetic analysis of Os-THT/TBT-like proteins (Figure 10B) that showed the unique exist of THT/TBT-like BAHD proteins in Gramineae, suggesting that the BAHD family of Os-THT/TBT-like proteins evolved conservatively in Gramineae.

The convergent evolution of BAHD enzymes has been reported, where spermidine acyltransferases from distinct subclades of the BAHD family exhibit similar biochemical function (Luo et al., 2007). Here, similar results were observed for the convergent evolution of putrescine acyltransferases (Figure 4A). It is interesting to find that, although biochemically redundant, they are biologically distinct (responsible for the variation of putrescine-derived PAs in leaf and grain, respectively) due to their specific expression patterns (Supplemental Data set 5). Our results suggest that tryptamine/tyramine acyltransferases are restricted to monocots or more specifically Gramineae, while data obtained from previous investigations have identified GNAT proteins to be present in both monocots and dicots (Figure 10B; Negrel and Martin, 1984; Louis and Negrel, 1991; Schmidt et al., 1999; Ishihara et al., 2000). The biological significance underpinning the convergent evolution of aryldiamine acyltransferases from both BAHD and GNAT families of proteins, especially within Gramineae, requires further investigation. The 3D structures of some BAHD acyltransferases have been resolved (Ma et al., 2005; Unno et al., 2007; Garvey et al., 2008; Lallemand et al., 2012; Manjasetty et al., 2012; Walker et al., 2013), but no structures are yet available for N-acyltransferases. Elucidation the structure of BAHD N-acyltransferase will greatly facilitate understanding these divergent and convergent evolutionary paths.

**METHODS**

**Plant Materials and Metabolite Profiling**

The two biological sample sets of two populations (Supplemental Data sets 6 and 7)
were grown in normal rice-growing season (2012 and 2013) at Huazhong Agricultural University, Wuhan, China. All seeds were planted in a seed bed in mid-May, and transplanted to the field in mid-June. The planting density was 16.5 cm between plants in a row, and the rows were 26 cm apart. For each sample set, the accessions were grown in randomized design and field management, including irrigation, fertilizer application and pest control, followed essentially the normal agricultural practice. The leaves were harvested at five-leaf stage (only for RIL population) and hundreds of individual mature seeds were collected (for two populations) using liquid nitrogen from 3 different plants per line in the field for metabolite extraction. The materials were then combined to make one biological replicate of each sample. The freeze-dried samples were extracted as previously described before analysis using an LC-ESI-MS/MS system (Chen et al., 2013). A previously described relative quantification method were used to analyze samples (Chen et al., 2014). Data collected from the two biological sample sets were then used to calculate $H^2$.

Genome-wide Association and Linkage Map Analyses

The sequence data are from the website RiceVarMap (http://ricevarmap.ncpgr.cn). Only SNPs with MAF $\geq 0.05$ and the number of varieties with the minor allele $\geq 6$ in a (sub) population were used to carry out GWAS. GWAS methods were described previously (Chen et al., 2014). The calculated genome-wide significance threshold, based on the original Bonferroni calculation of 0.05/Me was $6.15 \times 10^{-8}$, the calculated genome-wide suggestive threshold, based on the original Bonferroni calculation of $1/Me$ was $1.23 \times 10^{-6}$, Me stands for the effective number of independent SNPs (Chen et al., 2014). QTL mapping was performed using the ZS97×MH63 RIL population as described in previous work (Gong et al., 2013).

Statistical Analysis

The values of coefficient of variation (CV) were independently calculated for each metabolite (using the mean of the biological replicates of the untransformed m-trait
data) as below: $\sigma/\mu$, $\sigma$ and $\mu$ represent the standard deviation and the mean of each metabolite in the population, respectively. Broad-sense heritability ($H^2$) was calculated using the following equation by treating accessions as a random effect and the biological replication as a replication effect using one-way ANOVA: $H^2 = \frac{\text{Var}(G)}{\text{Var}(G)+\text{Var}(E)}$, where $\text{Var}(G)$ and $\text{Var}(E)$ represent variance derived from genetic and environmental effects, respectively. Linkage disequilibrium (LD) was estimated using standardized disequilibrium coefficients ($D'$) and squared allele-frequency correlations ($r^2$) for pairs of SNP loci according to the TASSEL program. TASSEL was also used to identify SNP-trait associations by generating a general linear model (GLM). LD plots were generated in HaploView (Barrett, 2009), indicating $r^2$ values between pairs of SNPs multiplied by 100; white, $r^2 = 0$; shades of gray, $0 < r^2 < 1$; black, $r^2 = 1$.

**Analysis of Two-locus Interactions**

To identify epistatic interactions between significant loci, we calculated the pairwise epistatic interactions between the significant loci against the average accumulation of 13 metabolites, respectively, within the full population using two-way ANOVA (Zhou et al., 2012). The calculation was based on unweighed cell means, and the sums of squares were multiplied by the harmonic means of the cell sizes to form the test criteria. Those that showed significant interactions at $P \leq 0.001$ were subjected to permutation tests, in which the positions of the phenotype scores in the dataset were randomized to perform the two-way ANOVA again. This process was repeated 1,000 times. If no more than 1% of the random $F$ value was larger than the $F$ from the real data, it was regarded to be significant at $P \leq 0.001$.

**Phylogenetic Analysis**

The rice protein sequences in this study were extracted from the TIGR (http://rice.plantbiology.msu.edu/index.shtml). The sequences of all biochemically characterized acyltransferases and THT/TBT-like proteins from other plant species
were extracted from NCBI (http://www.ncbi.nlm.nih.gov/) and Phytozome 10.3 database (http://phytozome.jgi.doe.gov/pz/portal.html). The amino acid sequences were aligned using the CLUSTALW (version 1.83) program (Supplemental Data sets 8 and 9). The neighbor-joining tree was constructed using MEGA5 software with default parameters.

**Vector Construction and Rice Transformation**

An overexpression vector (pJC034) for rice was constructed from the Gateway overexpression vector pH2GW7, with the 35S promoter of pH2GW7 was replaced by the maize *ubiquitin* promoter. The overexpression constructs for *Os04g56900*, *Os09g37180*, *Os11g42290* and *Os11g42370* were generated by directionally inserting the full cDNAs from Nipponbare first into the entry vector pDONR207 and then into the destination vector pJC034 using the Gateway recombination reaction (Invitrogen). The constructs were introduced into *Agrobacterium* strain EHA105 and then transferred into *japonica* ZH11 as described previously. For each construct, at least three independent overexpression plants were selected for targeted metabolite analysis. The freeze-dried samples were extracted as previously described before analysis using an LC-ESI-MS/MS system (Chen et al., 2013).

**Quantitative RT-PCR Analysis**

We isolated total RNA from rice using an RNA extraction kit (TRIzol reagent; Invitrogen) according to the manufacturer’s instructions. Total RNA was treated with DNase I (RNase-free, Thermo scientific). The first-strand cDNA was synthesized using 3 µg RNA and 200 U M-MLV (Invitrogen) reverse transcriptases according to the manufacturer’s protocol. Real-time RT-PCR was performed on an optical 96-well plate in an AB StepOnePlus PCR system (Applied Biosystems) by using SYBR Premix reagent F-415 (Thermo scientific). *Actin1* was used as an endogenous control. The expression measurements were obtained using the relative quantification method (Livak and Schmittgen, 2001). Primers used in this study are listed in Supplemental
Recombinant Protein Expression and Enzyme Assay

The protein expression construct was generated by directionally inserting the full cDNA into the entry vector pDONR207 (Invitrogen) and then error-free clones were moved into the expression vector pDEST17 by LR recombination (Invitrogen). Recombinant proteins were expressed in BL21 (DE3) cells (Novagen) following induction by addition of 0.4 mM IPTG and growing continually for 12 h at 16 °C. Cells were harvested and cell pellets were resuspended in 50 mM sodium phosphate buffer (pH 7.8) and lysed by sonication. The crude extract was collected and clarified by centrifugation at 14,000 g for 40 minutes at 4 °C. The supernatant was incubated with 1 mL Glutathione Sepharose (GE Healthcare Bio-Sciences) that been equilibrated with 50 mM sodium phosphate buffer (pH 7.8) at 4°C for 3 h. The column was washed with 20 mL equilibration buffer, followed by washing with 5 mL of buffer containing 15 mM glutathione. SDS-PAGE was performed and the proteins in the gels were visualized using Coomassie Brilliant Blue R-250. The enzyme-containing fractions were collected and concentrated using BioPhotometer (Eppendorf).

The enzyme reactions in vitro assay for transferases were performed in a total volume of 100 μL containing 200 μM acyl donor, 1 mM serotonin (tryptamine) and 50 mM potassium phosphate buffer (pH 7.4) and enzyme. Prior to addition of enzyme, the mixture was preincubated at 37°C for 10 minutes, and the reaction was started by addition of the enzyme. After incubating at 37°C for 30 min, the reaction was stopped by adding 200 μL of ice-cold 0.5% trifluoroacetic acid. The reaction mixture was then filtered through a 0.2 μm filter (Millipore) before being used for LC/MS analysis.

HPLC conditions for the analysis of polyamines and phenolamides were as follows: column, shim-pack VP-ODS (150L× 4.6); flow rate, 0.8mL/min; solvent A, 0.04% (by volume) acetic acid in water; solvent B, 0.04% acetic acid in acetonitrile. After injection (40 μL) into a column that had been equilibrated with 5% solvent B (by
volume), the column was initially developed isocratically with 5% solvent B, followed by a linear gradient from 5% to 95% solvent B for 20 min. The column was then washed isocratically with 95% solvent B for 2 min, followed by a linear gradient from 95% to 5% solvent B for 0.1 min. The column was isocratically with 5% solvent B for 5 min. The chromatograms were obtained with detection at 280nm. Peak identification of each component was confirmed post-run by LC/MS/MS analysis. The amounts of phenolamides were determined from peak integration using authentic samples for calibration.

**Enzyme Kinetics**

To determine the kinetic constants of Os-THT1/2 and Os-TBT1/2 for acyl donors, their activities were determined using 0 to 500 μM different CoA esters at a fixed concentration of 800 μM tryptamine. The kinetic constants for acyl acceptor were assayed using 0 to 800 μM different polyamine (except 0 to 3mM agmatine) at a fixed concentration of 800 μM p-coumaroyl-CoA or benzoyl-CoA. Hydroxycinnamoyl CoAs were purchased from TransMIT GmbH (Germany) and all other compounds were purchased from Sigma-Aldrich. All Kinetic parameters were calculated using Michaelis-Menten model (Sigma Plot, version 12.5). All the reactions were run in duplicate, and each experiment was repeated twice.

**Characterization of Conserved Protein Motifs**

Protein sequences were collected in a fasta file and subjected to motif analysis using the MEME (http://meme-suite.org/tools/meme) web interface with default settings. Previous work suggested that a motif width of 5 is typical; the actual settings in our study ranged from 4 to 8. To find clade IV-specific motifs, sequences containing 26 proteins from 5 clades and 8 proteins from clade IV were analyzed separately. Well-conserved sequences with low E-value, which were only available in clade IV, were considered to be specific motifs. Similar methods were utilized to find clade IVa and IVb specific motifs.
**Quantification of Phenolamides in Rice**

Soluble phenolamides were extracted, separated and quantified by LC/MS/MS as previously described (Chen et al., 2013). An authentic standard of \(p\)-coumaric acid was purchased from Sigma-Aldrich, USA. The standard stock solution was prepared using methanol as a solvent and stored at -20°C. A standard solution of \(p\)-coumaric acid was prepared just before use by diluting the stock solution with aqueous methanol. Phenolamides were quantified based on comparison with standards of \(p\)-coumaric acid.

**Accession Numbers**

Sequence data from this article can be found in the GenBank/EMBL databases under the following accession numbers:

- Os-\(AHT1\), KX430015; Os-\(PHT1\), KX430016; Os-\(PHT2\), KX430017; Os-\(PHT3\), KX430018; Os-\(THT1\), KX430019; Os-\(THT2\), KX430020; Os-\(TBT1\), KX430021; Os-\(TBT2\), KX430022;
- Os04g56910, XP_015634184.1; Os09g37200, XP_015651357.1; Os06g08640, XP_015641749.1; Hv-ACT, AAO73071; Na-AT1, AET80688.1; Bradi1g72230, KQK23250.1; Bradi3g23280, XP_003573858.1;
- GRMZM2G114918, XP_008651745.1; GRMZM2G030436, XP_008660165.1; *H. vulgare* BAJ88765, BAJ88765.1; Seita.9G249300.1, XP_004983083.1.

**Supplemental Data**

- **Supplemental Figure 1.** Natural Variation of Phenolamides within Rice Subspecies in Grain.
- **Supplemental Figure 2.** Manhattan Plots of 8 PAs Detected in Rice Leaf.
- **Supplemental Figure 3.** Manhattan Plots of 5 PAs Detected in Rice Grain.
- **Supplemental Figure 4.** QTL Mapping Results for 5 Phenolamides in Rice Leaf and Grain.
- **Supplemental Figure 5.** Locus Interactions among Significant Loci Controlling PA Accumulation.
- **Supplemental Figure 6.** In Vitro Enzymatic Assays for Recombinant Proteins.
- **Supplemental Figure 7.** Clade-Specific Motifs Identified by MEME.
- **Supplemental Table 1.** Ranges, Variance, and Broad-sense Heritability of 11 Phenolamides in Rice.
Supplemental Table 2. Summary of Significant Locus-Metabolite Associations Identified by GWAS.

Supplemental Table 3. Results of Analysis of Two-locus Interaction.

Supplemental Table 4. Summary of BAHD Transferase Candidate Genes from Phenolamide Mapping in Distinct Rice Tissues.

Supplemental Table 5. Relative Activity (RA) of Recombinant BAHD Acyltransferases on Different Acyl Donors and Acceptors.

Supplemental Table 6. Primers Used in This Study.

Supplemental Data set 1. Lead SNPs and Loci Detected in Leaf and Grain.

Supplemental Data set 2. Metabolic Quantitative Trait Loci (mQTLs) Results for PAs in Leaf and Grain.

Supplemental Data set 3. Associated Information of SNPs at Candidate Genes.

Supplemental Data set 4. Comparison of Coding Region within Two Candidates between ZS97 and MH63.

Supplemental Data set 5. Gene Expression of BAHD Transferase Candidates in Leaf and Grain of Two Rice Accessions.

Supplemental Data set 6. Collected 480 Rice Varieties.


Supplemental Data set 8. Text File of the Alignment Corresponding to the Phylogenetic Analysis of Acyltransferases in Figure 4A.

Supplemental Data set 9. Text File of the Alignment Corresponding to the Phylogenetic Analysis of THT/TBT-like Proteins in Figure 10B.

ACKNOWLEDGMENTS

We thank Cathie Martin for comments on the manuscript. This work was supported by the State Key Program of National Natural Science Foundation of China (No. 31530052), the Major State Basic Research Development Program of China (973 Program, No. 2013CB127001), the National High Technology R&D Program of China (863 Program, No. 2012AA10A304), and the Program from Fundamental Research Funds for the Central Universities (No. 2662015PY196).

AUTHOR CONTRIBUTIONS

J.L and M.P designed the research. Y.G and M.P performed most of the experiments. W.W and J.W cooperated the bioinformatics work. W.C identified the phenolamides in rice, and W.C, Y.G and M.P identified BAHD family genes. Y.G performed the
transgenic work. M.P conducted the enzyme activity assay. J.L and M.P analyzed the data. J.L and M.P wrote the article.

REFERENCE


28


Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., Altmann, T., Stitt, M., Willmitzer,


**FIGURE LEGENDS**

**Figure 1.** Structures of CoA-thioesters and Amine Moieties for Biosynthesis of Phenolamides. Structures of benzoyl-CoA and hydroxycinnamoyl-CoAs are shown as R1 - R5. Putrescine and agmatine are aliphatic amines. Tyramine, tryptamine and serotonin are aromatic amines. R indicates the differential replacement by R1 - R5. Eleven rice phenolamides in this study refer to different combinations of CoA-thioesters and amine moieties.

**Figure 2.** Natural Variation of Phenolamides within Rice Subspecies in Leaf. (A) Heat map visualization of the relative difference of phenolamides in 480 rice varieties. The content value of each phenolamide was normalized to complete linkage hierarchical clustering. Each rice variety is visualized in a single column and each of phenolamides is represented by a single row. Yellow indicates high abundance, whereas low relative phenolamides are blue (color key scale above heat map).
(B) Box plot for the content of eight phenolamides in *japonica* (yellow) and *indica* (blue) subspecies. DW, dry weight. The full names for abbreviations of the metabolites are given in Table 1.

**Figure 3.** Integrated Mapping Results of Association and Linkage Analysis for 11 Rice Phenolamides.

Manhattan plots of 11 rice phenolamide traits across twelve rice chromosomes are integrated and shown upward. The strength of association is indicated as the negative logarithm of the *P* value for the linear mixed model. All metabolite-SNP associations with *P* values below 1.23×10⁻⁶ (horizontal dashed lines in all Manhattan plots) are plotted against genome location in intervals of 1 Mb. Triangle in light grey represents metabolite-SNP associations with *P* values below 1.0×10⁻³⁰. Integrated linkage mapping of 5 phenolamides in rice are shown downward using a ZS97 × MH63 RIL population. Horizontal dashed line indicates significance thresholds (LOD = 6) in this study. Triangle in black indicates value of LOD below 30. Loci within a physical distance of 1 Mb were assigned to identify genes in this study. L, leaf; G, grain. AHT, agmatine: hydroxycinnamoyl-CoA transferase; PHT, putrescine: hydroxycinnamoyl-CoA transferase; THT, tryptamine: hydroxycinnamoyl-CoA transferase; TBT, tryptamine: benzoyl-CoA transferase.

**Figure 4.** Phylogenetics of Plant BAHD Family and GNAT Family Acyltransferases.

(A) An unrooted phylogenetic tree was constructed as described in Methods. Twenty-seven biochemically characterized BAHD proteins and 5 GNAT proteins were collected. Bootstrap values >70% (based on 1,000 replications) are indicated at each node (Bar: 0.1 amino acid substitutions per site). Os, *Oryza sativa*; Hv, *Hordeum vulgare*; Na, *Nicotiana tabacum*; At, *Arabidopsis thaliana*; Gt, Gentiana triflora; Dm, *Dendranthema x morifolium*; Dv, *Dahlia variabilis*; Ch, *Clarkia breweri*; Tc, *Taxus cuspidata*; De, Dianthus caryophyllus; As, *Avena sativa*; Na, *Nicotia attenuata*; Cr, Catharanthus roseus; Rs, Rauvolfia serpentine; Ca, *Capsicum annuum*; St, *Solanum tuberosum*; Sl, *Solanum lycopersicum*.

(B) Graphical map of candidate BAHD *N*-transferases (showed in dark grey) in the rice genome. Gene numbers and symbols are shown below and above each gene, respectively.

**Figure 5.** Biochemical Properties of the Recombinant Os-THTs and Os-TBTs *in vitro*.

(A) LC-MS chromatograms of *in vitro* enzyme assays showing the enzyme activity of recombinant Os-THT1. Trm, tryptamine; Sen, serotonin; Tym, tyramine; Agm, agmatine.

(B) Substrate preference of Os-THT1/2 and Os-TBT1/2 with different amines as acyl acceptors. Asterisk indicates the relative activity value is less than 1%. ND, no determined.

(C) Substrate preference of Os-THT1/2 and Os-TBT1/2 with different CoA esters as acyl donors. Ben-CoA, benzoyl-CoA; Cou-CoA, *p*-coumaroyl-CoA; Caf-CoA,
caffeoyl-CoA; Fer-CoA, feruloyl-CoA. Asterisk indicates the relative activity value is less than 1%.

Figure 6. Analysis of Functional SNPs within Candidate Genes Underlying the Natural Variation of Tryptamine Conjugates.
A representation of the pairwise $r^2$ value (a measure of linkage disequilibrium) among polymorphic sites in Os-TBT1 (A) and Os-TBT2 (B), where the darkness of the box corresponds to the $r^2$ value according to the legend. Asterisk represents the proposed functional sites. Corresponding SNPs are listed in Supplemental Data set 3. DNA sequence polymorphisms of Os-TBT1 (C) and Os-TBT2 (D) coding sequences in ZS97 relative to MH63. ZS97, Zhenshan 97, colored in black. MH63, Minghui 63, colored in light grey. Indel variations are underlined. Corresponding SNPs and Indels are listed in Supplemental Data set 4.

Figure 7. Metabolite Analysis of Transgenic Individuals in Rice.
The amounts of corresponding phenolamides in Os-PHT3 (A), Os-AHT1 (B), Os-TBT1 (C) and Os-TBT2 (D) overexpressing (OE) lines relative to the amounts in the wild-type (WT) lines are plotted. Error bars are SE of mean for n = 3. DW, dry weight.

Figure 8. Schematic Summary of the Activity of Eight BAHD Transferases in the Phenolamide Biosynthetic Pathway.
The full names for abbreviations of the metabolites are shown in Table 1. PAL, Phenylalanine ammonia-lyase; 4C4H, cinnamate-4-hydroxylase; C3H, coumarate-3-hydroxylase; COMT, caffeate $O$-methyltransferase; 4CL, 4-coumarate: CoA ligase; CNL, Cinnamate: CoA ligase.

Figure 9. Comparison of Amino Acid Sequences and Specific Activities of Mutant Proteins.
(A) Alignment of amino acid sequences in the N terminal regions of the N-hydroxycinnamoyltransferases. Identical amino acids are denoted as asterisk, and the gaps are indicated by dash line.
(B) Constructs of mutant proteins and measurements of their specific activities. Os-THT1::RRRR indicates insertion of four tandem arginine residues in the wild-type Os-THT1. Os-TBT1ΔRRRR indicates deletion of the four tandem arginine residues in the wild-type Os-TBT1. Dash indicates the $K_m$ value was too large to be determined precisely. ND, not determined. WT, wild type.

Figure 10. Biochemical Assays and Phylogenetic Analysis of Os-THT/TBT-like Proteins.
(A) In vitro biochemical assays of six proteins belonging to the Os-THT/OsTBT-like proteins.
(B) An unrooted tree of Os-THT/TBTs and homolog proteins from other species (> 40% identity). Bootstrap values (based on 1000 replications) >70% are shown for
### Table 1. Phenolamides Identified in Rice

<table>
<thead>
<tr>
<th>Trait</th>
<th>Abbreviations</th>
<th>RT (min)</th>
<th>m/z</th>
<th>Main fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-Caffeoyl agmatinea</td>
<td>Caf-Agm</td>
<td>3.37</td>
<td>293.1617</td>
<td>163.2, 233.1, 135.1, 89.1</td>
</tr>
<tr>
<td>N-Coumaroyl agmatinea</td>
<td>Cou-Agm</td>
<td>5.05</td>
<td>277.166</td>
<td>147.2, 119.1, 91.1</td>
</tr>
<tr>
<td>N-Feruloyl agmatinea</td>
<td>Fer-Agm</td>
<td>5.32</td>
<td>307.1758</td>
<td>177.2, 291.1, 145.2, 117.3, 89.1</td>
</tr>
<tr>
<td>N-Caffeoyl putrescinea</td>
<td>Caf-Put</td>
<td>2.73</td>
<td>251.1394</td>
<td>163.2, 135.1, 117.3, 89.1</td>
</tr>
<tr>
<td>N-Coumaroyl putrescineab</td>
<td>Cou-Put</td>
<td>3.36</td>
<td>235.1441</td>
<td>147.2, 119.1, 91.1</td>
</tr>
<tr>
<td>N-Feruloyl putrescineab</td>
<td>Fer-Put</td>
<td>3.74</td>
<td>265.1551</td>
<td>177.2, 145.2, 117.2, 89.1</td>
</tr>
<tr>
<td>N-Benzyol tryptaminea</td>
<td>Ben-Trm</td>
<td>13.50</td>
<td>265.1332</td>
<td>144.0, 204.9, 104.9, 77.1</td>
</tr>
<tr>
<td>N-Cinnamoyl tryptaminea</td>
<td>Cin-Trm</td>
<td>14.53</td>
<td>291.1494</td>
<td>143.9, 103.0, 157.0, 273.1</td>
</tr>
<tr>
<td>N-Coumaroyl tryptamineb</td>
<td>Cou-Trm</td>
<td>12.30</td>
<td>317.1386</td>
<td>147.1, 91.1, 144.2, 119.0</td>
</tr>
<tr>
<td>N-Coumaroyl serotoninb</td>
<td>Cou-Sen</td>
<td>9.68</td>
<td>323.1386</td>
<td>147.1, 119.0, 160.0, 91.3</td>
</tr>
<tr>
<td>N-Benzoyl serotoninb</td>
<td>Ben-Sen</td>
<td>10.90</td>
<td>281.3082</td>
<td>160.1, 264.0, 115.1, 105.0, 77.1</td>
</tr>
</tbody>
</table>

aDetected in rice leaf, bdetected in rice grain.

### Table 2. Kinetic Parameters of Os-THT1 and Os-THT2

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Os-THT1</th>
<th>Os-THT2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_m$ (μM)</td>
<td>$k_{cat}$ (s$^{-1}$)</td>
</tr>
<tr>
<td>Donors$^a$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzyoyl-CoA</td>
<td>369.6 ± 85.3</td>
<td>1.05 ± 0.07</td>
</tr>
<tr>
<td>$p$-Coumaroyl-CoA</td>
<td>91.4 ± 30.9</td>
<td>4.09 ± 0.44</td>
</tr>
<tr>
<td>Caffeoyl-CoA</td>
<td>138.9 ± 39.3</td>
<td>0.28 ± 0.01</td>
</tr>
<tr>
<td>Feruloyl-CoA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Acceptors$^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tryptamine</td>
<td>82.0 ± 25.7</td>
<td>6.15 ± 0.57</td>
</tr>
<tr>
<td>Serotonin</td>
<td>112.9 ± 9.1</td>
<td>6.81 ± 0.18</td>
</tr>
<tr>
<td>Tyramine</td>
<td>178.0 ± 53.4</td>
<td>5.84 ± 0.53</td>
</tr>
<tr>
<td>Agmatine</td>
<td>1439 ± 382</td>
<td>1.69 ± 0.23</td>
</tr>
</tbody>
</table>

$^a$400 μM tryptamine was used as the acyl acceptor.

$^b$500 μM $p$-Coumaroyl-CoA was used as the acyl donor.

NA, not available, $K_m$ value was too large to be determined precisely.

ND, not determined due to the low activity.
### Table 3. Kinetic Parameters of Os-TBT1 and Os-TBT2

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Os-TBT1</th>
<th></th>
<th>Os-TBT2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_m$ ($\mu$M)</td>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>$k_{cat}/K_m$</td>
<td>$K_m$ ($\mu$M)</td>
</tr>
<tr>
<td>Donors$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzyol-CoA</td>
<td>474.7 ± 54.7</td>
<td>2.68 ± 0.90</td>
<td>0.006</td>
<td>81.3 ± 3.7</td>
</tr>
<tr>
<td>p-Coumaroyl-CoA</td>
<td>101.9 ± 29.1</td>
<td>4.74 ± 0.48</td>
<td>0.047</td>
<td>ND</td>
</tr>
<tr>
<td>Caffeoyl-CoA</td>
<td>979 ± 126</td>
<td>0.52 ± 0.02</td>
<td>0.001</td>
<td>ND</td>
</tr>
<tr>
<td>Feruloyl-CoA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>ND</td>
</tr>
<tr>
<td>Acceptors$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tryptamine</td>
<td>10.5 ± 4.7</td>
<td>2.21 ± 0.17</td>
<td>0.210</td>
<td>45.4 ± 6.8</td>
</tr>
<tr>
<td>Serotonin</td>
<td>19.6 ± 6.6</td>
<td>2.50 ± 0.18</td>
<td>0.127</td>
<td>37.5 ± 6.1</td>
</tr>
<tr>
<td>Tyramine</td>
<td>15.0 ± 4.2</td>
<td>2.39 ± 0.13</td>
<td>0.159</td>
<td>114.3 ± 11.2</td>
</tr>
</tbody>
</table>

$^a$200 $\mu$M tryptamine was used as the acyl acceptor.

$^b$500 $\mu$M p-Coumaroyl-CoA and benzoyl-CoA were used as the acyl donor, respectively.

NA, not available, $K_m$ value was too large to be determined precisely.

ND, not determined due to the low activity.
Figure 1. Structures of CoA-thioesters and Amine Moieties for Biosynthesis of Phenolamides. Structures of benzoyl-CoA and hydroxycinnamoyl-CoAs are shown as R1 - R5. Putrescine and agmatine are aliphatic amines. Tyramine, tryptamine and serotonin are aromatic amines. R indicates the differential replacement by R1 - R5. Eleven rice phenolamides in this study refer to different combinations of CoA-thioesters and amines moieties.
Figure 2. Natural Variation of Phenolamides within Rice Subspecies in Leaf.

(A) Heat map visualization of the relative difference of phenolamides in 480 rice varieties. The content value of each phenolamide was normalized to complete linkage hierarchical clustering. Each rice variety is visualized in a single column and each of phenolamides is represented by a single row. Yellow indicates high abundance, whereas low relative phenolamides are blue (color key scale above heat map).

(B) Box plots for the contents of eight phenolamides in *japonica* (yellow) and *indica* (blue) subspecies. DW, dry weight. The full names for abbreviations of the metabolites are given in Table 1.
Manhattan plots of 11 phenolamides traits across the twelve rice chromosomes are integrated and shown upward. The strength of association is indicated as the negative logarithm of the P value for the linear mixed model. All metabolite-SNP associations with P values below $1.23 \times 10^{-6}$ (horizontal dashed lines in all Manhattan plots) are plotted against genome location in intervals of 1 Mb. Triangle in light grey represents metabolite-SNP associations with P values below $1.0 \times 10^{-30}$. Integrated linkage mapping of 5 phenolamides in rice is shown downward using a ZS97 × MH63 RIL population. Horizontal dashed line indicates significance thresholds (LOD = 6) in this study. Triangle in black indicates value of LOD below 30. Loci within a physical distance of 1 Mb were assigned to identified genes in this study. L, leaf; G, grain. AHT, agmatine: hydroxycinnamoyl-CoA transferase; PHT, putrescine: hydroxycinnamoyl-CoA transferase; THT, tryptamine: hydroxycinnamoyl-CoA transferase; TBT, tryptamine: benzoyl-CoA transferase.

Figure 3. Integrated Mapping Results of Association and Linkage Analysis for 11 Rice Phenolamides.
Figure 4. Phylogenetics of Plant BAHD Family and GNAT Family A cyltransferases.

(A) An unrooted phylogenetic tree was constructed as described in Methods. Twenty-seven biochemically characterized BAHD proteins and 5 GNAT proteins were collected. Bootstrap values >70% (based on 1,000 replications) are indicated at each node (Bar: 0.1 amino acid substitutions per site). Os, Oryza sativa; Hv, Hordeum vulgare; Na, Nicotiana tabacum; At, Arabidopsis thaliana; Gt, Gentian a triflora; Dm, Dendranthema x morifolium; Dv, Dahlia variabilis; Ph, Petunia hybrida; Cb, Clarkia breweri; Tc, Taxus cuspidata; Dc, Dianthus caryophyllus; As, Avena sativa; Na, Nicotiana attenuata; Cr, Catharanthus roseus; Rs, Rauvolfia serpentine; Ca, Capsicum annuum; St, Solanum tuberosum; Sl, Solanum lycopersicum.

(B) Graphical map of candidate BAHD N-transferases (showed in dark grey) in the rice genome. Gene numbers and symbols are shown below and above each gene, respectively.
Figure 5. Biochemical Properties of the Recombinant Os-THTs and Os-TBTs in vitro.

(A) LC-MS chromatograms of in vitro enzyme assays showing the enzyme activity of recombinant Os-THT1. Trm, tryptamine; Sen, serotonin; Tym, tyramine; Agm, agmatine.

(B) Substrate preference of Os-THT1/2 and Os-TBT1/2 with different amines as acyl acceptors. Asterisk indicates the relative activity value is less than 1%. ND, no determined.

(C) Substrate preference of Os-THT1/2 and Os-TBT1/2 with different CoA esters as acyl donors. Ben-CoA, benzoyl-CoA; Cou-CoA, p-coumaroyl-CoA; Caf-CoA, caffeoyl-CoA; Fer-CoA, feruloyl-CoA. Asterisk indicates the relative activity value is less than 1%.
Figure 6. Analysis of Functional SNPs within Candidate Genes Underlying the Natural Variation of Tryptamine Conjugates.

A representation of the pairwise $r^2$ value (a measure of linkage disequilibrium) among polymorphic sites in Os-TBT1 (A) and Os-TBT2 (B), where the darkness of the box corresponds to the $r^2$ value according to the legend. Asterisk represents the proposed functional sites. Corresponding SNPs are listed in Supplemental Data set 3. DNA sequence polymorphisms of Os-TBT1 (C) and Os-TBT2 (D) coding sequences in ZS97 relative to MH63. ZS97, Zhenshan 97, colored in black. MH63, Minghui 63, colored in light grey. Indel variations are underlined. Corresponding SNPs and Indels are listed in Supplemental Data set 4.
Figure 7. Metabolite Analysis of Transgenic Individuals in Rice
The amounts of corresponding phenolamides in Os-PHT3 (A), Os-AHT1 (B), Os-TBT1 (C) and Os-TBT2 (D) overexpressing (OE) lines relative to the amounts in the wild type (WT) lines are plotted. Error bars are SE of mean for n = 3. DW, dry weight.
Figure 8. Schematic Summary of the Activity of Eight BAHD Transferases in the Phenolamide Biosynthetic Pathway.

The full names for abbreviations of the metabolites are shown in Table 1. PAL, Phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylase; C3H, coumarate-3-hydroxylase; COMT, caffeate O-methyltransferase; 4CL, 4-coumarate:CoA ligase; CNL, Cinnamate:CoA ligase.
Figure 9. Comparison of Amino Acid Sequences and Specific Activities of mutant proteins.

(A) Alignment of amino acid sequences in the N-terminal regions of the N-hydroxycinnamoyltransferases. Identical amino acids are denoted as asterisk, and the gaps are indicated by dash line.

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HvACT</td>
<td>80</td>
<td>DASGGR</td>
<td></td>
<td>AI</td>
<td>L</td>
<td>LNDAGA</td>
<td>RF</td>
<td>VEA</td>
<td>99</td>
</tr>
<tr>
<td>04g56910</td>
<td>79</td>
<td>DGDGDR</td>
<td></td>
<td>AI</td>
<td>L</td>
<td>LNDGA</td>
<td>RF</td>
<td>VEA</td>
<td>98</td>
</tr>
<tr>
<td>OsAHT1</td>
<td>79</td>
<td>DGDGNR</td>
<td></td>
<td>AI</td>
<td>L</td>
<td>LNDGA</td>
<td>RF</td>
<td>VEA</td>
<td>98</td>
</tr>
<tr>
<td>OsTHT1</td>
<td>81</td>
<td>DERGRK</td>
<td></td>
<td>CFRLN</td>
<td>N</td>
<td>NAGA</td>
<td>RV</td>
<td>LEA</td>
<td>100</td>
</tr>
<tr>
<td>OsTBT1</td>
<td>75</td>
<td>DQGRRR</td>
<td></td>
<td>RRRR</td>
<td></td>
<td>VLHNNE</td>
<td>G</td>
<td>VLVDA</td>
<td>98</td>
</tr>
<tr>
<td>OsTBT2</td>
<td>75</td>
<td>DQGRRR</td>
<td></td>
<td>RRRR</td>
<td></td>
<td>VLHNDE</td>
<td>G</td>
<td>ALVLD</td>
<td></td>
</tr>
</tbody>
</table>

(B) Constructs of mutant proteins and measurements of their specific activities. Os-THT1::RRRR indicates insertion of four tandem arginine residues in the wild-type Os-THT1. Os-TBT1ΔRRRR indicates deletion of four tandem arginine residues in the wild-type Os-TBT1. Dash indicates the $K_m$ value was too large to be determined precisely. ND, not determined. WT, wild type.

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os-THT1::RRRR</td>
<td></td>
<td>serotonin</td>
<td>agmatine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os-THT1-WT</td>
<td></td>
<td>79.3±5.2</td>
<td></td>
<td>1439 ± 382</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os-TBT1ΔRRRR</td>
<td></td>
<td>70.3±2.4</td>
<td></td>
<td>1530 ± 225</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os-TBT1-WT</td>
<td></td>
<td>19.6±6.6</td>
<td></td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 10. Biochemical Assays and Phylogenetic Analysis of Os-THT/TBT-like proteins

(A) In vitro biochemical assays of six proteins belonging to the Os-THT/TBT-like proteins.

(B) An unrooted tree of Os-THT/TBTs and homolog proteins from other species (> 40% identity). Bootstrap values (based on 1000 replications) >70% are shown for corresponding nodes. The scale measures evolutionary distance in substitutions per site. The numbers of homolog proteins in *Aquilegia coerulea*, *Nelumbo nucifera* and *Amborella trichopoda* are shown. Brast, *Brachypodium stacei*; Bradi, *Brachypodium distachyon*; GRMZM, *Zea mays*; Pahal, *Panicum hallii*; Pavir, *Panicum virgatum*; Sevir, *Setaria viridis*; Seita, *Setaria italic*; Ccy, *Cynara cardunculus*; Pt, *Populus trichocarpa*; Cco, *Coffea canephora*. At, *Arabidopsis thaliana*; Nt, *Nicotiana tabacum*. Dots indicate the six proteins shown in the Figure 10A. Triangles indicate the four characterized enzymes in this study.


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title


Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title
Evolutionarily Distinct BAHD N-acyltransferases are Responsible for Natural Variation of Aromatic Amine Conjugates in Rice
Meng Peng, Yanqiang Gao, Wei Chen, Wensheng Wang, Shuangqian Shen, Jian Shi, Cheng Wang, Yu Zhang, Li Zou, Shouchuang Wang, Wan Jian, Xianqing Liu, Liang Gong and Jie Luo

*Plant Cell*; originally published online June 27, 2016; DOI 10.1105/tpc.16.00265

This information is current as of October 20, 2017

<table>
<thead>
<tr>
<th>Supplemental Data</th>
<th>/content/suppl/2016/06/27/tpc.16.00265.DC1.html</th>
</tr>
</thead>
<tbody>
<tr>
<td>eTOCs</td>
<td>Sign up for eTOCs at:</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>CiteTrack Alerts</td>
<td>Sign up for CiteTrack Alerts at:</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>Subscription Information</td>
<td>Subscription Information for <em>The Plant Cell</em> and <em>Plant Physiology</em> is available at:</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a></td>
</tr>
</tbody>
</table>