Molecular Characterization of Soybean Pterocarpan 2-Dimethylallyltransferase in Glyceollin Biosynthesis: Local Gene and Whole-Genome Duplications of Prenyltransferase Genes Led to the Structural Diversity of Soybean Prenylated Isoflavonoids

Keisuke Yoneyama, Tomoyoshi Akashi and Toshio Aoki*

Department of Applied Biological Sciences, Nihon University, Fujisawa, Kanagawa, 252-0880 Japan
*Corresponding author: E-mail, aoki.toshio@nihon-u.ac.jp; Fax, +81-466-84-3353.
(Received August 2, 2016; Accepted October 15, 2016)

Soybean (Glycine max) accumulates several prenylated isoﬂavonoid phytoalexins, collectively referred to as glyceollins. Glyceollins (I, II, III, IV and V) possess modiﬁed pterocarpan skeletons with C5 moieties from dimethylallyl diphosphate, and they are commonly produced from (6aS,11aS)-3,9,6a-trihydroxypterocarpan [(−)-glycinol]. The metabolic fate of (−)-glycinol is determined by the enzymatic introduction of a dimethylallyl group into C-4 or C-2, which is reportedly catalyzed by regiospeciﬁc prenyltransferases (PTs). 4-Dimethylallyl (−)-glycinol and 2-dimethylallyl (−)-glycinol are precursors of glyceollin I and other glyceollins, respectively. Although multiple genes encoding (−)-glycinol biosynthetic enzymes have been identiﬁed, those involved in the later steps of glyceollin formation mostly remain unidentiﬁed, except for (−)-glycinol 4-dimethylallyltransferase (G4DT), which is involved in glyceollin I biosynthesis. In this study, we identiﬁed four genes that encode isoﬂavonoid PTs, including (−)-glycinol 2-dimethylallyltransferase (G2DT), using homology-based in silico screening and biochemical characterization in yeast expression systems. Transcript analyses illustrated that changes in G2DT gene expression were correlated with the induction of glyceollins II, III, IV and V in elicitor-treated soybean cells and leaves, suggesting its involvement in glyceollin biosynthesis. Moreover, the genomic signatures of these PT genes revealed that G4DT and G2DT are paralogs derived from whole-genome duplications of the soybean genome, whereas other PT genes [isoﬂavone dimethylallyltransferase 1 (IDT1) and IDT2] were derived via local gene duplication on soybean chromosome 11.

Keywords: Gene duplication • Isoﬂavonoid • Phytoalexin • Prenyltransferase • Soybean.

Abbreviations: C4DT, coumestrol 4-dimethylallyltransferase; CDS, coding sequence; DMAPP, dimethylallyl diphosphate; ESI, electrospray ionization; EST, expressed sequence tag; G2DT, (−)-glycinol 2-dimethylallyltransferase; G4DT, (−)-glycinol 4-dimethylallyltransferase; HGT, homogentisate geranylgeranyltransferase; HID, 2-hydroxyisoflavanone dehydratase; HPT, homogentisate phytlytransferase; IDT, isoflavone dimethylallyltransferase; IFS, 2-hydroxyisoflavanone synthase; LGD, local gene duplication; m/z, mass to-charge ratio; MS/MS, tandem mass spectrometry; NCBI, the National Center for Biotechnology Information; NMR, nuclear magnetic resonance; P450, cytochrome P450 monooxygenase; P6aH, pterocarpan 6a-hydroxylase; PT, prenyltransferase; TLC, thin-layer chromatography; TQMS, triple quadrupole mass spectrometry; TMHMM, transmembrane hidden Markov model; UPLC, ultra-performance liquid chromatography; WGD, whole-genome duplication.

All nucleotide sequences reported in this paper have been submitted to the DDBJ/EMBL/GenBank database under accession numbers LC140926 (PT3), LC140927 (C4DT), LC140928 (IDT1), LC140929 (IDT2) and LC140930 (G2DT).

Introduction

Flavonoids comprise a widespread class of plant metabolites, and those with one or more dimethylallyl (C5) or geranyl (C10) groups present on their aromatic rings are known as prenylated flavonoids. The structures of prenylated flavonoids vary regarding the numbers and lengths of prenyl groups and modiﬁcations such as cyclization and hydroxylation. Hence, prenylated flavonoids are highly diverse, and they include approximately 1,000 compounds that are increasingly examined for their antioxidant, antitumor, antibacterial, antiviral and estrogenic activities (Yazaki et al. 2009). Although some of the properties are common to flavonoids with no prenyl groups, it is generally accepted that prenyl moieties increase lipophilicity and membrane permeability, and tend to potentiate the bioactivities of flavonoids (Botta et al. 2005, Yazaki et al. 2009, Yang et al. 2015). Prenylation of flavonoids is catalyzed by prenyltransferases (PTs), which transfer prenyl moieties from allylic prenyl diphosphate to flavonoid skeletons. The flavonoid-speciﬁc PT narin- genin 8-dimethylallyltransferase (SNBDT) was initially identiﬁed in the leguminous plant Sophora ﬂavescent (Sasaki et al. 2008), and several plant PTs that act on flavonoids and

PCP PLANT & CELL PHYSIOLOGY

Regular Paper

Plant Cell Physiol. 57(12): 2497–2509 (2016) doi:10.1093/pcp/pcw178, available online at www.pcp.oxfordjournals.org
© The Author 2016. Published by Oxford University Press on behalf of Japanese Society of Plant Physiologists.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
other aromatic compounds have been molecularly and biochemically characterized (Akashi et al. 2009, Sasaki et al. 2011, Shen et al. 2012, Chen et al. 2013, Karamat et al. 2014, Li et al. 2014, Munakata et al. 2014, Wang et al. 2014, Li et al. 2015, Munakata et al. 2016). A recent report on hop PTs demonstrated that a complex of two PTs, as a metabolon, catalyzes three steps of sequential aromatic prenylation in \(-b\)-bitter acid biosynthesis (Li et al. 2015). However, there is a wide gap between the number of identified PTs and that of known prenylated flavonoids and aromatic compounds. The fact that these characterized PTs have stringent substrate specificities indicates the presence of additional unidentified PTs.

Isoflavonoids comprise a subclass of flavonoids, in which phenyl rings (B-rings) are attached to C-3 instead of C-2 in the flavonoid skeleton. Isoflavonoids are further classified as isoflavones, pterocarpans, coumestans or isoflavans, all of which are predominantly unique to leguminous plants (Aoki et al. 2000). More than half of these isoflavonoids are prenylated, and many prenylated isoflavonoids are considered to be inducible antimicrobial phytoalexins. Glyceollins, a series of prenylated pterocarpans, are well known to be the phytoalexins of soybean (Glycine max) (Tahara and Ibrahim 1995). During glyceollin biosynthesis, the dimethylallyl moiety is introduced at either C-4 or C-2 of the pterocarpan precursor \((-)-\)glycinol [(\(6aS, 11aS\))-3,9,6a-trihydroxypterocarpan; Fig. 1]. Although most glycinol biosynthetic enzymes and their genes have been characterized, little is known about the enzymes responsible for the succeeding prenylation and cyclization reactions, which lead to the formation of the characteristic structure of glyceollins. A homology-based approach has led to the isolation of a cDNA encoding dimethylallyl diphosphate (DMAPP)-(\(-)-\)glycinol 4-dimethylallyltransferase (G4DT), which yields 4-dimethylallylglycinol, the direct precursor of glyceollin I (Akashi et al. 2009), but not 2-dimethylallylglycinol, suggesting
stringent regiospecificity of G4DT for the C-4 position of glycinol. These data suggest that another PT is involved in the biosynthesis of glyceollins II–V (Fig. 1), which should be referred to as DMAPP:((–)-glycinol 2-dimethylallyltransferase (G2DT). Recent reports also illustrated that elicitor-treated soybeans accumulate glycinols and several prenylated isoflavonoids, including 4-dimethylallylcoumestrol, 3’-dimethylallyldaidzein, 3’-dimethylallylgenistein and 8-dimethylallyldaidzein (Cheng et al. 2011, Simons et al. 2011) (Fig. 1), further indicating the presence of several isoflavonoid PTs with varying substrate specificities and regiospecificities in soybean.

The structural diversity of plant specialized metabolites is generally associated with a functional diversification of enzymes caused by gene duplication and nucleotide substitutions, in some cases leading to changes in substrate specificity, regiospecificity and expression patterns (Ober 2005). In our previous studies, the tandem clusters of paralogous genes encoding biosynthetic enzymes in flavonoid and triterpenoid pathways, which suggests their origin from local gene duplication (LGD), were observed in the chromosomes of a model legume *Lotus japonicus* (e.g. Shimada et al. 2003, Sawai et al. 2006, Shimada et al. 2007). Hence, comprehensive identification of isoflavonoid PT genes in soybean, the whole-genome information of which is also available, and their mapping in chromosomes may reveal the molecular basis of the structural diversity of prenylated isoflavonoids in terms of molecular evolution.

In the present study, we used genome sequence information and expressed sequence tags (ESTs) from soybeans to characterize PTs that are involved in isoflavonoid metabolism. Candidate PTs of the isoflavonoid pathway were selected using homology-based in silico screening, and their catalytic functions were confirmed in recombinant yeast microsomes. In these investigations, we identified four novel PTs, one of which is G2DT, and revealed the characteristic arrangements of PT genes in the soybean genome. The present data indicate that soybean PT genes were diversified via LGD and whole-genome duplication (WGD) followed by nucleotide substitution.

### Results

#### Identification of candidate genes for isoflavonoid PTs using in silico screening

Previous studies revealed that (iso)flavonoid-specific PTs exhibit relatively high sequence homology with homogentisate phytyltransferase (HPT), which is encoded by *AtVTE2-1* [National Center for Biotechnology Information (NCBI) accession No. AY089963] of *Arabidopsis thaliana* (Sasaki et al. 2008, Akashi et al. 2009). Based on this notion, we isolated 11 candidate PTs from soybean by in silico screening of homologous sequences in the soybean genomic database (Phytozone v11.0) using *AtVTE2-1* as the query (Table 1). Among the 11 candidate genes, Glyma.13G097800 (DQ2331059), Glyma.10G070100 (AB4334690) and Glyma.01G134600 (GR848991) corresponded to PT1, PT2 and PT3, respectively, which were previously reported by Akashi et al. (2009). PT1 had 78% sequence identity with *AtVTE2-1*, PT2 was found to encode G4DT, whereas PT3 remained uncharacterized due to limited sequence information at that time (Akashi et al. 2009). EST database searches were also performed for transcripts of the candidate genes, but no ESTs for the four candidates Glyma.03G03100, Glyma.10G070100, Glyma.10G070300 and Glyma.11G210500 were found at NCBI. In addition, their transcripts were not detected by PCR using the cDNA templates prepared from yeast extract-treated cultured soybean cells, which were used for cDNA cloning. Thus, in the present study, we performed functional analyses of the candidate genes Glyma.10G070200, Glyma.11G210300, Glyma.11G210400 and Glyma.20G245100 (designated PT4, PT5, PT6 and PT7, respectively), and PT3 (Table 1).

PT3, PT4, PT5, PT6 and PT7 were predicted to encode polypeptides of 393, 412, 402, 410 and 408 amino acids in length, respectively, and they had 47–56% identity with *AtVTE2-1* (Table 1). Amino acid sequences were aligned using Clustal Omega (Supplementary Fig. S1), and putative N-terminal transit peptides of 26, 44, 35, 43, and 43 amino acids in length and 8–9 transmembrane regions, respectively, were identified using transmembrane hidden Markov model (TMHMM) and WoLF PSORT programs. Predicted amino acid sequences also contained two motifs that are highly conserved among divergent cation-dependent PTs (Liang et al. 2002; Supplementary Fig. S1).

On further examination, the phylogenetic relationships among PT1–PT7, related PTs and PTs that act on aromatic compounds including flavonoids and isoflavonoids were investigated using the Neighbor–Joining method (Fig. 2). In the phylogenetic tree, PT3–PT7 are included in a monophyletic clade together with previously reported isoflavonoid PTs from Leguminosae, such as SfG6DT, LaPT and G4DT, suggesting their involvement in the biosynthesis of prenylated (iso)flavonoids. In this clade, PT3–PT7 and G4DT are intimately related to each other compared with those from other leguminous plants, namely *Glycyrrhiza uralensis* (GuA6DT), *S. flavescens* (SfG6DT, SfIdDT

| Sequence ID of the Phytosome database (v11.0) | Identity of the nucleotide and deduced amino acid sequence (in parentheses) to *AtVTE2-1* (%) | Full-length coding sequence information available using EST data | Function of the encoded protein |
|-----------------------------------------------|---------------------------------------------------------------------------------|-------------------------------------------------|-------------------------------|
| Glyma.13G097800 (PT1)                         | 66 (78)                                                                         | Yes                                             | HPT*                          |
| Glyma.03G03100                                  | 54 (55)                                                                         | No                                              | –                             |
| Glyma.10G070100                                  | 58 (54)                                                                         | No                                              | –                             |
| Glyma.10G070300                                  | 57 (49)                                                                         | No                                              | –                             |
| Glyma.11G210500                                  | 60 (56)                                                                         | No                                              | –                             |
| Glyma.10G295300 (PT2)                          | 59 (49)                                                                         | Yes                                             | G4DT*                         |
| Glyma.01G134600 (PT3)                          | 58 (50)                                                                         | No                                              | –                             |
| Glyma.10G070200 (PT4)                          | 58 (49)                                                                         | Yes                                             | G4DT*                         |
| Glyma.11G210300 (PT5)                          | 60 (55)                                                                         | Yes                                             | IDT1†                         |
| Glyma.11G210400 (PT6)                          | 61 (56)                                                                         | Yes                                             | IDT2†                         |
| Glyma.20G245100 (PT7)                          | 58 (47)                                                                         | Yes                                             | G2DT†                         |

* Predicted according to high identity with *AtVTE2-1* (Akashi et al. 2009).
† Experimentally characterized previously (Akashi et al. 2009).
‡ Experimentally characterized in this study; see text.
and SfN8DT) and Lupinus albus (LaPT1). In contrast, PTs responsible for prenylated (iso)flavonoid biosynthesis in Moraceae are located at a distant branch from those of Leguminosae. The PTs of Apiaceae and Cannabaceae, which are involved in the biosyntheses of prenylated coumarins and bitter acids, respectively, also form separate clades. In contrast to PTs in these specialized metabolisms, HPT (encoded by the orthologs of AtVTE2-1), homogentisate geranylgeranyltransferase (HGGT) and homogentisate prenyltransferase (encoded by the orthologs of AtVTE2-2), which are responsible for the biosyntheses of tocopherols, tocotrienols and plastoquinone, respectively, form monophyletic clades, each of which contains PTs of the same function from various plant species (Fig. 2).

Cloning and functional analysis of cDNAs for PT3–PT7

Coding sequences (CDSs) were amplified by PCR using cDNA templates from elicitor-treated cultured soybean cells, although the CDS of PT6 was artificially synthesized due to unsuccessful PCR amplification. Because some plastid-localized PTs partially or completely failed to be functionally expressed in yeast with the transit peptide (Akashi et al. 2009, Shen et al. 2012, Munakata et al. 2014, Li et al. 2015), the CDSs of the N-terminal-truncated forms without putative transit peptides were also PCR-amplified and introduced into yeast expression vectors. Subsequently, 10 constructs, namely the native CDSs of PT3, PT4, PT5, PT6 and PT7 and their truncated forms /C1TP-PT3, /C1TP-PT4, /C1TP-PT5, /C1TP-PT6 and /C1TP-PT7, respectively, were expressed in yeast. Recombinant yeast microsomes were then incubated with DMAPP, and the (iso)flavonoid substrates (2RS)-naringenin, (2S)-liquiritigenin, daidzein, genistein, coumestrol and (−)-glycinol, in the presence of Mg2+, and the reaction mixtures were analyzed using HPLC. Recombinant microsomes expressing ΔTP-PT4 and ΔTP-PT7 displayed dimethylallyltransferase activity toward coumestrol and (−)-glycinol, respectively, and produced single products (Fig. 3A, D). Conversely, microsomes expressing ΔTP-PT5 and ΔTP-PT6 produced single products when incubated with daidzein (Fig. 3B, C). Prenylation products were then confirmed using ultra-performance liquid chromatography (UPLC)–triple quadrupole mass spectrometry (TQMS) analysis in the electrospray ionization (ESI)-positive ion mode. A mass-to-charge ratio (m/z) of 337 was determined for the reaction product obtained from ΔTP-PT4-expressing microsomes and
was matched to that of protonated ions of coumestrol with a dimethylallyl group (Fig. 3E). Furthermore, the m/z values of the reaction products from ΔTP-PT5- and ΔTP-PT6- (323 and 323) expressing microsomes corresponded to those of protonated ions of dimethylallyl-conjugated daidzein (Fig. 3F, G). The reaction product from ΔTP-PT7-expressing microsomes gave an m/z of 323, which was in good agreement with that of dehydrated and protonated ions of (~)-glycinol with a dimethylallyl group (Fig. 3H). The native (untruncated) CDS constructs of PT4, PT5, PT6 and PT7 in recombinant microsomes yielded the same products as those from their truncated forms. Microsomes expressing ΔTP-PT3 or native PT3 exhibited no dimethylallyltransferase activity toward any of the tested substrates.

The enzymatic product from microsomes expressing ΔTP-PT4 was purified and analyzed using 1H-nuclear magnetic resonance (1H-NMR) to investigate further the chemical structure, and the data obtained were compared with those from the authentic sample of the substrate coumestrol. The dimethylallyl moiety of the ΔTP-PT4 product was confirmed by characteristic signals at 1.64 (3H, s, H-5') and 1.81 (3H, s, H-4') for two methyl groups, 3.47 (2H, d, J = 6.9 Hz, H-1') for one methylene...
and 5.22 (1H, m, H-2’) for one methine. The C-4 position, to which the dimethylallyl moiety is attached, was defined by the disappearance of the coumestrol H-4 signal and by a change from the coumestrol double-doublet H-2 signal to a doublet signal of the product. These data indicate that the enzymatic product of PT4-expressing microsomes is 4-dimethylallylcoumestrol. Accordingly, the enzyme encoded by PT4 was designated as DMAPP:coumestrol 4-dimethylallyltransferase (C4DT).

In further 1H-NMR analyses, the chemical structure of the purified enzymatic product from the microsomes expressing ΔTP-PT7 was determined, and the signals for the dimethylallyl moiety were identified as follows: 1.72 (6H, s, H-4’, H-5’) for two methyl groups, 3.28 (2H, d, J = 6.9 Hz, H-1’) for one methylene and 5.35 (1H, m, H-2’) for one methine. In addition, prenylation at the C-2 position was indicated by the disappearance of the glycinol H-2 signal and by the replacement of doublet H-1 and H-4 signals from glycinol to singlet product in the reaction product (Akashi et al. 2009). Taken together, these data reveal that the product from PT7-expressing microsomes is 2-dimethylallylglycinol, which is a direct precursor of glyceollins II and III. Accordingly, the enzyme encoded by PT7 was defined as DMAPP-(6α, 11α)-3,9,6α-trihydroxypterocarpan 2-dimethylallyltransferase [(−)-glycinol 2-dimethylallyltransferase; G2DT].

Microsomes expressing ΔTP-PT5 and ΔTP-PT6 exhibited dimethylallyltransferase activities against daidzein and yielded products with distinct HPLC retention times of 18.7 and 18.9 min, respectively (Fig. 3B, C). Hence, the enzymes encoded by PT5 and PT6 were designated DMAPP:isoflavone dimethylallyltransferase I (IDT1) and IDT2, respectively. However, weak enzymatic activity in microsomes expressing ΔTP-PT5 and ΔTP-PT6 did not yield products sufficient for determining chemical structures via NMR analyses. Hence, the position of the dimethylallyl group was estimated using tandem mass spectrometry (MS/MS) signals from a fragment ion at m/z 137, which was generated from the A-ring of daidzein following a retro Diels–Alder reaction (Maul et al. 2008). The presence of this signal indicated that the A-ring of the reaction product obtained by IDT1 was not enzymatically modified, suggesting prenylation of the B-ring (Supplementary Fig. S2A). In contrast, this signal was absent in experiments with IDT2, suggesting prenylation of the A-ring (Supplementary Fig. S2B).

**Biochemical characterizations of C4DT, IDT1, IDT2 and G2DT**

The substrate specificities of four N-terminal-truncated soybean PTs were tested using (2RS)-naringenin, (2S)-liquiritigenin, daidzein, genistein, coumestrol and (−)-glycinol. 2,7,4’-Trihydroxyisoflavonane, which is the direct daidzein precursor (Fig. 1), was also used for the assays of IDT1 and IDT2. In these experiments, C4DT and G2DT only exhibited activity toward coumestrol and (−)-glycinol, respectively, whereas IDT1 and IDT2 exhibited comparable activity toward daidzein and genistein but not toward 2,7,4’-trihydroxyisoflavonane (Table 2).

In further experiments, C4DT, IDT1, IDT2 or G2DT was incubated without divalent cations and in the presence of 5 mM EDTA, and no products were detected, demonstrating the enzymatic requirement for divalent cations for their catalysis. Moreover, all four enzymes preferred Mg 2+ over Mn 2+ (Supplementary Fig. S3). Lineweaver–Burk plots of G2DT revealed apparent Km values for (−)-glycinol and DMAPP of 183 and 121 μM, respectively.

**Elicitation responses and expression analyses in cultured cells and isolated leaves of soybean**

Prenylated isoflavonoid contents and the corresponding transcript levels of PT genes were analyzed using soybean cells in suspension cultures following treatments with yeast extract [0.3% (w/v) medium]. Yeast extract treatments led to increased accumulation of glycinol I and III, reflecting enhancement of the enzymatic activities of C4DT and G2DT (Akashi et al. 2009). Moreover, the C4DT product 4-dimethylallylcoumestrol was also detected with the two glyceollin isomers (Fig. 4A), and the glyceollin precursor glycinol was increasingly present 10–72 h after treatments with yeast extract. In this system, glycyllin I was the major glyceollin isomer, and it was increasingly present 24–72 h after yeast extract treatment; conversely, the glyceollin III level was approximately a quarter of those of glycinol and glyceollin I. 4-Dimethylallylcoumestrol content was slightly increased following elicitor treatment (Fig. 4A). In contrast, none of the intermediates of glycyllin I and III, 4-dimethylallylglycinol and 2-dimethylallylglycinol, other glyceollin isomers or putative products of IDT1 and IDT2 (dimethylallyldaidzein or dimethylallylgenistein) was detected. Real-time PCR analyses demonstrated that elicitor treatment caused transient up-regulation of C4DT, IDT1, G2DT and G4DT, with co-ordinated transient increases in the expression of the glycinol biosynthetic genes 2-hydroxyisoflavanone synthase 1 (IFS1), IFS2, 2-hydroxyisoflavanone dehydratase (HID) and pterocarpan 6a-hydroxylase (P6aH) (Fig. 4B). The expression of G2DT was induced by elicitor treatment, but it reached a maximum transcript level of no more than 5-fold higher than the steady-state expression level.

Aiming at an alternative phytoalexin induction system in which 2-dimethylallylglycinol-derived glyceollins (II–V) are

**Table 2 Substrate specificities of recombinant PTs for prenyl acceptors**

| Prenyl acceptors             | Relative activities (pmol mg microsomal protein−1 s−1) |
|------------------------------|-------------------------------------------------------|
|                              | C4DT | IDT1 | IDT2 | G2DT |
| (2RS)-Naringenin             | ND   | ND   | ND   | ND   |
| (2S)-Liquiritigenin          | ND   | ND   | ND   | ND   |
| Genistein                    | ND   | 6.9 ± 0.1 | 1.1 ± 0.1 | ND   |
| Daidzein                     | ND   | 8.7 ± 0.7 | 1.1 ± 0.1 | ND   |
| Coumestrol                   | 1.800 ± 60 | ND   | ND   | ND   |
| (−)-Glycinol                 | ND   | ND   | 15.1 ± 1.0 | ND   |
| 2,7,4’-Trihydroxyisoflavonane | ND | ND | ND | ND |

Relative activities of C4DT, IDT1, IDT2 and G2DT toward various flavonoids and isoflavonoids.

Values are presented as mean ± SD of three biological samples.

ND, not detected; −, not tested.
principally induced, isolated soybean leaves were treated with 1 mM cupric chloride (CuCl₂), and glyceollin production was determined according to previous reports on leguminous phytoalexin induction (Dewick 1977, Keen et al. 1986, Yuk et al. 2011). These experiments confirmed the accumulation of glyceollins II–V, again implying increased G2DT activity (Supplementary Text S1). In particular, the levels of glyceollins II–V were increased following CuCl₂ treatment, reaching 150–250 μg g⁻¹ FW at 72 h. Conversely, glyceollin I and dimethylallylisoflavones (dimethylallyldaidzein and dimethylallylgenistein) were not detected. Coumestrol and 4-dimethylallylcoumestrol were also induced by CuCl₂ treatment (Fig. 5A).

Subsequent real-time PCR experiments demonstrated that the expression of glyceollins II–V, again implying increased G2DT activity (Supplementary Text S1). In particular, the levels of glyceollins II–V were increased following CuCl₂ treatment, reaching 150–250 μg g⁻¹ FW at 72 h. Conversely, glyceollin I and dimethylallylisoflavones (dimethylallyldaidzein and dimethylallylgenistein) were not detected. Coumestrol and 4-dimethylallylcoumestrol were also induced by CuCl₂ treatment (Fig. 5A).

Subsequent real-time PCR experiments demonstrated that the expression of IFS1, IFS2, HID, P6aH, G4DT, C4DT, IDT1 and G2DT was increased by CuCl₂ treatment (Fig. 5B), as observed in cultured cells treated with yeast extract. In particular, G2DT transcripts increased by 100-fold. The transcripts of the four genes Glyma.03G033100, Glyma.10G070100, Glyma.10G070300 and Glyma.11G210500 were also undetected in the CuCl₂-treated isolated leaves, as observed in the yeast extract-treated cultured cells.

Arrangements of PT genes in soybean chromosomes

To investigate the evolutionary processes of soybean PT genes responsible for the presence of numerous prenylated isoflavonoids, we investigated the locations of PT2–PT7 and the annotations of adjacent genes using the soybean genome database. Fig. 6A shows a schematic of gene arrangements within approximately 75 kb upstream and downstream regions of each PT gene. The isoflavone PT genes IDT1 and IDT2 are located within 40 kb of each other on chromosome 11, and Glyma.11G210500 is located close to this tandem cluster. C4DT is also accompanied by Glyma.10G070100 and Glyma.10G070300. G4DT and G2DT, which encode enzymes that act on (−)-glycinol, are located in syntenic regions on chromosomes 10 and 20, respectively (Fig. 6A).

Discussion

More than half of the isoflavonoids found in leguminous plants have additional cyclic or acyclic prenyl moieties that originate from dimethylallyl or geranyl groups, and these prenyl modifications often enhance biological activities (Tahara and Ibrahim 1995). The identification of PTs that accept various isoflavonoid substrates is critical to understanding of the structural variations of prenylated isoflavonoids in leguminous plants. In addition, genome arrangements of PT genes might be partly linked to the evolutionary diversification of prenylated isoflavonoids in soybean. Soybean is a good resource for understanding isoflavonoid PT genes, as it produces a vast variety of prenylated isoflavonoids represented by glyceollins, and its detailed whole-genome information is available for gene mining. Characterization of the enzymes responsible for phytoalexin production may facilitate breeding of leguminous crops with enhanced resistance to pathogens.

In the present study, we performed molecular and functional characterization of soybean isoflavonoid PTs using a homology-based approach and identified the four PTs C4DT, IDT1, IDT2 and G2DT. Their deduced transmembrane domains and divalent cation requirements (Supplementary Figs. S1,
S3) were almost consistent with those of previously reported (iso)flavonoid PTs (Sasaki et al. 2008, Akashi et al. 2009, Sasaki et al. 2011, Shen et al. 2012, Chen et al. 2013, Li et al. 2014, Wang et al. 2014). Moreover, the putative N-terminal transit peptide sequences of the soybean PTs implied their plastid localizations and their functions in prenylation reactions occurring in the plastid or its related subcellular compartment, as described in previous works on SfN8DT and G4DT (Sasaki et al. 2008, Akashi et al. 2009), although their subcellular localizations were not experimentally verified. Among the PT candidates obtained by in silico screening, we did not analyze functionally the four genes, Glyma.03G033100, Glyma.10G070100, Glyma.10G070300 and Glyma.11G210500 because their transcripts were not observed in elicitor-treated cultured cells and the isolated leaves of soybean, as well as in the EST database at NCBI. Their spatio-temporal expression may be strictly regulated.

The roles of soybean PTs in the biosynthesis of prenylated isoflavonoid

Our data illustrated that C4DT catalyzes the production of 4-dimethylallylcoumestrol from coumestrol (Fig. 3A). To our knowledge, this is the first identified plant PT that acts on coumestans including coumestrol as substrates. 4-Dimethylallylcoumestrol was induced in yeast extract-treated suspension cultures and in isolated leaves treated with CuCl₂. In agreement with this, transcription of C4DT was up-regulated prior to the increase in 4-dimethylallylcoumestrol levels in both of these systems (Figs. 4A, B, 5A, B). A previous study revealed that the disease resistance-inducing herbicide lactofen induces 4-dimethylallylcoumestrol accumulation in the ‘minimal-wound snapped cotyledon assay’ (Cheng et al. 2011). Dimethylallylcoumestrol accumulation coupled with the expression patterns of C4DT on CuCl₂ treatment suggests its involvement in defensive responses to biotic or abiotic stresses.

IDT1 and IDT2 were found to catalyze dimethylallylation of isoflavones (daidzein and genistein), presumably producing B- and A-ring prenylated isoflavones, respectively (Fig. 3B, C). In the previously reported ‘minimal-wound snapped cotyledon assay’, production of the B- and A-ring prenylated isoflavones 3'-dimethylallyldaidzein (3'-dimethylallylgenistein) and 8-dimethylallyldaidzein, respectively, which are possible products of IDT1 and IDT2, respectively, was induced by lactofen and by a glucan from the cell walls of the pathogenic bacterium Phytophthora sojae (Cheng et al. 2011). In the present study, IDT1 transcript levels increased by approximately 10,000-fold after elicitor treatment of cultured cells and isolated leaves, implying that it is involved in defense responses (Figs. 4B, 5B). However, 3'-dimethylallyldaidzein, 3'-dimethylallylgenistein, 8-dimethylallyldaidzein and other prenylated isoflavones were not detected in these induced tissues, and IDT2 transcripts were not detected in cultured cells or isolated leaves. Hence, these data do not support the roles of IDT1 or IDT2 in the biosynthesis of 3'-dimethylallyldaidzein (3'-dimethylallylgenistein) or 8-dimethylallyldaidzein. Further investigation will be required to elucidate their roles in soybean isoflavonoid biosynthesis.

G2DT was found to catalyze the conversion of (−)-glycinol into 2-dimethylallylglycinol, which is a precursor of glyceollins II, III, IV and V (Fig. 3D). The regiospecificity of this prenylation reaction was unequivocally confirmed in comparisons of 1H-NMR spectra for glycinol and the G2DT products. The apparent $K_m$ values of G2DT for (−)-glycinol (183 μM) and DMAPP
(121 μM) were comparable with those of G4DT (68 and 150 μM, respectively), whereas the affinity of G4DT for the prenyl acceptor was slightly higher (Akashi et al. 2009). To test the role of G2DT in plant tissues, we employed two induction systems of glyceollins: suspension-cultured cells, which were previously used for the analysis of G4DT (Akashi et al. 2009), and isolated leaves, in which glyceollins II and III (Keen et al. 1986) and IV and V (Yuk et al. 2011) were strongly induced by elicitation (Pseudomonas syringae pv. pisi and iodoacetate) and wounding, respectively. Following yeast extract treatment of the suspension-cultured cells, up-regulation of G2DT transcripts by approximately 5-fold and slight accumulation of glyceollin III were observed (Fig. 4A, B). In additional experiments using CuCl2 in isolated soybean leaves, glyceollins II–V were markedly induced (Fig. 5A), although our elicitation procedures differed from those of previous studies. Moreover, G2DT transcription was increased by >100-fold prior to the accumulation of glyceollin isomers (Fig. 5B). Taken together, the pathway via 2-dimethylallylglycinol was induced more intensively in CuCl2-treated isolated leaves than in yeast extract-treated cultured cells, and the relative transcript level of G2DT was consistent with the accumulation levels of 2-dimethylallylglycinol-derived glyceollins (II–V) in the induction systems. These results support the notion that G2DT is involved in glyceollin II–V biosynthesis. In contrast, glyceollin I was not detected in elicited leaves despite the considerable 100-fold increase in G4DT transcript levels (Fig. 5B). This discrepancy suggests the conversion of glyceollin I to an unknown end-product or inactivation by (post)translational regulation of G4DT, warranting further investigation of metabolite profiling and G4DT protein expression.

In glyceollin biosynthesis, the subsequent cyclization reactions are reportedly catalyzed by cytochrome P450 monooxidases (P450s), but none of them has been molecularly characterized (Welle and Grisebach 1988, Ayabe and Akashi 2006). The molecular clones for G4DT and G2DT will be used to prepare
the substrates for the P450 cyclases, facilitating their molecular identification (Fig. 1). Because G4DT (and probably also G2DT) is plastid localized, the identification of the P450s involved in the next step of G4DT and G2DT will help to reveal the intriguing subcellular compartment in which glyceollins are produced.

Molecular evolution of isoflavonoid PTs in leguminous plants

Phylogenetic analysis of the PTs responsible for the prenylation of aromatic compounds revealed that clades are formed in two distinct manners (Fig. 2). One includes the PTs involved in the biosyntheses of tocophorols, tocotrienols and plastoquinone, which are essential for the basic activities of plants at the cellular level and thus can be considered to reflect ‘primary’ metabolism. In this case, a clade consists of the PTs that possess the same catalytic function but originate from various plant species. In contrast, PTs involved in plant specialized metabolism, such as the biosyntheses of prenylated (iso)flavonoids, coumarins and bitter acids, form their clades in the other manner, in which close phylogenetic relationships are observed among the PTs that originate from the same plant species but not among those that possess the same or similar function. The latter pattern was only clarified after multiple (iso)flavonoid PTs were identified from two species, *S. flavescens* and soybean. These phylogenetic relationships suggest that primary metabolism PTs were established prior to speciation in the angiosperm and that PTs in plant specialized metabolism were diversified after the establishment of each plant lineage. PTs that act on isoflavones, such as IDT1, IDT2, LaPT, SIG6DT, MalDT and CtiDT, have thus far been identified from several plant lineages, and they are dispersedly located in the phylogenetic tree, suggesting their convergent origins. Non-leguminous MalDT and CtiDT, which are located in the distant branch, are likely to have originated from the PTs responsible for plastoquinone biosynthesis, whereas the other isoflavone PTs from leguminous plants probably originated from those for tocopherols (Fig. 2).

The functional diversification process of soybean isoflavonoid PTs suggested by the genome organization

The genomic distribution of the six PT genes investigated in our studies and their adjacent genes in soybean chromosomes led to two notable findings (Fig. 6A). First, IDT1 and IDT2, which encode PTs with similar specificity for isoflavone substrates but distinct regiospecificities of prenylation, were located in tandem on the same chromosome. Furthermore, other PT genes were found adjacent to the IDT1–IDT2 cluster and G4DT, although their expression was not detected in this study. The two genes (IDT1 and IDT2) were reportedly produced by an LGD after the recent WGD (Wang et al. 2015). Taken together, clusters of PT genes including putative isoflavonoid PT genes indicate that they emerged via LGD events in soybean. Accordingly, subsequent nucleotide substitution mutations in the duplicate genes probably changed the properties of encoded PTs. In the genome of a model legume *Lotus japonicus*, tandem clusters of genes that encoded enzymes of similar but distinct catalytic functions were observed, including those encoding the two oxidosqualene cyclases, cycloartenol synthase and lupeol synthase (Sawai et al. 2006), Type I and II chalcone isomerases (Shimada et al. 2003), and pterocarpan reductases with various stereospecificities (Akashi et al. 2006, Shimada et al. 2007). The PT gene clusters on chromosomes 10 and 11 of soybean provide evolutionary examples of biosynthetic genes in specialized metabolism via LGD followed by nucleotide substitution.

Secondly, we observed a syntenic relationship between the approximately 150 kb flanking region of G4DT on chromosome 10 and that of G2DT on chromosome 20 (Fig. 6A). Recent genomic investigations of angiosperm indicated frequent WGDs caused by polyploidization and suggested that repeated WGDs were responsible for speciation between angiosperm lineages (Fawcett et al. 2009, Van de Peer et al. 2009, Severin et al. 2011). Soybean has a relatively large number of chromosomes (2n = 40) compared with other leguminous plants and has a complicated genome organization, in which >75% of genes are present in multiple copies. This striking genomic feature of soybean is attributed to WGDs approximately 13 and 59 million years ago (Schmutz et al. 2010, Fig. 6B). The synteny observed in the flanking regions of G4DT and G2DT demonstrates that the two genes are located in homoeologous regions deriving from the WGD that reportedly occurred 13 million years ago (Severin et al. 2011, Wang et al. 2015), and they have been functionally specialized from a common ancestral PT gene. These findings indicate that diversification of genes in plant specialized metabolism can occur after WGDs as well as LGDs. The previously reported conserved gene duplicates in homoeologous genomic regions in soybean included PT3 and Glyma.03G033100, as well as G4DT and G2DT (Severin et al. 2011). Their functions remain unknown at present, but their closely related functions may be clarified in a future study.

Materials and Methods

**Chemicals**

DMAPP was purchased from Cayman Chemical, and (2R5)-naringenin and coumestrol were purchased from Tokyo Chemical Industry Co., Ltd. and Sigma-Aldrich, respectively. Daidzein and genistein were purchased from LC Laboratories, and (2S)-liquiritigenin was obtained from laboratory stock. (−)-Glycinol and glyceollins I and II were prepared as described previously (Akashi et al. 2009). 2,7,4-Trihydroxyisoflavone was prepared using Escherichia coli expressing the cDNA for CYP93C2 (Uchida et al. 2015).

**Plant materials**

Maintenance of soybean callus and treatment of suspension cultures with yeast extract were performed as described previously (Akashi et al. 2009). Soybean seeds (cv. Fuki) were purchased from Takii Seed Co., Ltd., germinated on an artificial soil and grown in a greenhouse at 25°C under a 16 h light/8 h dark photoperiod. Leaves were then isolated from 5-week-old plants and soaked in 1 mM CuCl₂ solution.

**Analytical methods**

HPLC analyses were performed with an LC-2000 series HPLC system (JASCO) equipped with a multivavelength detector (MD-2010, JASCO) using a TSK-Gel
ODS-80TS column (4.6 x 150 mm; Tosoh) and a flow rate of 1.0 ml min\(^{-1}\) at 40°C. Gradient elution was performed using water (A) and methanol (B) or acetonitrile (C) with the following gradient programs: program 1 for analysis of prenylated isoflavonoids in suspension cultures (0–30 min, 40–100%; B); program 2 for analysis of prenylated isoflavonoids in soybean leaves (0–20 min, 40–80%; C); and program 3 for enzyme activity assays of PTs (0–30 min, 20–80%; C). Eluates were monitored at 260 nm for isoflavones, 280 nm for (−)-gallocatechin and (-)-epigallocatechin, and 345 nm for microsomal and 4,6-dimethoxyflavonocoumarin.

UPLC-TQMS analyses were performed using a Quattro premier XE (Waters) instrument. UPLC separation was performed using an ACQUITY UPLC BEH C18 column (2.1 x 100 mm, Waters) with a flow rate of 0.38 ml min\(^{-1}\) at 40°C. Gradient elution was performed with water (D) and methanol (E) or acetonitrile (F) in the presence of 0.1% (v/v) formic acid as follows: program 1 for analyses of prenylated isoflavonoids in suspension cultures (0–2 min, 5–30%; 2–10 min, 30–70%; and 10–12 min, 70–95% E); and program 2 for analyses of prenylated isoflavonoids in soybean leaves and the enzyme activity assays of PTs (0–2 min, 5–30%; 2–10 min, 30–70%; and 10–12 min, 70–95% F). MS spectra were measured using Quattro Premier Ver 4.1 software (Waters) under the following conditions: ESI-positive mode; source temperature of 120°C, desolvation gas with a flow rate of 800 l h\(^{-1}\) at 400°C, cone gas with a flow rate of 50 l h\(^{-1}\), cone voltage of 40 V and capillary voltage of 3 kV. The MS/MS spectra of IDT1 and IDT2 products were acquired using a collision energy ramp of 10–40 eV.

In silico screening of candidate genes for putative isoflavonoid PTs

Putative genes encoding isoflavonoid PTs were retrieved from the soybean database at Phytozome v11.0 (https://phytozome.jgi.doe.gov/pz/portal.html) using the amino acid sequence of A. thaliana HPT (AtVTE2-1; AY089963) as the query, and corresponding transcripts were searched in the EST database at NCBI. The correspondence of the sequence ID in Wm82a2v1.1 is shown in Supplementary Table S1. Protein motifs were predicted using TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) and Wolf PSORT programs (http://www.genoscope.cns.fr/psort/wolf_psort.html). Predicted genes that were located close to isoflavonoid PT genes on soybean chromosomes were also examined using the soybean database at Phytozome v11.0.

Cloning of cDNAs for putative isoflavonoid PTs

Total RNA was prepared from the elicitor-treated soybean cells using an SV total RNA isolation system (Promega), and cDNA templates for PCR amplification of CDSs (excluding PT6) were synthesised using a SuperScript III First-Strand Synthesis System (Invitrogen). Because the PCR product of PT6 was not obtained using the cDNA samples, a template for PT6 was synthesized using Strand Synthesis System (Invitrogen). The PCR products of PT3 and PT7 were directly cloned into yeast expression vector pYES-DEST52 using a Gateway LR Clonase II Enzyme GW/TOPO or pENTR/D-TOPO Gateway Technology and introduced into the yeast expression vector pYES-DEST52 using a Gateway LR Clonase II Enzyme Mix (Invitrogen). The PCR products of PT3 and PT7 were directly cloned into the yeast expression vector pYES2.1/TOPO (Invitrogen).

PT assays

_Saccharomyces cerevisiae_ BJ2168 cells were transformed with expression vectors, and recombinant protein expression and microsome preparations were performed as described previously (Akashi et al. 1999). The protein contents of microsomes were determined using the Bradford method (Bradford 1976), and PT assays were performed as described previously (Akashi et al. 2009) with some modifications. Briefly, to a total reaction volume of 200 μl, substrates and microsomes were added at the indicated concentrations and incubated for the indicated times. The specific activities of C4DT, IDT1, IDT2 and G2DT were determined after incubating 400 μM prenyl acceptor, 400 μM DMAPP and recombinant microsomes containing 15–100 μg of protein at 30°C for 15–60 min. Kinetic studies of G2DT were performed in the presence of 5, 10, 20, 40, 80, 160 or 400 μM (−)-gallocatechin and (-)-epigallocatechin and a fixed concentration of DMAPP (400 μM) or with a fixed concentration of (−)-gallocatechin (400 μM) and varying concentrations (10, 20, 40, 80, 160 and 400 μM) of DMAPP after incubation with microsomes containing 50 μg of protein at 30°C for 15 min. Apparent Km values were calculated using Lineweaver–Burk plots.

Identification of C4DT and G2DT reaction products

To identify the C4DT product, 800 μg of coumestrol, 800 μg of DMAPP and recombinant microsomes expressing the truncated form of PT4 (~13 mg of protein) were incubated at 30°C overnight. Subsequently, the reaction product was purified from an ethyl acetate extract of the reaction mixture using silica gel thin-layer chromatography (TLC) on a plate (Kieselgel F254; Merck) with chloroform:methanol (9:1, v/v) as the eluting solvent. NMR spectra were acquired using a JMN ECA-500 system (JEOL). 4,6-Dimethoxyflavonocoumarin, 1H-NMR (DMSO-d6) δ: 1.64 (3H, s, H-5'), 1.81 (3H, s, H-4'), 3.47 (2H, d, J = 6.9 Hz, H-7'), 5.22 (1H, m, H-2'), 6.93 (1H, d, J = 8.6 Hz, H-2), 6.94 (1H, dd, J = 2.3, 8.6 Hz, H-8), 7.15 (1H, d, J = 2.3 Hz, H-10), 7.68 (1H, d, J = 8.6 Hz, H-7), 7.69 (1H, d, J = 8.6 Hz, H-1). To validate the signal assignments of 4,6-dimethoxyflavonocoumarin, authentic coumestrol was measured using NMR. Coumestrol, 3H-NMR (DMSO-d6) δ: 6.91 (1H, d, J = 2.3 Hz, H-4), 6.94 (1H, dd, J = 2.3, 8.6 Hz, H-2), 6.95 (1H, dd, J = 2.3, 8.6 Hz, H-8), 7.17 (1H, d, J = 2.3 Hz, H-10), 7.70 (1H, d, J = 8.6 Hz, H-7), 7.86 (1H, d, J = 8.6 Hz, H-1); 13C-NMR (DMSO-d6): δ: 98.7 (C-10), 102.0 (C-6a), 103.0 (C-4), 104.1 (C-11b), 113.8 (C-2), 114.6 (C-8b), 121.0 (C-7), 122.7 (C-2), 154.7 (C-4a), 156.0 (C-9), 157.0 (C-10a), 1576 (C-6), 1595 (C-11a), 161.3 (C-3). To identify the G2DT product, 400 μg of (−)-gallocatechin, 400 μg of DMAPP and recombinant microsomes expressing the truncated form of PT7 (~7 mg of protein) were incubated at 30°C overnight. The reaction product was then purified from ethyl acetate extract of the reaction mixture using silica gel TLC with hexane:ethyl acetate (3:7, v/v) as the eluting solvent. 2,3-Dimethoxy-7-glycinol, 1H-NMR (acetonitrile-d6): δ: 1.72 (6H, s, H-4, H-5'), 3.28 (2H, d, J = 6.9 Hz, H-7'), 3.97 (1H, d, J = 11.5 Hz, H-6), 4.09 (1H, d, J = 11.5 Hz, H-5), 5.23 (1H, s, H-11a), 5.35 (1H, m, H-2'), 6.25 (1H, d, J = 2.3 Hz, H-10), 6.35 (1H, s, H-4'), 6.42 (1H, dd, J = 2.3, 8.0 Hz, H-8), 7.16 (1H, s, H-1), 7.19 (1H, d, J = 8.0 Hz, H-7).

Prenyled isoflavonoid analyses in elicitor-treated soybean tissues

Aliquots (10 ml) of elicited suspension cultures (250 ml) were periodically collected and extracted with ethyl acetate using a Polytrom homogenizer as described previously (Akashi et al. 2009). Elicitor-treated leaves were periodically collected and ground into a fine powder in liquid nitrogen using a mortar and pestle and then extracted with 20 vol. of methanol overnight. Extracts were evaporated to dryness, dissolved in methanol and analyzed using HPLC and UPLC-TQMS. Coumestrol and prenylated isoflavonoids were identified by comparing retention times, UV-VIS spectra and MS spectra with those of authentic standards.

Real-time PCR of genes involved in the biosynthesis of prenylated isoflavonoids

Real-time PCR was performed using SYBR Green PCR Master Mix and a Fast Real-Time PCR system (Applied Biosystems) according to the manufacturer’s protocol. Data were analyzed using the comparative Ct method with the housekeeping gene SKP1/ASK-interacting protein 16 (SKIP16) as an internal standard (Hu et al. 2009). The primers for real-time PCR are listed in Supplementary Table S2.

Phylogenetic analysis

Amino acid sequences of selected PTs with specificity for aromatic substrates including (iso)flavonoids were analyzed using the ClustalW program (Thompson et al. 1994) of the DNA Data Bank of Japan. The phylogenetic tree was constructed using the Neighbor–Joining method from the results of 1,000 bootstrap replicates and was presented as a non-root tree using TreeView (Page 1996).

Supplementary data

Supplementary data are available at PCP online.
Funding

This work was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan [Grant-in-Aid for Scientific Research (C) (No. 21510232)] and the Nakato Scholarship Foundation (Japan) [a Grant-in-Aid to K.Y.].

Acknowledgements

We wish to thank Tsuyoshi Nakagawa (Nihon University) for technical assistance and Dr. Kai Uchida for critical reading of the manuscript and many helpful suggestions.

Disclosures

The authors have no conflicts of interest to declare.

References

Akashi, T., Aoki, T. and Ayabe, S. (1998) Identification of a cytochrome P450 cDNA encoding flavone 2-hydroxylase of licorice (Glycyrrhiza urchina L.; Fabaceae), which represents licidione synthase and flavone synthase II. FEBS Lett. 431: 287–290.

Akashi, T., Koshimizu, S., Aoki, T. and Ayabe, S. (2006) Identification of cDNAs encoding pterocarpan reductase involved in isoflavan phytoalexin biosynthesis in Lotus japonicus by EST mining. FEBS Lett. 580: 5666–5670.

Aoki, T., Sasaki, K., Aoki, T., Ayabe, S. and Yazaki, K. (2009) Molecular cloning and characterization of a cDNA for pterocarpan 4-dimethylallyltransferase catalyzing the key prenylation step in the biosynthesis of glyceollin, a soybean phytoalexin. Plant Physiol. 149: 683–693.

Ayabe, S. and Akashi, T. (2000) Flavonoids of leguminous plants: structure, biological activity, and biosynthesis. J. Plant Res. 113: 475–488.

Ayabe, S. and Akashi, T. (2006) Cytochrome P450s in flavonoid metabolism. Phytochem. Rev. 5: 271–282.

Botta, B., Vitali, A., Menendez, P., Misiti, D. and Delle Monache, G. (2005) Prenylated flavonoids: pharmacology and biotechnology. Curr. Med. Chem. 12: 717–739.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72: 248–254.

Chen, R., Liu, X., Zou, J., Yin, Y., Ou, B., Li, J. et al. (2013) Regio- and stereospecific prenylation of flavonoids by Lotus japonicus. Anal. Biochem. 428: 25–36.

Cheng, J., Yuan, C. and Graham, T.L. (2011) Potential defense-related prenylated isoflavones in lactofen-induced soybean. Phytochemistry. 72: 875–881.

Dewick, P.M. (1977) Biosynthesis of pterocarpan phytoalexins in Trifolium pratense. Phytochemistry 16: 93–97.

Fawcett, J.A., Maere, S. and Van de Peer, Y. (2009) Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. Proc. Natl. Acad. Sci. USA 106: 5737–5742.

Hu, R., Fan, C., Li, H., Zhang, Q. and Fu, Y.F. (2009) Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT–PCR. BMC Mol. Biol. 10: 93–104.

Karamat, F., Olyry, A., Munakata, R., Koeduka, T., Sugiyama, A., Paris, C., et al. (2014) A coumarin-specific prenyltransferase catalyzes the crucial biosynthetic reaction for furanocoumarin formation in parsley. Plant J. 77: 627–638.

Keen, N.T., Lyne, R.L. and Hymowitz, T. (1986) Phytoalexin production as a chemosystematic parameter within the genus Glycine. Biochem. Syst. Ecol. 14: 481–486.

Li, J., Chen, R., Wang, R., Liu, X., Xie, D., Zou, J., et al. (2014) GuA6DT, a regiospecific prenyltransferase from Glycyrrhiza uralensis, catalyzes the 6-prenylation of flavones. Chembiochem 15: 1673–1681.

Li, H., Ban, Z., Qin, H., Ma, L., King, A.J. and Wang, G. (2015) A heteromeric membrane-bound prenyltransferase complex from Humulus lupulus catalyzes three sequential aromatic prenylations in the bitter acid pathway. Plant Physiol. 167: 650–659.

Liang, P.H., Ko, T.P. and Wang, A.H.J. (2002) Structure, mechanism and function of prenyltransferases. Eur. J. Biochem. 269: 3339–3354.

Maul, R., Scheeb, N.H. and Kulling, S.E. (2008) Application of LC and GC hyphenated with mass spectrometry as tool for characterization of unknown derivatives of isoflavonoids. Anal. Bioanal. Chem. 391: 239–250.

Munakata, R., Inoue, T., Koeduka, T., Karamat, F., Olyry, A., Sugiyama, A., et al. (2014) Molecular cloning and characterization of a GDP-specific aromatic prenyltransferase from Citrus limon. Plant Physiol. 166: 80–90.

Munakata, R., Olyry, A., Karamat, F., Courdavault, V., Sugiyama, A., Date, Y., et al. (2016) Molecular evolution of parsnip (Pastinaca sativa) membrane-bound prenyltransferases for linear and/or angular furanocoumarin biosynthesis. New Phytol. 211: 332–344.

Oba, D. (2005) Seeing double: gene duplication and diversification in plant secondary metabolism. Trends Plant Sci. 10: 444–449.

Page, R.D. (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12: 357–358.

Sasaki, K., Tsurumaru, Y., Yamamoto, H. and Yazaki, K. (2008) Cloning and characterization of naringenin 8-prenyltransferase, a flavonoid-specific prenyltransferase of Sophora flavescens. Plant Physiol. 146: 1075–1084.

Severin, A.J., Cannon, S.B., Shoemaker, R.C. and Grant, D. (2014) Genome-wide analyses of the structural gene families involved in flavonoid biosynthesis in Trema orientalis. J. Plant Res. 127: 247–257.

Simons, R., Vincken, J.P., Bohin, M.C., Kuijpers, T.F., Verbruggen, M.A. and Schmutz, J. (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183.

Shen, G., Huhman, D., Le, Z., Snyder, J., Sumner, L.W. and Dixon, R.A. (2012) Characterization of an isoflavonoid-specific prenyltransferase from Lupinus albus. Plant Physiol. 159: 70–80.

Shimada, N., Aoki, T., Sato, S., Nakamura, Y., Tabata, S. and Ayabe, S. (2003) A cluster of genes encodes the two types of chalcone isomerase involved in the biosynthesis of general flavonoids and legume-specific 5-deoxy(iso)flavonoids in Lotus japonicus. Plant Physiol. 131: 941–951.

Shimada, N., Sato, S., Akashi, T., Nakamura, Y., Tabata, S. and Ayabe, S. (2007) Genome-wide analyses of the structural gene families involved in the legume-specific 5-deoxyisoflavonoid biosynthesis of Lotus japonicus. DNA Res. 14: 25–36.

Simons, R., Vincken, J.P., Bohin, M.C., Kuipers, T.F., Verbruggen, M.A. and Gruppen, H. (2011) Identification of prenylated pterocarpons and other isoflavonoids in Rhizopus spp. elicted soya bean seedlings by electrospray ionisation mass spectrometry. Rapid Commun. Mass Spectrom. 25: 55–65.

Tahara, S. and Ibrahim, R.K. (1995) Prenylated isoflavonoids—an update. J. Plant Res. 113: 181–189.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673–4680.
Uchida, K., Akashi, T. and Aoki, T. (2015) Functional expression of cytochrome P450 in *Escherichia coli*: an approach to functional analysis of uncharacterized enzymes for flavonoid biosynthesis. *Plant Biotechnol.* 32: 205–213.

Van de Peer, Y., Fawcett, J.A., Proost, S., Sterck, L. and Vandepoele, K. (2009) The flowering world: a tale of duplications. *Trends Plant Sci.* 14: 680–688.

Wang, R., Chen, R., Li, J., Liu, X., Xie, K., Chen, D., et al. (2014) Molecular characterization and phylogenetic analysis of two novel regio-specific flavonoid prenyltransferases from *Morus alba* and *Cudrania tricuspidata*. *J. Biol. Chem.* 289: 35815–35825.

Wang, J., Chu, S., Zhu, Y., Cheng, H. and Yu, D. (2015) Positive selection drives neofunctionalization of the UbiA prenyltransferase gene family. *Plant Mol. Biol.* 87: 383–394.

Welle, R. and Grisebach, H. (1988) Induction of phytoalexin synthesis in soybean: enzymatic cyclization of prenylated pterocarps to glyceollin isomers. *Arch. Biochem. Biophys.* 263: 191–198.

Yang, X., Jiang, Y., Yang, J., He, J., Sun, J., Chen, F., et al. (2015) Prenylated flavonoids, promising nutraceuticals with impressive biological activities. *Trends Food Sci. Technol.* 44: 93–104.

Yazaki, K., Sasaki, K. and Tsurumaru, Y. (2009) Prenylation of aromatic compounds, a key diversification of plant secondary metabolites. *Phytochemistry* 70: 1739–1745.

Yuk, H.J., Curtis-Long, M.J., Ryu, H.W., Jang, K.C., Seo, W.D., Kim, J.Y., et al. (2011) Pterocarpan profiles for soybean leaves at different growth stages and investigation of their glycosidase inhibitions. *J. Agric. Food Chem.* 59: 12683–12690.