

Enzymatic and metabolic engineering for efficient production of syringin, sinapyl alcohol 4-*O*-glucoside, in *Arabidopsis thaliana*



Yang Chu, Tackmin Kwon, Jaesung Nam*

Department of Molecular Biotechnology, Dong-A University, Busan 604-714, South Korea

ARTICLE INFO

Article history:

Received 22 May 2013

Received in revised form 21 February 2014

Available online 22 March 2014

Keywords:

Arabidopsis thaliana

Cruciferae

Syringin

Sinapyl alcohol 4-*O*-glucoside

Coniferin

Coniferyl alcohol 4-*O*-glucoside

Monolignol biosynthesis

Glucosyltransferase 72E family

Phenylpropanoid pathway

Gene stacking

ABSTRACT

To promote efficient production of syringin, a plant-derived bioactive monolignol glucoside, synergistic effects of enzymatic and metabolic engineering were combined. Recombinant *UGT72E3/E2* chimeras, generated by exchanging parts of the C-terminal domain including the Putative Secondary Plant Glycosyltransferase (PSPG) motif of *UGT72E3* and *UGT72E2*, were expressed in leaves of transgenic *Arabidopsis* plants; syringin production was measured *in vivo* and by enzymatic assays *in vitro*. In both tests, *UGT72E3/2* displayed substrate specificity for sinapyl alcohol like the parental enzyme *UGT72E3*, and the syringin production was significantly increased compared to *UGT72E3*. In particular, in the *in vitro* assay, which was performed in the presence of a high concentration of sinapyl alcohol, the production of syringin by *UGT72E3/2* was 4-fold higher than by *UGT72E3*. Furthermore, to enhance metabolic flow through the phenylpropanoid pathway and maintain a high basal concentration of sinapyl alcohol in the leaves, *UGT72E3/2* was combined with the sinapyl alcohol synthesis pathway gene *F5H* encoding ferulate 5-hydroxylase and the lignin biosynthesis transcriptional activator *MYB58*. The resulting *UGT72E3/2+F5H+MYB58* OE plants, which simultaneously overexpress these three genes, accumulated a 56-fold higher level of syringin in their leaves than wild-type plants.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

Syringin (sinapyl alcohol 4-*O*-glucoside) (**2**) (Fig. 1) is a natural product distributed widely throughout many types of plants, with massive accumulation restricted to some medicinal plants. It has been reported that it has various pharmacological effects with little toxicity, including anti-inflammatory, anti-nociceptive, immunomodulatory, and anti-diabetic effects (Cho et al., 2001; Choi et al., 2004; Liu et al., 2008; Niu et al., 2008). These health-promoting effects of syringin (**2**) attract considerable interest for novel applications. Although its biosynthesis pathway is well-characterized,

many questions need to be elucidated to improve syringin (**2**) production and accumulation in plants.

Syringin (**2**) biosynthesis branches from the phenylpropanoid pathway, which is largely a highly conserved means of lignin production in the plant kingdom (Anterola and Lewis, 2002). Sinapyl alcohol (**1**), a monolignol synthesized via the phenylpropanoid pathway, is converted to syringin (**2**) by plant family 1 UDP-dependent glucosyltransferases (UGTs). In *Arabidopsis thaliana*, 120 *UGT* genes have been identified. Among them, a small cluster of three closely related genes (*At3g50740*, *At5g66690*, *At5g26310*) encode the *UGT72E* family, which are responsible for monolignol 4-*O*-glucoside production (Lanot et al., 2006; Lim et al., 2001) (Fig. 1). *UGT72E2* preferentially glycosylates coniferyl alcohol (**3**) and has a high catalytic activity, which results in efficient production of coniferin (**4**). In contrast, *UGT72E3* has high specificity for the glycosylation of sinapyl alcohol (**1**), but its low catalytic activity results in inefficient production of syringin (**2**) (Lanot et al., 2006, 2008; Lim et al., 2001). An improvement of *UGT72E3* catalytic activity is thus a prerequisite for the application of this enzyme for syringin (**2**) production in plants.

Low availability of the substrate for UGT is another obstacle that must be surmounted to achieve efficient syringin (**2**) production in plants. There are three enzymatic branching points in the

Abbreviations: OE, overexpression; HPLC, high performance liquid chromatography; PSPG, Putative Secondary Plant Glycosyltransferase; UGT, UDP-glucose:alcohol glucosyltransferase; F5H, ferulate 5-hydroxylase; PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-(hydroxy) cinnamoyl CoA ligase; HCT, hydroxycinnamoyl CoA:shikimate/quinate hydroxycinnamoyltransferase; CHS, chalcone synthase; C3H, *p*-coumarate 3-hydroxylase; CCoAOMT, caffeoyl CoA 3-*O*-methyltransferase; CCR, cinnamoyl CoA reductase; COMT, caffeic acid *O*-methyltransferase; CAD, cinnamyl alcohol dehydrogenase.

* Corresponding author. Address: Department of Molecular Biotechnology, Dong-A University, Saha-gu, Hadan-2-dong, Busan 604-714, South Korea. Tel.: +82 51 200 7518; fax: +82 51 200 7505.

E-mail address: jnam@dau.ac.kr (J. Nam).

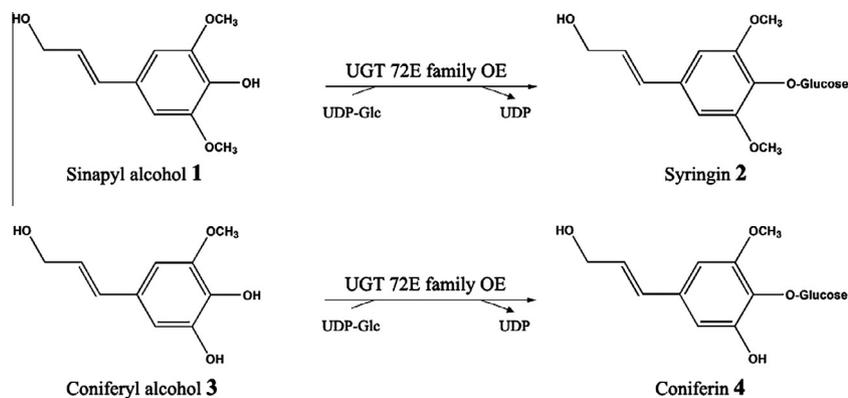


Fig. 1. Chemical structures of monolignols and their glucosides.

phenylpropanoid pathway that control the flux of carbon towards sinapyl alcohol (**1**): chalcone synthase (CHS), hydroxycinnamoyl-Coenzyme A shikimate/quinate hydroxycinnamoyl transferase (HCT) and ferulate 5-hydroxylase (F5H). At these critical enzymatic reaction points, CHS and HCT direct *p*-coumaroyl CoA, a versatile metabolite generated by the phenylpropanoid pathway, into flavonoid synthesis and guaiacyl (G)-type monolignol synthesis, respectively (Anterola and Lewis, 2002; Whetten and Sederoff, 1995). F5H further catalyzes the conversion of G-type monolignol to syringyl (S)-type monolignol (**1**), which is a direct precursor for syringin (**2**) production (Vanholme et al., 2008). The details of the transcriptional network of monolignol biosynthesis in *Arabidopsis* have been elucidated (Zhou et al., 2009; Zhao et al., 2010). The lignin-specific transcriptional factor MYB58 directly activates most monolignol biosynthesis genes except *F5H*, which is regulated by a secondary cell wall master regulator SND1.

Light exposure induces accumulation of the monolignol glucosides, coniferin (**4**) and syringin (**2**), in *Arabidopsis* roots. This accumulation is a result of both increased production of these compounds by the phenylpropanoid pathway and altered utilization of these compounds by the lignin synthesis pathway. However, light-induced accumulation of monolignol glucosides does not occur in aerial tissues (Hemm et al., 2004).

In this study, the synergistic effects of enzymatic and metabolic engineering methods for the production of syringin (**2**) in *Arabidopsis* leaves was demonstrated. Transgenic plants overexpressing a novel recombinant *UGT72E3/2*, generated by domain swapping between *UGT72E3* and *UGT72E2*, improved syringin (**2**) production compared to transgenic plants overexpressing *UGT72E3* or *UGT72E2*. Further, metabolic activation for S-type monolignol biosynthesis increased metabolic flux towards sinapyl alcohol (**1**). Consequently, a combination of these effects via gene stacking resulted in a drastic increase of syringin (**2**) production in the leaves of transgenic *Arabidopsis* plants when compared to wild-type plants.

Results

Generation of chimeric *UGT72E2/3* and *UGT72E3/2* by domain exchanging between parental *UGT72E2* and *UGT72E3*

Although the *UGT72E2* and *UGT72E3* of *A. thaliana* display 85% sequence identity and 91% similarity at the amino acid level, they exhibit distinct enzymatic characteristics (Lim et al., 2001). *UGT72E2* displays both high specificity for its coniferyl alcohol (**3**) substrate and a high catalytic activity, while *UGT72E3* displays high specificity for its sinapyl alcohol (**1**) substrate and a low catalytic activity. To generate a novel chimeric UGT for the efficient production of syringin (**2**) in plants, a partial domain swapping strategy

was used. Initially, predicted secondary structures for these enzymes with the SWISS-MODEL workspace were obtained (Schwede et al., 2003; <http://swissmodel.expasy.org>) and these were compared (Fig. 2A), and found to be very similar. The main differences were observed in the N β 3–N α 3, C β 3a–C β 3b, and C α 8 regions. Focus was next on a portion of the C-terminal domain (amino acid residues 340–481) that contains the PSPG motif responsible for stabilization of UDP-glucose in the donor binding pocket. The exchange of this sequence between *UGT72E2* and *UGT72E3* resulted in the chimeric enzymes *UGT72E2/3* and *UGT72E3/2* (Fig. 2A). Using the SWISS-MODEL Workspace, the predicted three-dimensional (3D) protein structures of four different UGTs: *72E2*, *72E3*, *72E2/3*, and *72E3/2* (Fig. 2B) were compared. The template used in the 3D structure modeling process, *Arabidopsis* *UGT72B1* (PDB ID code: 2VCE; Brazier-Hicks et al., 2007), is expected to provide reliable predicted 3D structures, because it belongs to the same phylogenetic group (Group E) as the *UGT72E* family with 40% amino acid sequence identity to both *UGT72E2* and *UGT72E3* (Li et al., 2001). Slight differences in the predicted secondary structures of *UGT72E2* and *UGT72E3* were reflected in the predicted 3D structures, and the chimeric enzymes *UGT72E2/3* and *UGT72E3/2* exhibited similar structures to the parental enzymes *UGT72E2* and *UGT72E3* in both N- and C-terminal domains (Fig. 2B).

Enzymatic activities of parental *UGT72E2* and *UGT72E3* and chimeric *UGT72E2/3* and *UGT72E3/2* in transgenic *Arabidopsis* plants

Full length ORFs encoding *UGT72E2*, *UGT72E3*, *UGT72E2/3*, and *UGT72E3/2* were individually overexpressed in *Arabidopsis* under the control of a super promoter (Lee et al., 2007). Independent transgenic plants were subjected to RT-PCR analysis to investigate the steady-state transcript levels for *UGT* transgenes. Representative transgenic plants with equivalent expression levels were selected for each construct and used for further experiments (Fig. 3A).

For qualitative analysis of the enzymatic activities encoded by each *UGT* transgene, leaf color of the transgenic plants under long wavelength (366 nm) UV light was examined (Fig. 3B). The leaves of the transgenic plants were redder than that of wild-type plants, because overexpression of *UGTs* leads to a depletion of sinapoyl malate that absorbs UV light in wild-type plants (Lanot et al., 2008). This result suggests that, like the parental *UGT72E2* and *UGT72E3* enzymes, the chimeric *UGT72E2/3* and *UGT72E3/2* enzymes may change the profiles of soluble phenolic metabolites in leaves.

High performance liquid chromatography (HPLC) was used to perform a quantitative analysis of monolignol glycosylation reactions in the leaves of the transgenic plants. In accordance with previous reports, the parental-type *UGT72E2* and *UGT72E3* OE plants showed a different substrate specificity (Lanot et al., 2006, 2008).

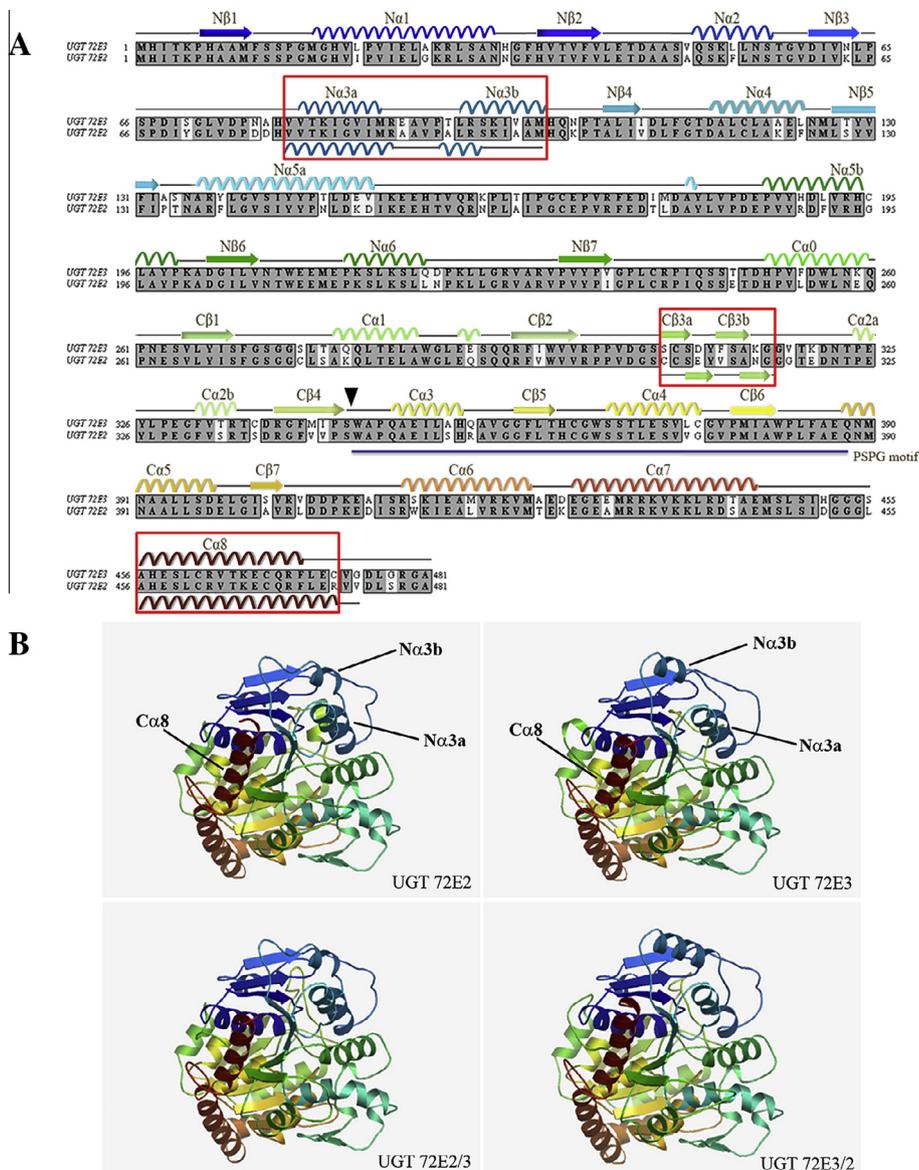


Fig. 2. Amino acid sequence alignment and structure predictions for UGT72E, UGT72E3, and their recombinants. (A) Amino acid sequence alignment and secondary structure prediction for UGT72E2 and UGT72E3. Sequence identity (85.7%) is shaded dark gray, and similarity (91.1%) is shaded light gray. The diagnostic PSPG motif of plant family 1 glucosyltransferases is underlined. The exchange site for recombinant UGT72E2/3 and UGT72E3/2 is indicated with an arrowhead. The secondary structures of UGT72E2 and UGT72E3 were predicted using the SWISS-MODEL workspace. The structures are almost identical, but differences are marked with boxes. A β -strand and an α -helix structure are illustrated by an arrow and a wavy line, respectively. (B) The 3D structures of UGT72E2, UGT72E3 and their recombinant UGTs, UGT72E2/3 and UGT72E3/2, were predicted with the SWISS-MODEL workspace using UGT72B1 as a template. Each protein was diagrammed with color-ramped way from the N (blue) to the C terminus (red). Regions where UGT72E2 and UGT72E3 display different secondary structures are indicated.

The ratio between syringin (**2**) and coniferin (**4**) concentrations was used as the specificity constant (C_s/C_c) to determine the preferred substrate for each transgenic plant. The C_s/C_c values in the leaves of UGT72E2 and UGT72E3 OE plants were 0.22 and 3.30, respectively, whereas the C_s/C_c value was 1.47 in wild-type plants (Fig. 3C). These results suggest that the preferred substrates of UGT72E2 and UGT72E3 are coniferyl alcohol (**3**) and sinapyl alcohol (**1**), respectively. These transgenic plants also showed significant differences in coniferin (**4**) or syringin (**2**) production. The leaves of UGT72E2 OE plants accumulated large amounts of coniferin (**4**) ($4.43 \pm 0.59 \mu\text{mol coniferin/g DW}$), but UGT72E3 OE plants contained a relatively small amount of syringin (**2**) ($1.19 \pm 0.19 \mu\text{mol syringin/g DW}$) (Fig. 3C). Interestingly, the chimeric UGT72E2/3 ($C_s/C_c = 0.25$) and UGT72E3/2 ($C_s/C_c = 2.19$) enzymes had the same preferred substrates as the parental-type UGT72E2 and UGT72E3. These results indicate that amino acid

residues 1–339, which comprise the N-terminal domain and part of the C-terminal domain of the UGT enzyme play a critical role in determining the preferred acceptor specificity for the glycosylation reaction.

However, significant changes in coniferin (**4**) or syringin (**2**) production were observed in the chimeric UGT OE plants. The accumulation of coniferin (**4**) ($3.50 \pm 0.48 \mu\text{mol coniferin/g DW}$) in the UGT72E2/3 OE plants was reduced by as much as 79%, when compared to the parental-type UGT72E2 OE plants. In contrast, the accumulation of syringin (**2**) ($1.77 \pm 0.31 \mu\text{mol syringin/g DW}$) in the UGT72E3/2 OE plants was increased by as much as 145% when compared to the parental-type UGT72E3 OE plants (Fig. 3C).

When cultured in liquid medium under light, the roots of wild-type plants accumulated large amounts of syringin (**2**) ($8.49 \pm 0.44 \mu\text{mol syringin/g DW}$) and coniferin (**4**) ($22.71 \pm 4.19 \mu\text{mol coniferin/g DW}$). This accumulation is due to activation

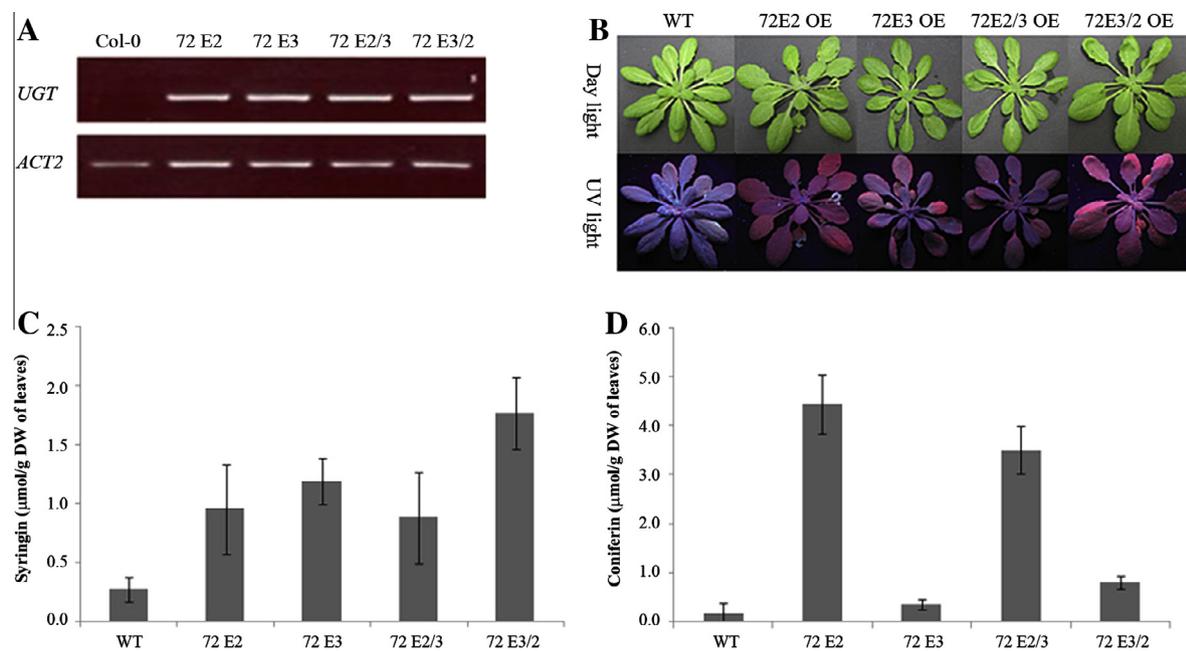


Fig. 3. Characterizations of transgenic plants overexpressing *UGT72E2*, *UGT72E3*, and their recombinants. (A) Semi-quantitative reverse transcriptase-mediated PCR analysis of RNA extracted from 4-week-old wild-type and *UGT*-overexpressing transgenic plants. The actin 2 gene (*ACT2*) was used as a loading control. (B) Fluorescence of *UGT*-overexpressing transgenic plants. Four-week-old plants were photographed under daylight (top row) or ultra-violet illumination (bottom row). Quantified (C) syringin (2) and (D) coniferin (4) levels in 3-week-old leaves of *UGT*-overexpressing transgenic plants with HPLC.

of the phenylpropanoid pathway by light (Hemm et al., 2004). However, the effect of overexpression of the chimeric *UGTs* on the production of coniferin (4) and syringin (2) in the roots was similar to that observed in leaves. For instance, coniferin (4) production in the roots of *UGT72E2/3* OE plants was decreased by as much as 87% of that of the *UGT72E2* OE plants, and the syringin (2) production in the roots of *UGT72E3/2* OE plants was increased by as much as 120% of that of the *UGT72E3* OE plants (Fig. S1).

In vitro enzyme assay of crude protein extracts from homozygous transgenic plants overexpressing *UGT72E* family enzymes

Although the *UGT72E3/2* OE plants accumulated more syringin (2) in leaves than the *UGT72E3* OE plants, its concentration was much lower than that observed in the roots of wild-type plants cultured under continuous light. Because exposure of roots to continuous light activates the phenylpropanoid pathway to generate a larger pool of monolignols (Hemm et al., 2004), it can be hypothesized that lack of proper substrate for *UGT72E3/2* reaction resulted in a relatively low syringin (2) production in the leaves of the transgenic plants. Thus, *in vitro* enzyme activity of crude protein extracts from each transgenic plant was investigated in the presence of high concentrations of substrates to determine the effect of *UGT72E* family enzymes on coniferin (4) and syringin (2) production. Extracts from wild-type plants showed an insignificant increase in coniferin (4) and syringin (2) concentrations for the 60-min reaction, suggesting that endogenous *UGT72E2* and *UGT72E3* activities are negligible (Fig. 4). As in the *in vivo* assays, the extracts of *UGT72E2* OE and *UGT72E3* OE plants showed strong preferences for coniferyl alcohol (3) and sinapyl alcohol (1), respectively, resulting in different coniferin (4) and syringin (2) production levels. These extracts, however, showed only slight increases in coniferin (4) and syringin (2) production in the presence of high concentrations of the corresponding substrates during reactions (Fig. 4). Interestingly, the *UGT72E3/2* OE plant extract displayed 4-fold higher syringin (2) production than the *UGT72E3* OE plant extract in the presence of high concentration of sinapyl alcohol

(2), but no similar increase in coniferin (4) was seen with coniferyl alcohol (3) (Fig. 4). These results suggest that an increase in substrate and enzyme concentrations may enhance syringin (2) production in these plants.

Enhancement of the syringin (2) biosynthesis pathway using metabolic engineering

Overexpression of either *HCT* or *F5H*, which regulate metabolic flux towards sinapyl alcohol (1), was not enough to produce syringin (2) efficiently in the leaves of transgenic plants (data not shown and Fig. 5). Although overexpression of *UGT72E3/2* in combination with *F5H* led to increased syringin (2) production in the leaves of transgenic plants, the *UGT72E3/2*+*F5H* OE plants accumulated only 1.7-fold more syringin (2) than the *UGT72E3/2* OE plants (Fig. 5A). This result is consistent with previous reports that overexpression of *F5H* alone is not enough to direct metabolic flux towards sinapyl alcohol (1) in *Arabidopsis* (Stewart et al., 2009). To explore the full potential for syringin (2) production in *Arabidopsis*, the overall activity of the phenylpropanoid pathway needed to be enhanced in combination with the overexpression of *UGT72E3/2* and *F5H*. Thus, the transcriptional activator *MYB58*, which positively regulates the expression of most lignin biosynthesis genes except *F5H*, in the *UGT72E3/2*+*F5H* OE plants was overexpressed. The *72E3/2*+*F5H*+*MYB58* OE plants stacking three transgenes had a synergistic effect on syringin (2) production in leaves, resulting in 56-fold more syringin (2) accumulation in the transgenic plants than wild-type plants (Fig. 5A). In contrast, coniferin (4) production was higher in the *72E3/2*+*MYB58* OE plants than in the *72E3/2*+*F5H*+*MYB58* OE plants. This result suggests that overexpression of the *F5H* transgene causes a depletion of the metabolic source of coniferyl alcohol (3) (Fig. 5B). Consistent with the drastic effect of *MYB58* overexpression on syringin (2) production, most of the genes regulating early reaction steps in the phenylpropanoid pathway were highly expressed in the *72E3/2*+*F5H*+*MYB58* OE plants (Fig. 6). *PAL*, *C4H*, *4CL1*, *HCT*, *C3H*, and *CCoAOMT1* genes have

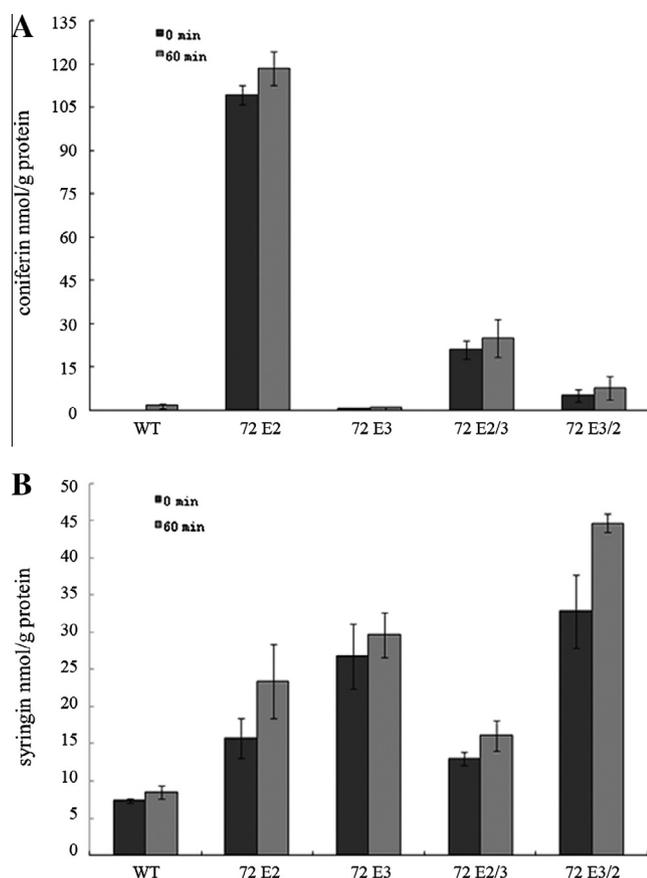


Fig. 4. Syringin (2) or coniferin (4) production by crude protein extracts from *UGT*-overexpressing transgenic plants in the presence of coniferyl alcohol (3) or sinapyl alcohol (1). Effects of coniferyl alcohol (3) and sinapyl alcohol (1) supplements on coniferin (4) and syringin (2) production in the protein extracts of transgenic plants overexpressing *UGT72E2*, *E3*, *E2/3*, and *E3/2* were investigated following the method described in section 'Glucosyltransferase activity assay with crude protein extract from transgenic plants'. Syringin (2) and coniferin (4) production levels were defined as nmol of generated glucosides per 1 g protein before and after reaction for 60 min. (A) Generated coniferin (4) and (B) syringin (2) levels in the *in vitro* reactions. Results are represented as mean \pm S.D. from at least three replicates.

increased expression and *CCR2*, *COMT*, and *CAD4* genes have normal high expression.

Discussion

Recent structural data for several plant UGTs clearly demonstrate that these proteins have two functionally distinct domains. The acceptor binding site is mainly formed by residues in the N-terminal domain, and the donor binding site is formed in the C-terminal domain and flanked on one side by a PSPG motif (Paquette et al., 2003). These domains pack very tightly by inter- and intra-domain interactions, resulting in formation of a deep, narrow catalytic site where the acceptor and donor binding sites come close enough together to catalyze the glucosyltransfer reaction at the interface (Osmani et al., 2009). Using these function-structure relationships of UGTs, a partial domain swapping between *UGT72E2* and *UGT72E3* was performed to generate a novel UGT that efficiently catalyzes syringin (2) production via the transfer of glucose from UDP-glucose to the C4-OH position of sinapyl alcohol (1).

The successful engineering of chimeric enzymes is most likely to occur when the parent enzymes have a high sequence identity and high conservation of intra- and inter-domain interactions (Osmani et al., 2009). For example, the chimeric enzyme *UGTN1C3*,

whose parent enzymes share 81% sequence similarity, is constructed by combining the entire N-terminal domain of *AtUGT71C1* with the entire C-terminal domain of *AtUGT71C3*. The acceptor specificity of the chimeric enzyme is similar to *AtUGT71C1*, while the donor specificity of the chimeric enzyme matches *AtUGT71C3*. Its catalytic activity is between that of *AtUGT71C3* and *AtUGT71C1* (Weis et al., 2008). Another chimeric UGT was described in which the entire N-terminal domain of *AtUGT74F2* was fused to the entire C-terminal domain of *AtUGT74F1*. The parent enzymes for this chimera share 90% sequence similarity. The chimeric UGT displays similar acceptor specificity to the parent enzyme *AtUGT74F2*, but its catalytic activity is significantly decreased relative to the parent enzyme *AtUGT74F1* (Cartwright et al., 2008). These results indicate that the N- and C-terminal domains can be substituted to alter acceptor and donor specificities in chimeric enzyme, but these specific recognitions are not sufficient for efficient catalytic reactions. The overall fit of acceptor and donor into their binding pockets as well as exact positioning and stabilization of acceptor and donor at the interface between domains play an essential role in the glucosyltransfer reaction from donor to acceptor (Osmani et al., 2009). Thus, a minimal C-terminal region containing the PSPG motif between the parent enzymes *UGT72E2* and *UGT72E3* that share 91% sequence similarity enhanced the catalytic activity of *UGT72E3* towards sinapyl alcohol (1), leading to higher production of syringin (2). Based on established substrate preferences of *UGT72E* enzymes (Lanot et al., 2006; Lim et al., 2001), transgenic plants were generated expressing the parental genes *UGT72E2* and *UGT72E3* and the chimeric genes *UGT72E2/3* and *UGT72E3/2* to compare their syringin (2) and coniferin (4) productions *in planta*.

As expected from their predicted 3D structures (Fig. 2B), chimeric *UGT72E2/3* and *UGT72E3/2* retained the substrate specificity of the parental enzymes *UGT72E2* and *UGT72E3*. In contrast, the syringin (2) or coniferin (4) production of these chimeric enzymes towards their specific substrates was significantly changed; the leaves of the *UGT72E2/3* OE plants accumulated 21% less coniferin (4) than its parental *UGT72E2* OE plants, while those of *UGT72E3/2* OE plants accumulated 45% more syringin (2) than *UGT72E3* OE plants (Fig. 3C). These results suggest that the substitution of a partial C-terminal region containing the PSPG motif between highly homologous but catalytically different UGTs is a good strategy to generate a novel chimeric UGTs altering the catalytic activity without disturbing the acceptor specificity. The C-terminal 30% of the *UGT72E2* sequence including the PSPG motif contributes to the determination of donor specificity and stabilization of the donor in the donor binding pocket, meanwhile the N-terminal 70% of the *UGT72E3* sequence including the entire N-terminal region, the inter-domain linker, and part of the C-terminal region determines acceptor specificity and the correct positioning of acceptor and donor at the interface between domains (Osmani et al., 2009). The kinetic data of *UGT72E* enzymes and their chimeras combined with site directed mutagenesis may enable identification of relevant amino acids.

Although the chimeric *UGT72E3/2* increased syringin (2) production in the leaves of *UGT72E3/2* OE plants, it was not sufficient to produce it in plants effectively. To find the other factors affecting syringin (2) production, an *in vitro* assay was performed using crude protein extracts from homozygous transgenic plants overexpressing the *UGT72E* gene family. In contrast to *UGT72E3*, *UGT72E3/2* was able to increase syringin (2) production significantly in the presence of high concentration of sinapyl alcohol (1), indicating that the limited production of syringin (2) in the leaves of the *UGT72E3/2* OE plant was due to a lack of sinapyl alcohol (1) (Fig. 4). In leaves of *Arabidopsis* plants, most sinapyl compounds are converted to sinapoyl malate to protect the plant

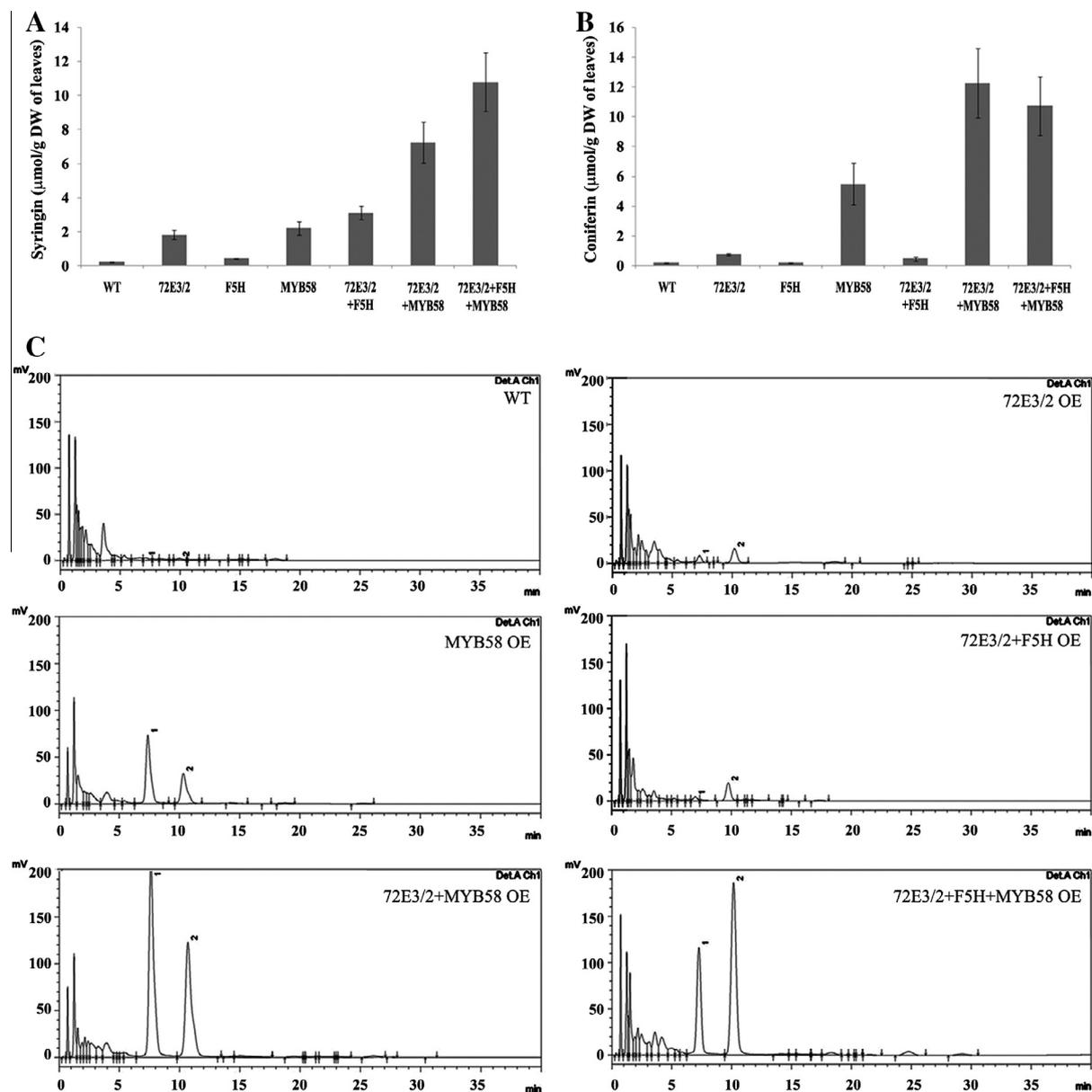


Fig. 5. Accumulation of monolignol glucosides in the leaves of wild-type and transgenic plants overexpressing different combinations of *UGT72E3/2*, *F5H*, and *MYB58*. (A) Syringin (2) and (B) coniferin (4) levels in 3-week-old leaves of wild-type and transgenic plants overexpressing different combinations of *UGT72E3/2*, *F5H*, and *MYB58* and (C) their selected HPLC chromatograms at 268 nm. Peaks: 1, coniferin; 2, syringin. Results are represented as mean \pm S.D. from at least three replicates.

from UV radiation (Kusano et al., 2011). Therefore, it is important to enhance the metabolic flux towards sinapyl alcohol (1) in combination with overexpression of *UGT72E3/2* when designing transgenic plants that accumulate large amounts of syringin (2) in their leaves.

To supply abundant sinapyl alcohol (1) as a substrate for syringin (2) production, the phenylpropanoid pathway was engineered to increase metabolic flux towards sinapyl alcohol (1) by overexpressing both *F5H* and *MYB58* in the *UGT72E3/2* OE plants. When stacking transgenes in plants, the use of binary vectors carrying the CaMV 35S promoter resulted in a severe cosuppression of transgenes (data not shown). Because small RNA-mediated cosuppression has evolved in plants to defend against viral invasion, vectors that stack multiple viral sequences, particularly promoter sequences, are very susceptible to cosuppression (Vaucheret et al., 2001). To avoid cosuppression, a novel plant transformation vector series was used that contains a super promoter developed

by the Gelvin research group (Lee et al., 2007). This vector series contains no CaMV 35S promoter. Successful stacking of three different transgenes in the *UGT72E3/2*+*F5H*+*MYB58* OE plant was achieved without severe cosuppression (Fig. 6). The *UGT72E3/2*+*F5H*+*MYB58* OE plants not only accumulated more syringin (2) than transgenic plants overexpressing *UGT72E3/2* or *MYB58* alone, but also accumulated even more of it in their leaves than the transgenic plants that stacked *UGT72E3/2* and *MYB58* together (Fig. 5). In the light-grown roots of the *UGT72E3/2*+*F5H*+*MYB58* OE plants, both syringin (2) and coniferin (4) production were similar to wild-type plants (Fig. S2), suggesting that the accumulation of large amounts of syringin (2) in the leaves of *UGT72E3/2*+*F5H*+*MYB58* OE plants was not caused by a simple redirection of metabolic flux from the roots to the leaves (Lanot et al., 2008). Instead, it appears that in the *72E3/2*+*F5H*+*MYB58* OE plants, *MYB58*-induced expression of genes regulating the phenylpropanoid pathway combines with overexpression of the *F5H* transgene to direct metabolic flux

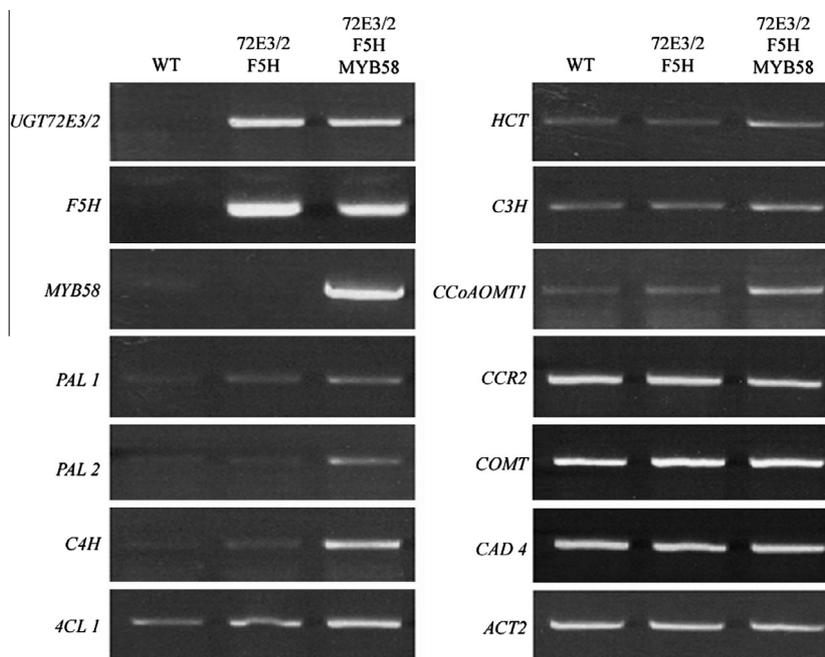


Fig. 6. Phenylpropanoid pathway genes regulated by overexpression of *MYB58* in the UGT72E3/2+F5H+MYB58 OE plants. Leaves from 4-week-old plants were used for expression analyses of phenylpropanoid pathway genes and the three stacked transgenes (see Abbreviations for full names of genes). Analysis was conducted using semi-quantitative reverse transcriptase-mediated PCR with gene specific primers (see section 'Reverse transcriptase-mediated RNA analysis'). The actin 2 gene (*ACT2*) was used as a loading control.

towards sinapyl alcohol (**1**). The resulting pool of sinapyl alcohol (**1**) might be efficiently glycosylated by the chimeric UGT72E3/2 enzyme. Taken together, these results demonstrate that the synergistic effect of combined overexpression of *UGT72E3/2*, *F5H*, and *MYB58* is a mechanism for efficient production of syringin (**2**) in plant leaves.

Conclusions

Although the syringyl-lignin biosynthetic pathway is highly conserved in the plant kingdom, only a few plant species, such as *Acanthopanax senticosus*, accumulate large amounts of syringin (**2**). In this study, a practical method was established for efficient production of syringin (**2**) in leaves of *A. thaliana*. The novel recombinant enzyme UGT72E3/2 was first generated by exchanging a partial C-terminal domain, which contains the PSPG motif that stabilizes UDP-glucose in the donor pocket, between UGT72E2 and UGT72E3. The leaves of UGT72E3/2 OE plants accumulated 45% more syringin (**2**) than UGT72E3 OE plants. To further explore the full potential of the UGT72E3/2 OE transgenic plants, metabolic flux was enhanced through the phenylpropanoid pathway towards sinapyl alcohol (**1**) for syringin (**2**) production by overexpressing *MYB58* and *F5H* together, resulting in 56-fold higher syringin (**2**) content compared to wild-type plants. These results demonstrated that the synergistic effect of improved syringin (**2**) production of the chimeric UGT72E3/2 and enhanced metabolic flux through the phenylpropanoid pathway is a mechanism for its efficient production in plant leaves.

Experimental

Plant materials and growth conditions

Wild-type *A. thaliana* used in the experiments was ecotype Columbia. For all experiments with plants, seeds were surface-sterilized, germinated, and grown on plates containing solid

Gamborg's B5 medium supplemented with or without antibiotics (kanamycin 50 $\mu\text{g ml}^{-1}$, hygromycin 20 $\mu\text{g ml}^{-1}$, phosphinothricin 10 $\mu\text{g ml}^{-1}$) for selection. All transgenic plants used in experiments were homozygous lines. The replicates of each sample were made from different plants in the same transgenic lines. For analyses of gene expression and crude plant extract glucosyltransferase activity, seedlings were transferred to soil and grown in a controlled environment (14/10 h light/dark cycle, 22 °C) for 4 weeks. For target metabolite analysis by HPLC, seedlings were transferred into flasks containing liquid Gamborg's B5 medium and aseptically cultured under continuous light at 22 °C for 3 weeks.

Binary vector constructs and generation of transgenic plants

Full length coding regions for *UGT72E2* (At5g66690) and *UGT72E3* (At5g26310) were amplified from *A. thaliana* genomic DNA with Pyrobest DNA Polymerase (*TaKaRa Taq*TM) using the primer combinations 72E2 For/72E2 Rev and 72E3 For/72E3 Rev, respectively. PCR cycle conditions were 92 °C for 30 s, 56 °C for 60 s, and 72 °C for 60 s with 35 cycles. Recombinant *UGT72E2/3* and *UGT72E3/2* genes were constructed using an overlapping extension PCR method (Wurch et al., 1998). Briefly, 1020 and 424 bp DNA fragments spanning the carboxyl terminal region and the amino terminal region of each gene were PCR-amplified with the primer combinations 72E2 For/72E Rev, 72E For/72E2 Rev, 72E3 For/72E Rev, and 72E For/72E3 Rev, respectively. Taking advantage of the overlapping sequences of 72E For/72E Rev, DNA fragments from *UGT72E2* and *UGT72E3* was reciprocally exchanged and PCR-ligated in frame with the primer combinations 72E2 For/72E3 Rev and Rev 72E3 For/72E Rev. Similarly, full length coding regions of *MYB58* (At1g16490), and *F5H* (At4g36220) were amplified from an *A. thaliana* cDNA library generated from seedlings with the primer combinations MYB58 For/MYB58 Rev, F5H For/F5H Rev, respectively. Sequence information for all primers used in the construction of vectors is provided in Table S1. All PCR products were cloned into pBluescriptIIKS, sequenced, excised

with appropriate restriction enzymes, and cloned into the same sites of the super promoter binary vector series; pE1801, pE1803, and pE1813 were used for plant transformation with three different selection markers (Lee et al., 2007). *Arabidopsis* plants were transformed by standard *Agrobacterium tumefaciens* methods (Clough and Bent, 1998), and primary transformants were selected with antibiotics according to the binary vector used. Transgene stacking was performed by standard genetic crosses between transgenic plants overexpressing different transgenes. Plants that were homozygous for two or three different transgenes were isolated from the F3 generation based on selection markers.

Reverse transcriptase-mediated RNA analysis

Total RNA was isolated from 4-week-old rosette leaves of *Arabidopsis* using QIAzol[®] Lysis Reagent (QIAGEN). The RNA was treated with DNase I (Invitrogen) for 30 min to remove any potential DNA contamination before the reverse transcription reaction. Synthesis of first-strand cDNA and semi-quantitative RT-PCR amplification was performed with the SuperScript[™] III One-Step RT-PCR System (Invitrogen) according to the manufacturer's instructions. To ensure transgene expression in transgenic plants, a combination of a gene specific primer and a 3'-UTR specific primer of vector was used for each gene. The primer for actin was used as a reference. Sequence information for all primers used for RNA analysis is provided in the Table S1.

Glucosyltransferase activity assay with crude protein extract from transgenic plants

Four-week-old leaves were harvested and frozen in liq. N₂. Frozen tissue (1 g) was ground to a fine powder in liq. N₂ using a mortar and pestle. Extraction buffer 1 ml (25 mM Tris-MES, pH 6.5, 10% (v/v) glycerol, 20 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride and 1% polyvinylpyrrolidone) (w/w of tissue) was added to the powder, and the slurry was thawed on ice. Subsequently, the slurry was mixed vigorously and centrifuged at 16,000×g at 4 °C for 20 min. The supernatant was collected, transferred to a 1.5 ml microcentrifuge tube, and further centrifuged at 16,000×g at 4 °C for 5 min. The resulting supernatant was collected for the glucosyltransferase activity assay (Lanot et al., 2006). Total protein concentration was determined using a protein assay kit (Bio-Rad Laboratories) with bovine serum albumin serving as a standard. The assay mix (200 μl) contained crude proteins 200 μg, 14 mM β-mercaptoethanol, 5 mM UDP-glucose, 1 mM phenylpropanoid substrate (coniferyl alcohol (3) or sinapyl alcohol (1)), and 100 mM Tris-HCl. For glucosyltransferase activity analysis, reactions leading to 4-O-glucosylation were carried out at 22 °C for 60 min. Reactions were stopped by the addition of MeOH 400 μl, quick-frozen, and stored at –20 °C prior to reversed-phase HPLC analysis.

HPLC analysis of target metabolites

To extract syringin (2) and coniferin (4) accumulated in plants, roots and aerial tissues were separately harvested from plants grown in liquid medium under sterile conditions for 3 weeks. Plant tissues were dried at 60 °C for 16 h and ground to a fine powder using a mortar and pestle. Dry tissue powder (80 mg for aerial tissues or 60 mg for root tissues) was extracted with MeOH (1.5 ml) at room temperature for 14 h. After filtration through a 0.45 μm filter, the supernatant was collected, and was used for reversed-phase HPLC analysis. An aliquot (10 μl) of the sample was injected with auto-injector for analysis. Reversed-phase HPLC (LC-20AD Liquid Chromatograph system and SPD-20A UV/VIS Detector; SHIMADZU) analysis was carried out using a Columbus

5 μm C18 column (150 × 4.60 mm; PerkinElmer) maintained at 25 °C with MeOH-H₂O (20:80, v/v) at a flow rate of 1.0 ml/min. Each peak in the chromatogram was scanned at 210 and 268 nm. Syringin (2) and coniferin (4) were identified by comparison to known standards (CAS No.531-29-3 and CAS No.118-34-3, National Institutes for Food and Drug Control, China). The acquired data were analyzed by SHIMADZU LC solution software. Quantitation was based on ultraviolet absorption at 268 nm. The peak areas were converted into mass measurements by comparison with an external standard calibration curve.

Acknowledgements

We thank Stanton B. Gelvin, Purdue University, for providing the super promoter vector series. This work was supported by a Grant awarded to J. Nam from Biogreen 21 (PJ008167) and by a scholarship awarded to Y. Chu from the BK21 program.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.phytochem.2014.03.003>.

References

- Anterola, A.M., Lewis, N.G., 2002. Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity. *Phytochemistry* 61, 221–294.
- Brazier-Hicks, M., Offen, W.A., Gershater, M.C., Revett, T.J., Lim, E.H., Bowles, D.J., Davies, G.J., Edwards, R., 2007. Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. *Proc. Natl. Acad. Sci. USA* 104, 20238–20243.
- Cartwright, A.M., Lim, E.K., Kleantous, C., Bowles, D.J., 2008. A kinetic analysis of regio-specific glucosylation by two glucosyltransferases of *Arabidopsis thaliana*: domain swapping to introduce new activities. *J. Biol. Chem.* 283, 15724–15731.
- Cho, J.Y., Nam, K.H., Kim, A.R., Park, J., Yoo, E.S., Baik, K.U., Yu, Y.H., Park, M.H., 2001. In-vitro and in-vivo immunomodulatory effects of syringin. *J. Pharm. Pharmacol.* 53, 1287–1294.
- Choi, J., Shin, K.M., Park, H.J., Jung, H.J., Kim, H.J., Lee, Y.S., Rew, J.H., Lee, K.T., 2004. Anti-inflammatory and antinociceptive effects of sinapyl alcohol and its glucoside syringin. *Planta Med.* 11, 1027–1032.
- Clough, S.J., Bent, A.F., 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743.
- Hemm, M.R., Rider, S.D., Ogas, J., Murry, D.J., Chapple, C., 2004. Light induces phenylpropanoid metabolism in *Arabidopsis* roots. *Plant J.* 38, 765–778.
- Kusano, M., Tohge, T., Fukushima, A., Kobayashi, M., Hayashi, N., Otsuki, H., Kondou, Y., Goto, H., Kawashima, M., Matsuda, F., Niida, R., Matsui, M., Saito, K., Fernie, A.R., 2011. Metabolomics reveals comprehensive reprogramming involving two independent metabolic responses of *Arabidopsis* to UV-B light. *Plant J.* 67, 354–369.
- Lanot, A., Hodge, D., Jackson, R.G., George, G.L., Elias, L., Lim, E.K., Vaistij, F.E., Bowles, D.J., 2006. The glucosyltransferase UGT72E2 is responsible for monolignol 4-O-glucoside production in *Arabidopsis thaliana*. *Plant J.* 48, 286–295.
- Lanot, A., Hodge, D., Lim, E.K., Vaistij, F.E., Bowles, D.J., 2008. Redirection of flux through the phenylpropanoid pathway by increased glucosylation of soluble intermediates. *Planta* 228, 609–616.
- Lee, L.Y., Kononov, M.E., Bassuner, B., Frame, B.R., Wang, K., Gelvin, S.B., 2007. Novel plant transformation vectors containing the superpromoter. *Plant Physiol.* 145, 1294–1300.
- Li, Y., Baldauf, S., Lim, E.K., Bowles, D.J., 2001. Phylogenetic analysis of the UDP-glucosyltransferase multigene family of *Arabidopsis thaliana*. *Biol. Chem.* 276, 4338–4343.
- Lim, E.K., Li, Y., Parr, A., Jackson, R., Ashford, D.A., Bowles, D.J., 2001. Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in *Arabidopsis*. *J. Biol. Chem.* 276, 4344–4349.
- Liu, K.Y., Wu, Y.C., Liu, I.M., Yu, W.C., Cheng, J.T., 2008. Release of acetylcholine by syringin, an active principle of *Eleutherococcus senticosus*, to raise insulin secretion in Wistar rats. *Neurosci. Lett.* 434, 195–199.
- Niu, H.S., Liu, I.M., Cheng, J.T., Lin, C.L., Hsu, F.L., 2008. Hypoglycemic effect of syringin from *Eleutherococcus senticosus* in streptozotocin-induced diabetic rats. *Planta Med.* 74, 109–113.
- Osmani, S.A., Bak, S., Møller, B.L., 2009. Substrate specificity of plant UDP-dependent glucosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry* 70, 325–347.
- Paquette, S., Møller, B.L., Bak, S., 2003. On the origin of family 1 plant glucosyltransferases. *Phytochemistry* 62, 399–413.
- Schwede, T., Kopp, J., Guex, N., Peisch, M., 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acid Res.* 31, 3381–3385.

- Stewart, J.J., Akiyama, T., Chapple, C., Ralph, J., Mansfield, S.D., 2009. The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol.* 150, 621–635.
- Vanholme, R., Morreel, K., Ralph, J., Boerjan, W., 2008. Lignin engineering. *Curr. Opin. Plant Biol.* 11, 278–285.
- Vaucheret, H., Béclin, C., Fagard, M., 2001. Post-transcriptional gene silencing in plants. *J. Cell Sci.* 114, 3083–3091.
- Weis, M., Lim, E.K., Bruce, N.C., Bowles, D.J., 2008. Engineering and kinetic characterization of two glucosyltransferases from *Arabidopsis thaliana*. *Biochimie* 90, 830–834.
- Whetten, R., Sederoff, R., 1995. Lignin biosynthesis. *Plant Cell* 7, 1101–1013.
- Wurch, T., Lestienne, F., Pauwels, P.J., 1998. A modified overlap extension PCR method to create chimeric genes in the absence of restriction enzyme. *Biotechnol. Tech.* 12, 653–657.
- Zhao, Q., Wang, H., Yin, Y., Xu, Y., Chen, F., Dixon, R.A., 2010. Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. *Proc. Natl. Acad. Sci. USA* 107, 14496–14501.
- Zhou, J., Lee, C., Zhong, R., Ye, Z.-H., 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell* 21, 248–266.