PCP
PLANT & CELL PHYSIOLOGY

Regular Paper

Lotus japonicus Triterpenoid Profile and Characterization of the *CYP716A51* and *LjCYP93E1* Genes Involved in Their Biosynthesis In Planta

Hayato Suzuki[1], Ery Odette Fukushima[1,2], Yuko Shimizu[1], Hikaru Seki[1], Yukiko Fujisawa[3], Masao Ishimoto[3], Keishi Osakabe[4], Yuriko Osakabe[4] and Toshiya Muranaka[1,*]

[1]Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka, 565-0871 Japan
[2]Universidad Regional Amazónica IKIAM, Km 7 Via Muyuna, Napo, Tena, Ecuador
[3]Institute of Crop Science, NARO, 2-1-2 Kannondai, Tsukuba, Ibaraki, 305-8518 Japan
[4]Division of Bioscience and Bioindustry, Graduate School of Technology, Industrial and Social Sciences, Tokushima University, Tokushima, Japan
*Corresponding author: E-mail, muranaka@bio.eng.osaka-u.ac.jp; Fax, +81-6-6879-7426.
(Received March 10, 2019; Accepted July 12, 2019)

Lotus japonicus **is an important model legume plant in several fields of research, such as secondary (specialized) metabolism and symbiotic nodulation. This plant accumulates triterpenoids; however, less information regarding its composition, content and biosynthesis is available compared with *Medicago truncatula* and *Glycine max*. In this study, we analyzed the triterpenoid content and composition of *L. japonicus*. *Lotus japonicus* accumulated C-28-oxidized triterpenoids (ursolic, betulinic and oleanolic acids) and soyasapogenols (soyasapogenol B, A and E) in a tissue-dependent manner. We identified an oxidosqualene cyclase (OSC) and two cytochrome P450 enzymes (P450s) involved in triterpenoid biosynthesis using a yeast heterologous expression system. OSC9 was the first enzyme derived from *L. japonicus* that showed α-amyrin (a precursor of ursolic acid)-producing activity. CYP716A51 showed triterpenoid C-28 oxidation activity. LjCYP93E1 converted β-amyrin into 24-hydroxy-β-amyrin, a metabolic intermediate of soyasapogenols. The involvement of the identified genes in triterpenoid biosynthesis in *L. japonicus* plants was evaluated by quantitative real-time PCR analysis. Furthermore, gene loss-of-function analysis of *CYP716A51* and *LjCYP93E1* was conducted. The *cyp716a51*-mutant *L. japonicus* hairy roots generated by the genome-editing technique produced no C-28 oxidized triterpenoids. Likewise, the complete abolition of soyasapogenols and soyasaponin I was observed in mutant plants harboring *Lotus retrotransposon 1* (*LORE1*) in *LjCYP93E1*. These results indicate that the activities of these P450 enzymes are essential for triterpenoid biosynthesis in *L. japonicus*. This study increases our understanding of triterpenoid biosynthesis in leguminous plants and provides information that will facilitate further studies of the physiological functions of triterpenoids using *L. japonicus*.**

Keywords: Betulinic acid • CRISPR/Cas9 • *LORE1* • Oleanolic acid • Soyasapogenols • Ursolic acid.

Accession numbers: The nucleotide sequence reported in this article has been submitted to the DNA Data Bank of Japan (DDBJ) under the accession numbers of AB706297 (*CYP716A51*), LC414182 (*LjCYP93E1*) and LC485316 (*OSC9*).

Introduction

Triterpenoids are a category of plant-specialized metabolites with vast structural diversity and pharmacological activities (Thimmappa et al. 2014). The first committed step of triterpenoid biosynthesis is the cyclization of 2,3-oxidosqualene by oxidosqualene cyclases (OSCs), which generate simple triterpene skeletons such as α-amyrin, lupeol and β-amyrin. These simple skeletons are modified by cytochrome P450 monooxygenases (CYPs or P450s) and UDP-dependent glycosyltransferases (UGTs), resulting in the production of triterpenoid sapogenins (aglycones) and triterpenoid saponins (glycosides) (Thimmappa et al. 2014, Seki et al. 2015).

Legume plants produce various triterpenoids in large amounts. The model legume *Medicago truncatula* produces hemolytic sapogenins and saponins from β-amyrin (Huhman et al. 2005, Confalonieri et al. 2009). Hemolysis refers to the ability of certain compounds to disrupt red blood cell membranes, which, for triterpenoids, is highly dependent on the backbone and the oxidized carbon positions (Vo et al. 2017). CYP716A12 is a key enzyme involved in hemolytic saponin biosynthesis that catalyzes a three-step oxidation at the C-28 position of β-amyrin to produce oleanolic acid (Carelli et al. 2011, Fukushima et al. 2011). CYP72A68 (Fukushima et al. 2013, Biazzi et al. 2015) and CYP72A67 (Biazzi et al. 2015) produce medicagenic acid through oxidation at the C-23 and C-2β positions, respectively, of oleanolic acid. *Cyp716a12* and *ugt73f3* (a UGT enzyme involved in hemolytic saponin biosynthesis) mutant plants exhibited a drastic dwarf phenotype, suggesting that triterpenoid sapogenins may be toxic or that hemolytic saponins are important for the growth and development of *M. truncatula* plants (Naoumkina et al. 2010, Carelli et al. 2011).

Soyasaponins are widely distributed in legume plants such as soybean (*Glycine max*) and *M. truncatula* (Fenwick et al. 1991, Huhman et al. 2005, Confalonieri et al. 2009, Carelli et al. 2011).

Plant Cell Physiol. 60(11): 2496–2509 (2019) doi:10.1093/pcp/pcz145, Advance Access publication on 16 August 2019, available online at https://academic.oup.com/pcp

Soyasaponins are classified as nonhemolytic saponins, as their aglycones (soyasapogenols B, A and E) do not have hemolytic activity. Some P450 genes are reportedly involved in soyasapogenol biosynthesis in the abovementioned two legume plants. In *G. max*, CYP93E1 and CYP72A61 oxidize β-amyrin at the C-24 and C-22 positions, respectively, to produce soyasapogenol B (Shibuya et al. 2006, Ebizuka et al. 2011, Krishnamurthy et al. 2019). Soyasapogenol A is biosynthesized by oxidation of the C-21 position of soyasapogenol B by CYP72A69 (Sg-5) (Yano et al. 2017). Soyasapogenol E is likely to be produced from soyasapogenol B by further oxidation at the C-22 position; however, the soyasapogenol B C-22 oxidase responsible has not yet been identified. *CYP93E2* and *CYP72A61v2* of *M. truncatula* are orthologs of *CYP93E1* and *CYP72A61* of *G. max*, respectively (Fukushima et al. 2011, Fukushima et al. 2013). The physiological functions of soyasaponins are still under discussion. However, the selective secretion of soyasaponins from the roots of hydroponic-cultured legume plants suggests roles for soyasaponins as chemical signals in the interaction between legume plants and other organisms in the rhizosphere (Tsuno et al. 2018).

Lotus japonicus is a model legume plant. Five OSCs and one P450 were functionally characterized as triterpenoid biosynthetic enzymes in heterologous expression systems and by gene-silencing analyses (Iturbe-Ormaetxe et al. 2003, Sawai et al. 2006a, Sawai et al. 2006b, Delis et al. 2011, Krokida et al. 2013). AMY2 is a multifunctional OSC enzyme that produces dihydro-lupeol, which is subsequently oxidized by CYP71D353 at the C-20 and C-28 positions to produce 20-hydroxy lupeol and 20-hydroxy betulinic acid, respectively, when transiently expressed in *Nicotiana benthamiana* leaves by agroinfiltration (Krokida et al. 2013). Although dihydro-lupeol, 20-hydroxy lupeol and 20-hydroxy betulinic acid were not detected in *L. japonicus* plants, *amy2*-silenced plants showed nonflowering and short and stunted root phenotypes (Krokida et al. 2013). *OSC3* encodes a lupeol synthase (Sawai et al. 2006b). Lupeol accumulation decreased in *osc3*-silenced plants and they showed higher expression of the *ENOD40* gene (an early nodulin gene), resulting in a rapid nodulation phenotype (Delis et al. 2011). This finding suggests a negative effect of lupeol or its derivatives on nodulation (Delis et al. 2011).

Although *L. japonicus*, as well as *M. truncatula* and *G. max*, has been used in studies of the biosynthesis and physiological functions of triterpenoids, the triterpenoid sapogenin profile of this plant has not been analyzed. In addition, no P450 enzyme involved in triterpenoid biosynthesis in planta has been identified. In this study, we qualitatively analyzed and quantified sapogenins in *L. japonicus* for newly detecting some triterpenoids (ursane, lupane and oleanane types) and for revealing their tissue-dependent accumulation. We functionally characterized one OSC (OSC9) and two P450s (CYP716A51 and LjCYP93E1) involved in the production of major triterpenoids in this plant using a yeast heterologous expression system. We also examined the expression patterns of identified triterpenoid biosynthetic genes in *L. japonicus* plants. Finally, we directly evaluated the roles of P450 genes via in planta loss-of-function analyses, using the clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) (Sander and Joung 2014, Wang et al. 2016) technology and *Lotus retrotransposon 1* (*LORE1*)-tagged lines (Fukai et al. 2012, Urbański et al. 2012).

Results

Triterpenoid profiling in *L. japonicus*

Leaves, stems, roots, flowers and fruits were collected from 2.5-month-old *L. japonicus*. Extracted triterpenoids were analyzed by gas chromatography-mass spectrometry (GC-MS) (Supplementary Figs. S1, S2). Quantification of the triterpenoid sapogenins in the acid-hydrolyzed extracts (**Table 1**) shows the total amount of nonglycosylated triterpenoids in plants and aglycones of triterpenoid saponins because sugar moieties were removed from saponins by acid hydrolysis. The chromatograms contain the peaks of simple triterpenoid skeletons, α-amyrin (1), lupeol (5) and β-amyrin (9), and their C-28-oxidized derivatives, ursolic acid (4), betulinic acid (8) and oleanolic acid (12) (**Fig. 1**; **Table 1**). Soyasapogenols [soyasapogenol B (15), soyasapogenol A (16) and soyasapogenol E (17)] and their biosynthetic intermediate [sophoradiol (14)] were also detected (**Fig. 1**; **Table 1**). Although a small peak at the same retention time (Rt) as the authentic standard of 24-hydroxy-β-amyrin (13) appeared in root extracts (Supplementary Fig. S2), we could not identify it as its intensity was too low and it co-eluted with an unknown metabolite.

Among the abovementioned triterpenoid sapogenins, α-amyrin (1), ursolic acid (4), betulinic acid (8), β-amyrin (9), oleanolic acid (12), sophoradiol (14), soyasapogenol A (16) and soyasapogenol E (17) were detected for the first time in *L. japonicus* in this study (**Fig. 1**). Together, our triterpenoid profiling experiments revealed a tissue-dependent accumulation of triterpenoids in *L. japonicus*. In brief, betulinic acid (8) and soyasapogenol B (15) accumulated mainly in roots and ursolic acid (4) and oleanolic acid (12) accumulated mainly in stems (**Table 1**).

To determine whether these triterpenoids are glycosylated or not in planta, we compared the GC-MS total ion chromatograms (TICs) of the extracts of each tissue with and without acid hydrolysis (Supplementary Fig. S1). Soyasapogenols were detected only in the acid-hydrolyzed extracts, whereas C-28 oxidized triterpenoids were also detected in nonhydrolyzed extracts. Furthermore, we quantified the amount of each triterpenoid in roots with and without acid hydrolysis using selected ion monitoring (SIM)-GC-MS analysis (Supplementary Table S1). We detected small amounts of soyasapogenol B (16) and soyasapogenol E (17) without acid hydrolysis. However, the amounts of these compounds in acid-hydrolyzed extracts were 830 times and 78 times higher, respectively, compared with nonhydrolyzed extracts. No soyasapogenol A (16) was detected without acid hydrolysis even in SIM mode. These results suggest that most soyasapogenols are glycosylated in this plant. The C-28 oxidized triterpenoids appear to accumulate as aglycone forms in *L. japonicus*.

Table 1 Quantitation of triterpenoids in 2.5-month-old *L. japonicus* by GC-MS

Triterpenoid (μg/g-dw)	Leaves	Stems	Roots	Flowers	Fruits
α-Amyrin (1)	378.3 ± 8.3	5.1 ± 1.7	1.9 ± 0.2	439.5 ± 14.0	40.3 ± 5.0
Ursolic acid (4)	114.4 ± 10.3	1,029.8 ± 104.4	nd	24.9 ± 3.8	22.2 ± 5.7
Lupeol (5)	nd	nd	12.3 ± 17.4	50.0 ± 9.9	56.8 ± 12.8
Betulinic acid (8)	nd	223.7 ± 13.9	6,490.0 ± 1,979.6	10.3 ± 1.4	4.6 ± 0.9
β-Amyrin (9)	116.7 ± 22.7	20.5 ± 4.9	13.1 ± 2.5	42.2 ± 3.3	12.1 ± 1.6
Oleanolic acid (12)	213.7 ± 4.8	5,670.5 ± 280.9	57.0 ± 10.7	72.5 ± 34.7	14.2 ± 7.5
24-OH-β-Amyrin (13)	nd	nd	nd	nd	nd
Sophoradiol (14)	79.8 ± 6.0	26.4 ± 2.1	131.4 ± 9.9	24.7 ± 0.6	15.7 ± 5.4
Soyasapogenol B (15)	609.2 ± 59.7	361.0 ± 46.6	5,443. 4 ± 442.3	364.8 ± 56.7	385.2 ± 120.0
Soyasapogenol A (16)	49.3 ± 7.5	56.8 ± 14.1	706.3 ± 67.5	3.2 ± 0.8	28.4 ± 2.4
Soyasapogenol E (17)	7.4 ± 0.7	5.0 ± 0.2	474.9 ± 129.0	11.6 ± 1.8	9.3 ± 5.0
Total	1,568.8 ± 56.4	7,399.9 ± 395.9	13,335.5 ± 2,032.2	1,043.5 ± 64.1	589.5 ± 154.5

The acid-hydrolyzed extracts were used for triterpenoid quantitation. Values are means ± standard deviation (SD) of three biological replicates calculated from the peak area of each compound and the internal standard (asiatic acid). Compounds were numbered according to **Fig. 1**. nd, not detected.

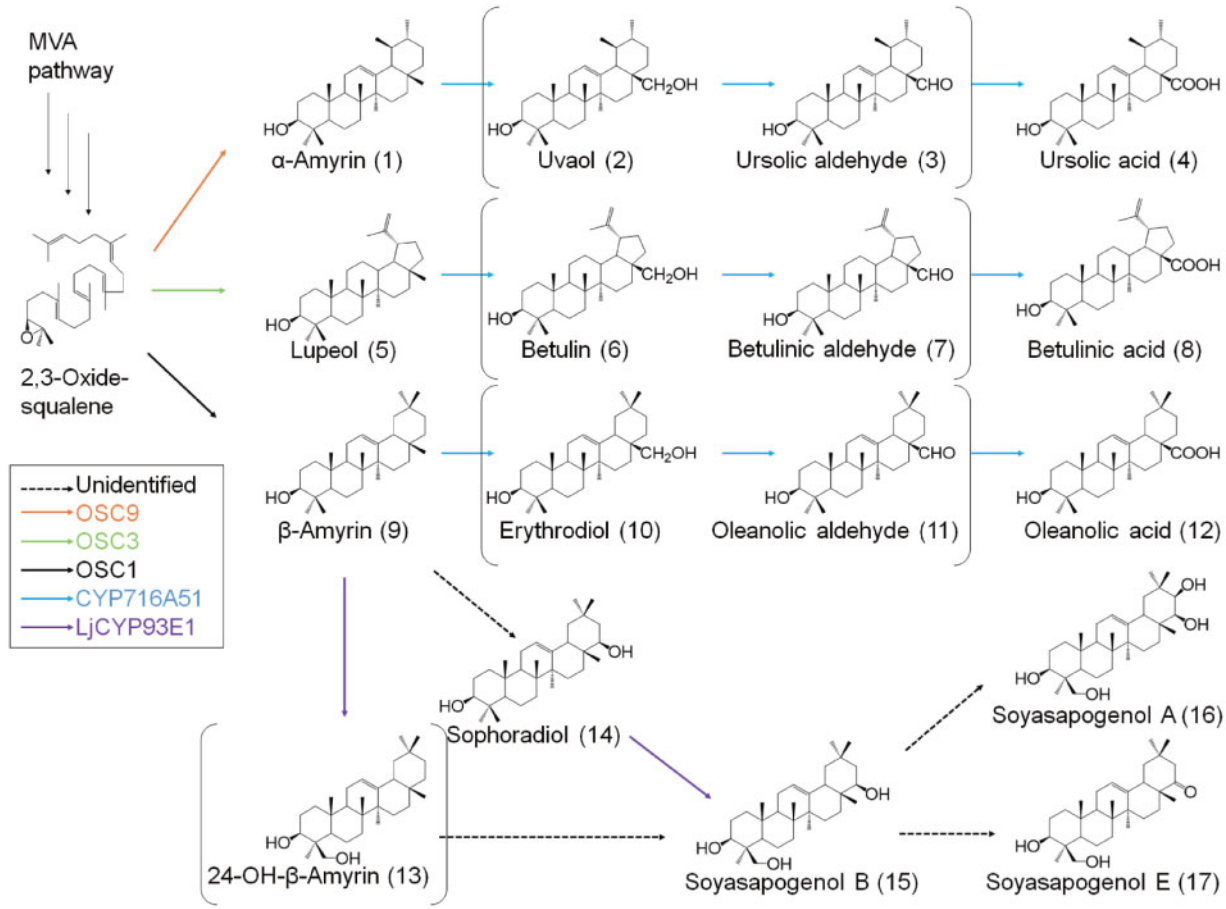

Fig. 1 Proposed triterpenoid biosynthetic pathways in *L. japonicus*. Compounds 1, 4, 5, 8, 9, 12, 14, 15, 16 and 17 were detected in *L. japonicus* plants. Compound 13 was not identified due to its small peak intensity in plant extracts. The presence of compounds 2, 3, 6, 7, 10 and 11 was not examined because authentic standards were not included in the GC-MS analysis of plant extracts. OSC enzymes reported thus far and OSC and P450 enzymes newly identified here catalyze these reactions. The compound numbers are consistent with those in **Table 1**.

Identification of α-amyrin synthase in *L. japonicus*

Our triterpenoid profiling experiments revealed the accumulation of α-amyrin (1) and ursolic acid (4) in *L. japonicus*. However, no OSC producing α-amyrin has been identified in this plant. We mined the *L. japonicus* genome database to find an unidentified OSC gene (Lj3g3v3617740), which was named *OSC9*.

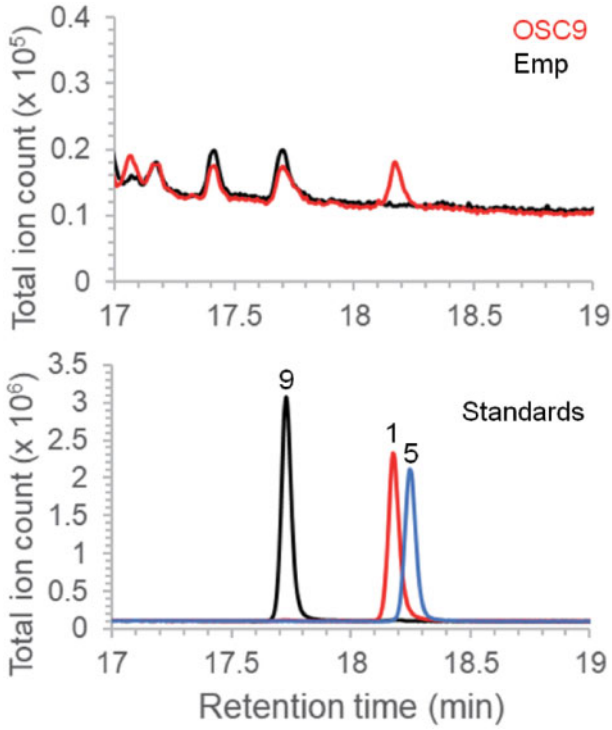

Fig. 2 GC-MS analysis of extracts of transgenic yeast expressing OSC9. α-Amyrin was detected in OSC9-expressing yeast extracts but not in empty vector control yeast extracts. Compounds were numbered according to **Fig. 1**.

To elucidate the function of OSC9, we employed a yeast heterologous expression system. α-Amyrin (1) was detected in the extracts of OSC9-expressing yeast culture but not in the extracts of empty vector control yeast culture (**Fig. 2**). No other triterpenoid compounds specifically produced in OSC9-expressing yeast were found in this assay. Therefore, OSC9 is an α-amyrin synthase (aAS) probably involved in the production of α-amyrin (1) and ursolic acid (4) in *L. japonicus*.

Identification of P450 enzymes involved in the production of major triterpenoids in *L. japonicus*

Recent studies have identified several P450 enzymes responsible for triterpenoid oxidation reactions in different plant species. Because *L. japonicus* accumulated C-28-oxidized triterpenoids and C-24-oxidized triterpenoids (soyasapogenols), it is reasonable that CYP716A and CYP93E subfamily P450s, previously reported as triterpenoid C-28 and C-24 oxidases, respectively, are involved in their production. One CYP716A gene (Lj4g3v3014670, *CYP716A51*) and one CYP93E gene (Lj1g3v3555800, *LjCYP93E1*, named by Dr. David Nelson, P450 Nomenclature Committee, in Du et al. 2016) were found in the *L. japonicus* genome using a BLAST search on miyakogusa.jp.

The activities of CYP716A51 and LjCYP93E1 were analyzed using a yeast heterologous expression system as well. The function of CYP716A51 was evaluated against three simple triterpene skeletons: α-amyrin (1), lupeol (5) and β-amyrin (9). Yeast strains expressing aAS (derived from *Olea europaea*)/cytochrome P450 reductase (CPR)/CYP716A51, lupeol synthase (*Glycyrrhiza uralensis* LUS)/CPR/CYP716A51 or β-amyrin synthase (*L. japonicus* bAS)/CPR/CYP716A51 produced ursolic acid (4), betulinic acid (8) or oleanolic acid (12), respectively, as final products (**Fig. 3A–C**). The metabolic intermediates, uvaol (2), betulin (6), erythrodiol (10), ursolic aldehyde (3) and oleanolic aldehyde (11), were also detected (**Fig. 3A–C**), similar to our previous report (Suzuki et al. 2018). Because the aAS used in this assay is a multifunctional OSC that produces α-amyrin and β-amyrin, β-amyrin derivatives were also produced in the yeast strain expressing aAS/CPR/CYP716A51 (**Fig. 3A**). The activity of LjCYP93E1 was tested against the β-amyrin skeleton. The peak corresponding to 24-hydroxy-β-amyrin (13) was detected in the GC-MS chromatogram of bAS/CPR/LjCYP93E1-expressing yeast extract (**Fig. 3D**). However, no oxidized triterpenoids were detected from the yeast strains expressing only the OSC and CPR genes (**Fig. 3**). These results demonstrate that CYP716A51 and LjCYP93E1 are C-28 and C-24 oxidases, respectively.

Expression patterns of *OSCs*, *CYP716A51* and *LjCYP93E1* in *L. japonicus*

Triterpenoids in *L. japonicus* accumulated in a tissue-dependent manner (**Table 1**). We performed quantitative real-time PCR (qPCR) analysis (**Fig. 4**) to examine the correlations between the triterpenoid accumulation patterns and expression patterns of triterpenoid biosynthetic genes in leaves, stems and roots. The expression of *OSC9* (encoding aAS) was observed in leaves and stems in which α-amyrin (1) and ursolic acid (4) accumulated. The *OSC9* expression level in roots was below the quantification limit. *OSC3* (encoding LUS) was highly expressed in roots that accumulated the most betulinic acid (8). *OSC1* (encoding bAS) is involved in the biosynthesis of both oleanolic acid (12) and soyasapogenols (15, 16 and 17) (**Fig. 1**). The *OSC1* expression was higher in stems (abundant oleanolic acid) and in roots (abundant soyasapogenols). The *CYP716A51* expression was too low to be quantified in leaves in which the smallest total amount of C-28 oxides accumulated. The most soyasapogenols accumulated in roots and the *LjCYP93E1* expression was similarly higher in roots. The expression of *OSC9*, *OSC3*, *OSC1*, *CYP716A51* and *LjCYP93E1* was variable in different tissues (**Fig. 4**) and these differences may cause the tissue-dependent accumulation of different triterpenoids in *L. japonicus* (**Table 1**).

Knockout of the *CYP716A51* gene in transgenic hairy roots

To directly confirm the involvement of CYP716A51 in C-28-oxidized triterpenoid production in planta, we used the CRISPR-Cas9 system. We previously described a CRISPR-Cas9 vector, pMgP237-2A-GFP, which can express multiplex gRNAs (Hashimoto et al. 2018, Nakayasu et al. 2018). Two target sequences on the first exon of *CYP716A51*, T1 and T2 (Supplementary Fig. S3A; Supplementary Table S2), were selected using CRISPRdirect software (Naito et al. 2015). Both target sequences were simultaneously integrated into the vector to generate T1/T2-pMgP237. Transgenic hairy roots

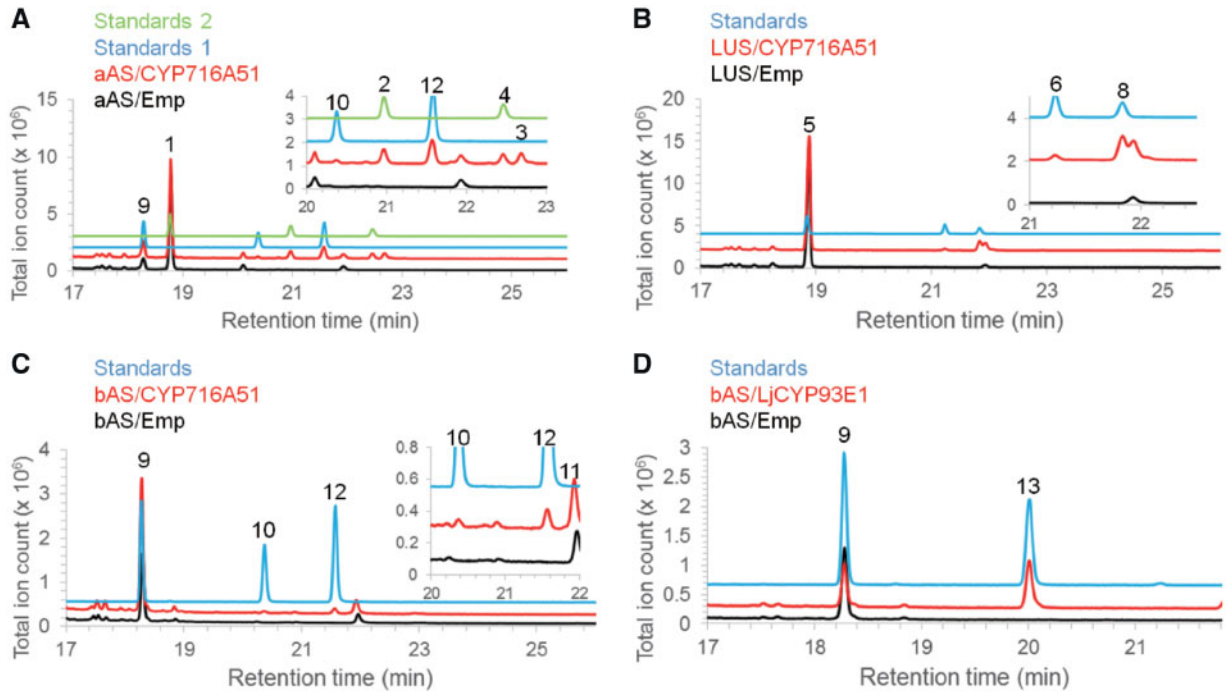

Fig. 3 GC-MS analysis of extracts of transgenic yeast after co-expression of OSCs, CPR and P450s. CYP716A51 was co-expressed with (A) aAS/CPR, (B) LUS/CPR or (C) bAS/CPR in yeasts. (D) LjCYP93E1 was introduced into bAS/CPR-expressing yeast. As controls, the corresponding empty vector lines were also analyzed, by GC-MS. Betulinic acid (8) and oleanolic aldehyde (11) co-eluted with a yeast endogenous metabolite. Compounds were numbered according to **Fig. 1**.

were induced by *Agrobacterium rhizogenes* ATCC15834 harboring T1/T2-pMgP237 or the empty vector as a control.

Putative *cyp716a51*-mutant hairy root lines were selected by PCR and electrophoresis (heteroduplex mobility assay). Some extra bands were observed in the hairy root lines (T1/T2_#7, #8, #11 and #15), but not in the vector control hairy root lines (VC_#1 and VC_#2) (Supplementary Fig. S3B), suggesting that mutations occurred in the *CYP716A51* gene and produced heteroduplex PCR fragments. Genomic DNA fragments around the T1 and T2 target sites were cloned and sequenced. Mutated alleles were not found in the control lines, VC_#1 and VC_#2 (Supplementary Fig. S3C). In addition, although wild-type (WT) sequences were detected in T1/T2_#7, #8 and #15, several types of mutations were observed in T1/T2_#7, #8, #11 and #15 (Supplementary Fig. S3C). All of the mutations resulted in frameshifts with the exception of a 36-bp deletion mutation around the T2 site in line T1/T2_#8, which resulted in the translation of a short protein with as-yet-unknown activity.

We analyzed the triterpenoid composition of the control and mutant lines. The control lines (VC_#1 and VC_#2) accumulated several oxidized triterpenoids, including oleanolic acid (12), betulinic acid (8), sophoradiol (14) and soyasapogenol B (15) (**Fig. 5**). The mutant lines T1/T2_#7, #8, #11 and #15 accumulated sophoradiol (14) and soyasapogenol B (15) but lacked oleanolic acid (12) and betulinic acid (8) (**Fig. 5**). The same results were confirmed in extracted ion chromatograms (EICs) at $m/z = 189$ and 203 (Supplementary Fig. S4). These results confirm that CYP716A51 is involved in oleanolic and betulinic acids production in *L. japonicus*. Although CYP716A51 produced ursolic acid (4) when heterologously co-expressed with aAS and CPR in yeast, hairy roots of *L. japonicus* did not accumulate ursolic acid (4) or α-amyrin (1).

Ljcyp93e1 loss-of-function mutant plants

The seeds of two independent mutant lines, 30030559 (X) and 30030927 (Y), both of which contain an *LORE1* insertion (Fukai et al. 2012, Urbański et al. 2012) into the third exon of *LjCYP93E1* were obtained (Supplementary Fig. S5A). WT allele-specific primer sets (A and B) and insertion allele-specific primer sets (A and C) were designed for the *LjCYP93E1* gene and other genes that may have *LORE1* insertions in their exons (Supplementary Fig. S5B; Supplementary Tables S3, S4) according to a previous report (Urbański et al. 2012). Heterozygous *Ljcyp93e1* mutants (#X12 and #Y1) and homozygous *Ljcyp93e1* mutants (#X1 and #Y4) were selected by PCR analysis (Supplementary Fig. S5C, D).

We analyzed the sapogenin composition of the leaves of WT, heterozygous *Ljcyp93e1*-mutant and homozygous *Ljcyp93e1*-mutant plants (Supplementary Fig. S6). 24-Hydroxy-β-amyrin (13) was not detected in any plants, including WT plants, but all of the plants accumulated sophoradiol (14). Soyasapogenols (15, 16 and 17) were detected in WT and heterozygous mutant lines (#X12 and #Y1), but not in homozygous *Ljcyp93e1*-mutant lines (#X1 and #Y4).

To reveal how *LORE1* insertion affected the expression of *LjCYP93E1*, we analyzed the transcript levels of *LjCYP93E1* and *OSC1* in WT plants and the progeny of #X1 line by qPCR analysis (Supplementary Fig. S7). The expression of *LjCYP93E1* decreased

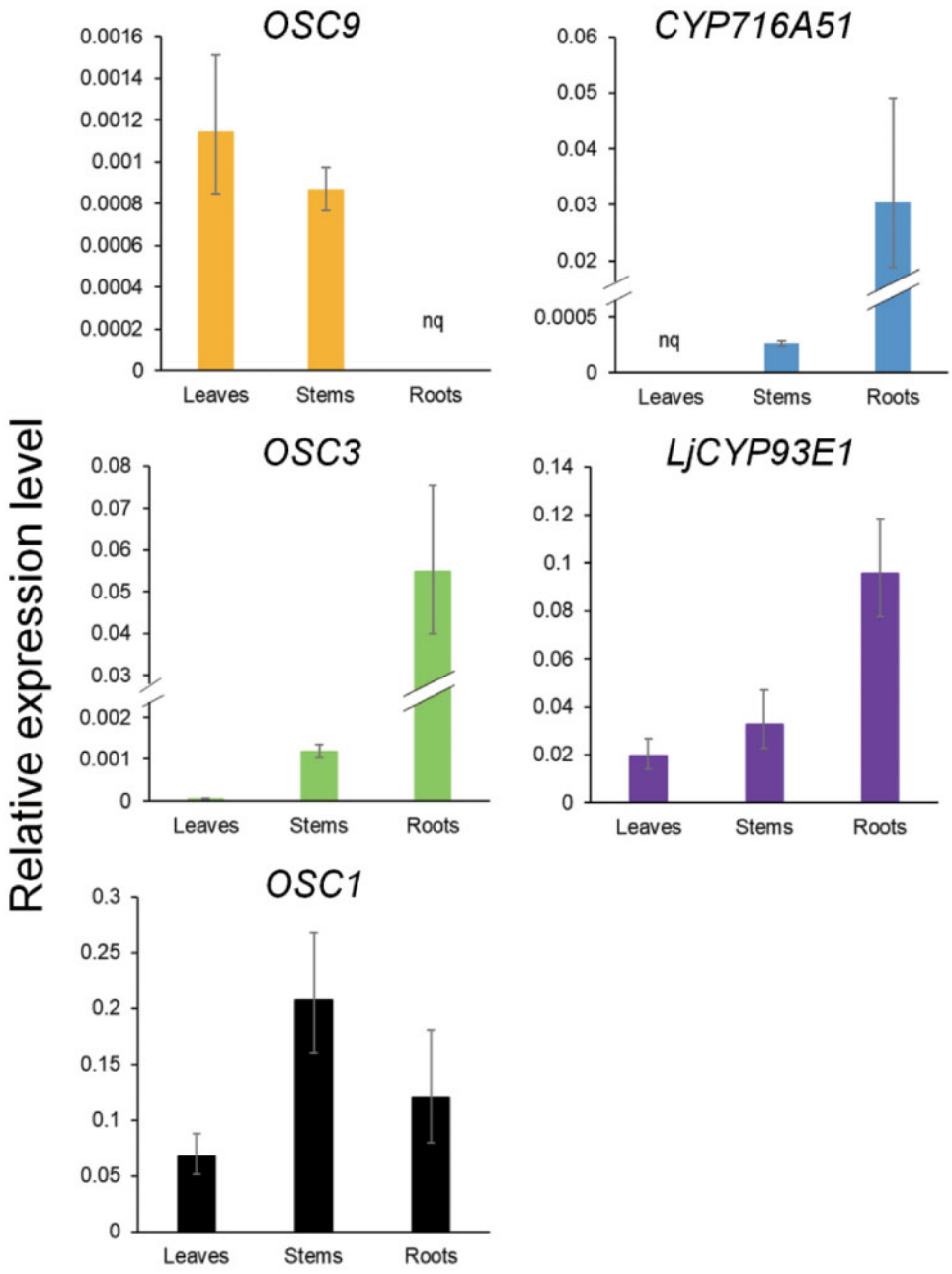

Fig. 4 qPCR analysis of *OSCs* and *P450s* in leaves, stems and roots. Relative expression levels of the genes were analyzed in triplicate. The values were normalized with the expression level of a reference gene (*LjUBQ1*). Means ± standard deviation (SD). nq, detected but not quantifiable.

significantly in #X1 plants, whereas *OSC1* expression was not affected. Amplification was observed in #X1 plants, even using a third exon-specific primer set (amplifying the downstream sequences of the *LORE1* insertion site). These observations suggest that *LORE1* insertion into the exon of *LjCYP93E1* caused misfolding and degradation of the transcripts (Doma and Parker 2007) and also resulted in the translation of nonactive proteins, although the *Ljcyp93e1*-mutant allele was fully transcribed.

To examine the metabolic changes between the WT and homozygous *Ljcyp93e1*-mutant lines in detail, we compared the triterpenoid compositions of the leaves (**Fig. 6A**), stems (**Fig. 6B**) and roots (**Fig. 6C**) of WT, #X1 and #Y4 lines with and without acid hydrolysis. Peaks corresponding to soyasapogenol B (15), soyasapogenol A (16), soyasapogenol E (17) and unknown compounds (u1–u4) were not detected in any tissues of *Ljcyp93e1*-mutant lines. These unknown compounds may be artifacts of the acid hydrolysis of soyasaponins, such as soyasapogenol C and F (Carelli et al. 2011, Tava et al. 2017). By contrast, α-amyrin (1), ursolic acid (4), lupeol (5), betulinic acid (8), β-amyrin (9), oleanolic acid (12) and sophoradiol (14) were detected in both WT and *Ljcyp93e1*-mutant plants (**Fig. 6**). The accumulation of β-amyrin (9) and sophoradiol (14) increased in *Ljcyp93e1*-mutant plants because they are the

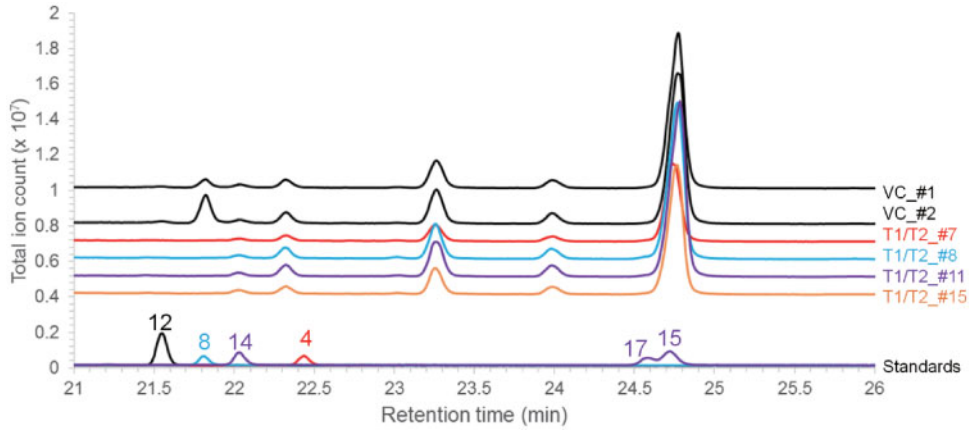

Fig. 5 GC-MS analysis of triterpenoids in *cyp716a51*-mutant hairy root lines. Peaks corresponding to oleanolic acid (12) and betulinic acid (8) were not detected in T1/T2 lines. Triterpenoids were extracted from 10 mg of dried materials and resuspended in 300 µl of chloroform/methanol (1:1) after acid hydrolysis. One hundred microliters of the solution were evaporated and derivatized followed by GC-MS analysis. Compounds were numbered according to **Fig.1**. VC, vector control.

precursors of the soyasapogenols that were reduced in the *Ljcyp93e1* mutants (**Fig. 6**). More sophoradiol was detected in the acid-hydrolyzed extracts than in the nonhydrolyzed extracts of *Ljcyp93e1* mutants, suggesting that the UGT enzymes predominantly glycosylating soyasapogenols in WT plants accepted sophoradiol as a substrate instead of soyasapogenols in *Ljcyp93e1*-mutant plants. The competing pathway product, oleanolic acid (12), also increased in the root extracts of *Ljcyp93e1*-mutant plants (**Fig. 6C**).

Unknown peaks, u5 and u6, were detected in the leaves and stems of *Ljcyp93e1* plants, and their intensities increased in the roots of *Ljcyp93e1* plants compared with WT plants (**Fig. 6**). Peaks u5 and u6 were not detected in the nonhydrolyzed samples, indicating that they are glycosylated in *L. japonicus* plants. The molecular ion of u5 was $m/z = 586.4$ (Supplementary Fig. S8), consistent with monohydroxy-β-amyrin + 2TMSi. Peak u6 had a molecular ion of $m/z = 674.5$ (Supplementary Fig. S8), consistent with dihydroxy-β-amyrin + 3TMSi. Because the mutant plants lacked the C-24 oxidation activity by LjCYP93E1, compound u5 may correspond to 21-hydroxy-β-amyrin and compound u6 may correspond to cantoniensistriol (21,22-dihydroxy-β-amyrin). However, we could not identify these peaks due to the lack of corresponding authentic standards. In addition, the amount was too small to purify and to elucidate the structure by nuclear magnetic resonance analysis.

Finally, we performed liquid chromatography-mass spectrometry (LC-MS) analysis to confirm changes in saponin composition in *Ljcyp93e1*-mutant plants. Soyasaponin I content was analyzed in the leaves, stems and root extracts of WT, #X1 and #Y4 lines (**Fig. 7A;** Supplementary Figs. S9, S10). Soyasaponin I is common in legume plants; it has soyasapogenol B as aglycone and glucuronic acid, galactose and rhamnose attached to C-3 (**Fig. 7B**). The #X1 and #Y4 lines completely lacked soyasaponin I (**Fig. 7A**). In addition, several peaks that may correspond to other soyasaponins found in the TICs of WT plant extracts also disappeared in *Ljcyp93e1*-mutant plant extracts (Supplementary Fig. S9).

Metabolic changes at the sapogenin and saponin levels indicate the essential activity of LjCYP93E1 in *L. japonicus* triterpenoid biosynthesis.

Discussion

Triterpenoid biosynthesis in *L. japonicus*

The presence of lupeol, soyasapogenol B and soyasaponins produced from the aglycones, soyasapogenol B, A and E, in *L. japonicus* plants has been reported (Delis et al. 2011, Tsuno et al. 2018). In addition to the abovementioned triterpenoids, we reported in this study the accumulation of ursane, lupane and oleanane triterpenoids in a tissue-dependent manner (**Fig.1**; **Table 1**). These results provide us with basic knowledge of the biosynthesis and physiological functions of triterpenoids in *L. japonicus*.

α-Amyrin (1) and ursolic acid (4) were detected throughout tissues; however, roots accumulated the smallest amount of α-amyrin and did not accumulate ursolic acid (**Table 1**). We identified OSC9, which produced α-amyrin when heterologously expressed in yeast. The accumulation pattern of α-amyrin (**Table 1**) correlated with the expression pattern of *OSC9* (**Fig. 4**), suggesting the involvement of *OSC9* in ursane triterpenoid biosynthesis in this plant. In the phylogenetic tree of OSC proteins, OSC9 is located in the same clade as β-amyrin synthases and multifunctional OSCs, some of which have α-amyrin producing activities (Supplementary Fig. S11). To date, no OSC producing α-amyrin as a single product has been found (Brendolise et al. 2011, Thimmappa et al. 2014, Srisawat et al. 2019). In this study, we detected no triterpenoid products other than α-amyrin in the OSC9-expressing yeast extract (**Fig. 2**). However, OSC9 may produce undetectable levels of other triterpenoids in this assay. Scale-up culture or a different expression system will be required to clarify the product specificity of OSC9.

It was previously suggested that lupeol or its further modified compound regulates the expression of the *ENOD40* gene

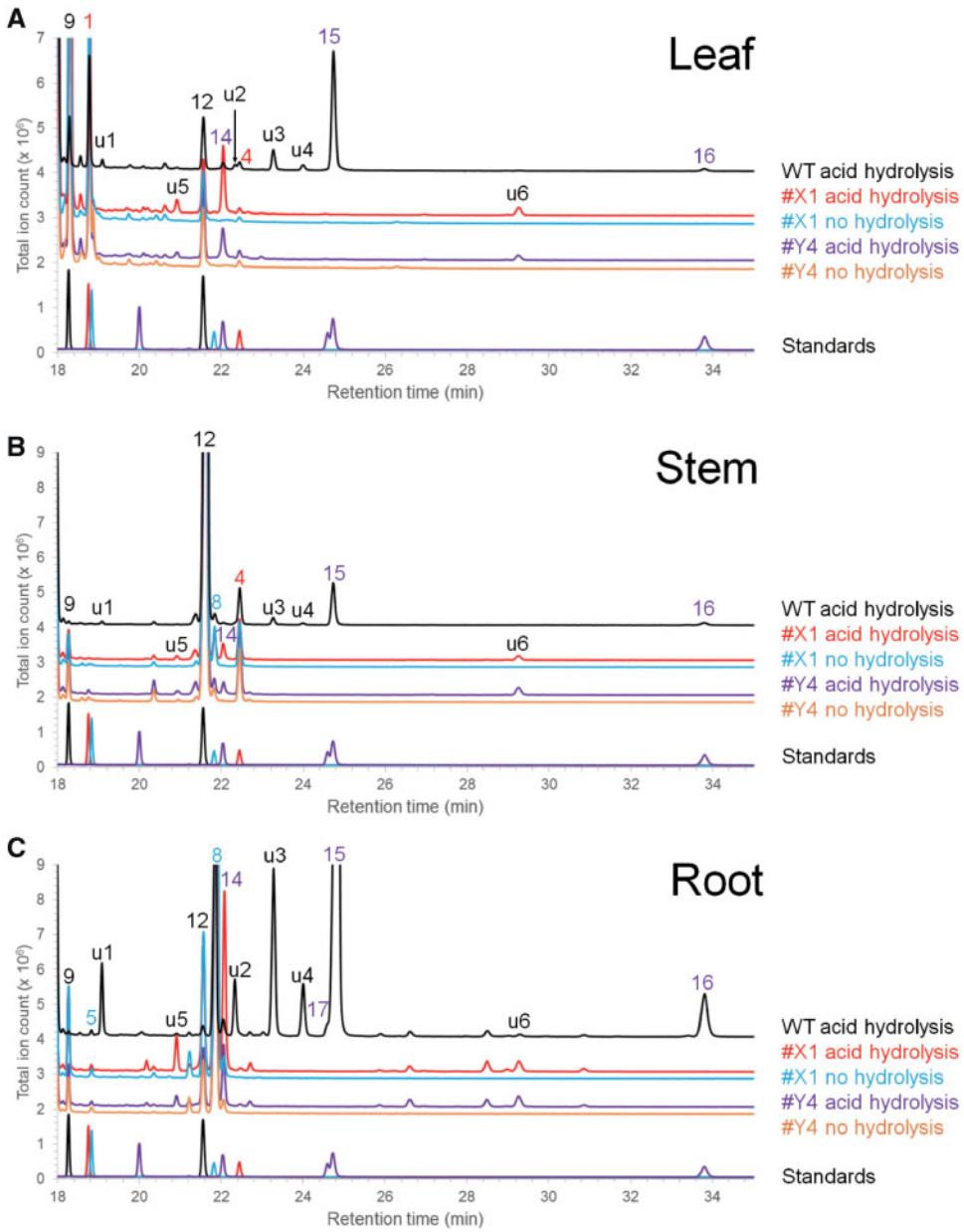

Fig. 6 Changes in triterpenoid production between WT and *Ljcyp93e1*-mutant plants. TICs of extracts from (A) leaves, (B) stems and (C) roots from WT and *Ljcyp93e1*-mutant plants with or without acid hydrolysis. Triterpenoids were extracted from 10 mg of dried materials and resuspended in 300 µl of chloroform:methanol (1:1). A portion of the solution (100 µl for leaves or 50 µl for stems and roots) was evaporated and derivatized followed by GC-MS analysis. Compounds were numbered according to **Fig.1**.

and affects nodule formation in *L. japonicus* (Delis et al. 2011). In our triterpenoid quantitation analysis, nonglycosylated betulinic acid (8) was predominant in roots (**Table 1**; Supplementary Fig. S1). Lupeol (5) could not be quantified in two of the three biological replicates of roots, suggesting that lupeol is rapidly converted into betulinic acid in roots. Thus, betulinic acid may be the real contributor to *ENOD40* gene expression regulation.

β-Amyrin (9) and oleanolic acid (12) were also found throughout the tissues. Stems accumulated the most oleanolic acid without glycosylation (**Table 1**; Supplementary Fig. S1). No hemolytic triterpenoid, except for oleanolic acid, was detected in *L. japonicus*, although the other model legume *M. truncatula* produces hemolytic saponins by further modifying oleanolic acid.

The presence of soyasapogenols B, A and E (15, 16 and 17) after acid hydrolysis of methanol extracts was consistent with the presence of 2,3-dihydro-2,5-dihydroxy-6-methyl-4*H*-pyran-4-one (DDMP) and groups B, A and E soyasaponins in this plant as reported previously (Tsuno et al. 2018). In addition to LjCYP93E1, CYP72A subfamily P450 enzymes are likely involved in *L. japonicus* soyasapogenols biosynthesis. Seven CYP72A subfamily genes are found in the *L. japonicus* genome. Lj3g3v3776580.1 is likely involved in soyasapogenol B

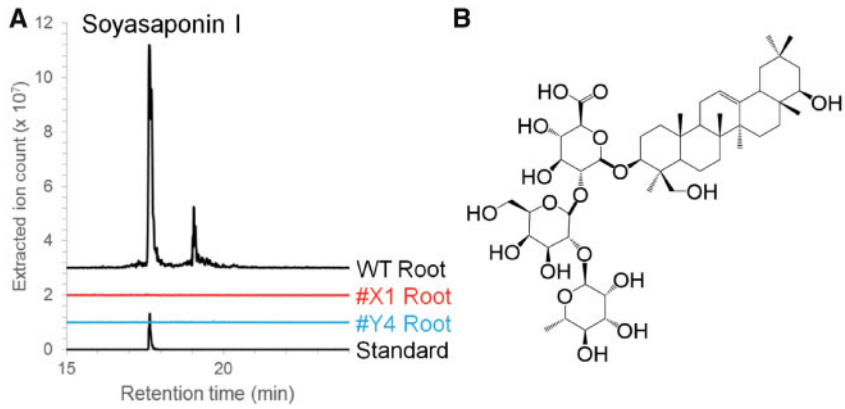

Fig. 7 Complete abolition of soyasaponin I in *Ljcyp93e1*-mutant *L. japonicus*. Saponins extracted from roots were analyzed by LC-MS. (A) EICs ($m/z = 941.5$) of roots from WT and *Ljcyp93e1*-mutant plants and the authentic standard of soyasaponin I (Rt = 17.65 min). (B) Structure of soyasaponin I (3-rhamnose-galactose-glucuronic acid-soyasapogenol B).

biosynthesis via the C-22β oxidation of the β-amyrin skeleton since it clusters tightly with previously reported C-22 oxidases, *G. max* CYP72A61 (Ebizuka et al. 2011), *M. truncatula* CYP72A61v2 (Fukushima et al. 2013), and *G. uralensis* CYP72A566 (Tamura et al. 2018) in a phylogenetic tree (Supplementary Fig. S12). CYP72A69 (Sg-5) was recently characterized as a C-21β-oxidase responsible for soyasapogenol A biosynthesis in *G. max* (Yano et al. 2017). By contrast, a P450 enzyme that converts soyasapogenol B into soyasapogenol E has not yet been identified. Further experiments to prove the involvement of CYP72A subfamily enzymes in soyasapogenols and soyasaponins biosynthesis in *L. japonicus* are currently ongoing.

CYP716A51 is involved in C-28-oxidized triterpenoid biosynthesis in planta

Lotus japonicus contains one full-length *CYP716A* gene (*CYP716A51*) in its genome. CYP716A51 converts α-amyrin, lupeol and β-amyrin (1, 5 and 9) skeletons into ursolic, betulinic and oleanolic acids (4, 8 and 12), respectively, in yeast via the corresponding hydroxyl and aldehyde intermediates (**Figs. 1, 3A–C**). Interestingly, although a similar total amount of C-28 oxides accumulated in the stems and roots (**Table 1**), the expression level of *CYP716A51* was 113 times higher in roots (**Fig. 4**). CYP716A subfamily enzymes accepted α-amyrin, lupeol and β-amyrin as substrates but had different conversion ratios (amount of final oxidized product/amount of substrate) when heterologously expressed in yeast (Suzuki et al. 2018). Similarly, the conversion efficiency of lupeol by CYP716A51 seemed to be lower than that of β-amyrin in yeast (**Fig. 3A–C**). This plant may express *CYP716A51* at higher levels in roots to overcome the low conversion efficiency of lupeol by CYP716A51 and to produce betulinic acid efficiently.

We performed in planta loss-of-function analysis of *CYP716A51* in transgenic hairy roots using the pMgP237-2A-GFP vector. The results indicated that this CRISPR-Cas9 vector can be used in *L. japonicus* hairy roots as well as potato hairy roots (Nakayasu et al. 2018) and tomato plants (Hashimoto et al. 2018). The genome of the vector control lines (VC_#1 and VC_#2) harbored only the WT allele of *CYP716A51* (Supplementary Fig. S3) and produced the CYP716A51 products, betulinic and oleanolic acids (**Fig. 5**; Supplementary Fig. S4). Although both WT and mutated *CYP716A51* alleles were detected in three *cyp716a51*-mutant hairy root lines (T1/T2_#7, #8 and #15), probably due to mosaicism, none of these lines accumulated C-28-oxidized triterpenoids (**Fig. 5**; Supplementary Fig. S4). The WT allele was not found in the genome of the *cyp716a51*-mutant line T1/T2_#11 (Supplementary Fig. S3C), and no C-28-oxidized triterpenoids were detected (**Fig. 5**; Supplementary Fig. S4). Incomplete disruption of the *CYP716A51* genome resulting in a lack of C-28-oxidized triterpenoids may have occurred because (1) cells expressing active CYP716A51 may not be able to produce a detectable level of C-28 oxides or (2) CYP716A51 may be expressed in specific cell types and *CYP716A51* could be completely disrupted in these cells. In any case, mutations in *CYP716A51* and the abolition of C-28-oxidized triterpenoids found specifically in T1/T2 hairy root lines strongly imply the in planta function of CYP716A51 as a triterpenoid C-28 oxidase, consistent with the yeast in vivo functional analysis. CRISPR-Cas9 may induce undesired off-target mutations. To discard the possibility that the abolition of C-28 oxidized triterpenoids in established *cyp716a51*-mutant hairy root lines may result from off-target effects, we surveyed the *L. japonicus* genome and found no putative off-target sequences for T1 or T2 sites.

The functions of CYP716 family enzymes are well studied (Ghosh 2017). Most of the CYP716A subfamily enzymes exhibit C-28 oxidation activities (Carelli et al. 2011, Fukushima et al. 2011, Suzuki et al. 2018). Some CYP716A members also oxidize different carbon positions not only C-28 positions (Moses et al. 2015, Yasumoto et al. 2016, Miettinen et al. 2017, Tamura et al. 2017, Pütter et al. 2019). CYP716C, CYP716E, CYP716S and CYP716Y subfamily members show the oxidation activities at different carbon positions of pentacyclic triterpenoid backbones (Moses et al. 2014a, Yasumoto et al. 2017, Miettinen et al. 2017, Nakamura et al. 2018, Sandeep et al. 2019). CYP716A51 is tightly clustered with other C-28 oxidases in

the phylogenetic tree (Supplementary Fig. S13), which further supports the result that CYP716A51 has a C-28 oxidase activity as indicated in our experiments (Figs. 3A–C, 5; Supplementary Figs. S3, S4).

The physiological functions of C-28-oxidized triterpenoids in plants are largely unknown. Until now, *CYP716A12*-knockout *M. truncatula* (Carelli et al. 2011) and *CYP716A252*- or *CYP716A253*-transiently silenced *Ocimum basilicum* (Misra et al. 2017) have been reported. *Cyp716a12*-mutant *M. truncatula* with altered hemolytic saponin production showed a drastic dwarf phenotype. Why the disruption of *CYP716A12* resulted in the phenotypical alteration is still under discussion. However, *ugt73f3*-mutant *M. truncatula* also exhibited a dwarf phenotype (Naoumkina et al. 2010), suggesting that triterpenoid sapogenins may be toxic or that hemolytic saponins are important for the normal growth of *M. truncatula*. In this study, we used hairy roots to characterize the in planta function of *CYP716A51*. Disruption of *CYP716A51* and the abolition of oleanolic and betulinic acids did not result in growth inhibition, at least in hairy roots (data not shown). The generation of whole *cyp716a51*-mutant plants will enable comparison of the two model legume plants, *M. truncatula* and *L. japonicus*, which accumulate different C-28-oxidized triterpenoids.

LjCYP93E1 is involved in soyasaponin biosynthesis in planta by catalyzing C-24 oxidation of the β-amyrin backbone

Several *CYP93E* subfamily genes from different leguminous plants encode β-amyrin C-24 oxidases, as determined by heterologous expression in yeast (Shibuya et al. 2006, Seki et al. 2008, Fukushima et al. 2011, Moses et al. 2014b). The in planta activity of CYP93E1 was demonstrated in an analysis of *cyp93e1*-mutant *G. max* (Krishnamurthy et al. 2019).

The genome of *L. japonicus* contains one *CYP93E* gene (*LjCYP93E1*). In vivo functional analysis of LjCYP93E1 in yeast (Fig. 3D) and triterpenoid composition analysis of WT, heterozygous *Ljcyp93e1*-mutant and homozygous *Ljcyp93e1*-mutant plants (Figs. 6, 7; Supplementary Fig. S6) indicated that LjCYP93E1 has β-amyrin C-24 oxidase activity and provided evidence for the involvement of *LjCYP93E1* in soyasaponin biosynthesis in planta. The correlation of soyasapogenol accumulation patterns (Table 1) and *LjCYP93E1* expression patterns (Fig. 4) also supports our conclusion. All of the functionally characterized CYP93E subfamily enzymes catalyze C-24 oxidation of the β-amyrin backbone. LjCYP93E1 reasonably showed around 80% amino acid identity with other CYP93E subfamily members.

Mutant legume plants with altered soyasaponin compositions have been reported. Takagi et al. (2011) manipulated the expression of the *bAS* gene (*GmBAS1*) and significantly reduced soyasaponin accumulation in soybean seeds by RNA interference with a seed-specific promoter, suggesting that soyasaponins are not essential for seed germination. Natural soybean mutants of *CYP72A69* (*Sg-5*) lacked group A soyasaponins but produced other types of soyasaponins (Yano et al. 2017, Rehman et al. 2018). Krishnamurthy et al. (2019) isolated mutants of *GmBAS1*, *CYP93E1* and *CYP72A61* from a *G. max* mutant library and showed that the *Gmbas1*-mutant plants lacked saponins; however, this did not affect their growth. In our study, *Ljcyp93e1*-knockout *L. japonicus* plants lacking soyasaponins in all tissues grew with no obvious morphological changes and could breed normally by self-pollination (data not shown). These results suggest that soyasaponins are not essential for the growth of *L. japonicus* plants as well. The *Ljcyp93e1*-mutant *L. japonicus* analyzed in this study will be useful to further study the physiological functions of soyasaponins in planta.

The utility of *L. japonicus* for triterpenoid research

Glycine max, *M. truncatula* and *L. japonicus* are useful for molecular biology research because their genomes have been sequenced (Sato et al. 2008, Schmutz et al. 2010, Young et al. 2011), and they are more readily stably transformed than other legume plants (Finer and McMullen 1991, Chabaud et al. 2007, Hirota et al. 2012). In addition, retrotransposon insertion mutants of *M. truncatula* and *L. japonicus* are available (Tadege et al. 2008, Fukai et al. 2012, Urbański et al. 2012).

The abovementioned three legume plants have similar soyasaponin compositions, suggesting possession of common biosynthetic pathways and physiological functions. Many soyasaponin-deficient *G. max* mutants have been isolated to identify soyasaponin biosynthesis-related genes (Krishnamurthy et al. 2019). However, the genome of *G. max* is two to three times larger than that of *L. japonicus* due to two whole-genome duplication events (Schmutz et al. 2010). Functionally redundant genes may lead to failure to manifest soyasaponin-deficient phenotypes in mutants in a particular soyasaponin biosynthetic gene. Thus, the application of high-throughput mutant screening strategy used for *G. max* (Krishnamurthy et al. 2019) to *L. japonicus* may facilitate the discovery of new soyasaponin-deficient mutant plants that were not found in *G. max*.

Among the abovementioned plants, only *M. truncatula* produces hemolytic saponins. In *Medicago* species, triterpenoid saponins, including hemolytic saponins, are thought to function in the defense against insects and fungi (Tava and Avato 2006). Overexpression of *bAS* (*AsOXA1*) in *M. truncatula* enhanced saponin accumulation and nodulation (Confalonieri et al. 2009). Instead of hemolytic saponins, *L. japonicus* accumulated nonglycosylated oleanolic acid (a common precursor of diverse hemolytic saponins in *M. truncatula*), ursolic acid and betulinic acid (Table 1; Supplementary Fig. S1). Generally, nonglycosylated triterpenoids are components of cuticular wax, which protect the plant from biotic and abiotic stresses (Lee and Suh 2015). A previous study (Delis et al. 2011) and our results (Table 1; Supplementary Fig. S1) suggest that betulinic acid regulates nodulation in *L. japonicus*. The differences between *M. truncatula* and *L. japonicus* make the latter useful for analyzing the differentiation of the mechanisms by which triterpenoid sapogenins and saponins regulate defense and nodulation in legume plants.

Triterpene saponin biosynthesis activating regulator 1 (TSAR1) and TSAR2 are basic helix–loop–helix (bHLH) transcription factors (TFs) that regulate the biosynthesis of soyasaponins and hemolytic saponins, respectively, in *M. truncatula* (Mertens et al. 2016). Interestingly, GubHLH3, the TF that regulates soyasaponin biosynthesis in the medicinal legume *G. uralensis* (licorice), is phylogenetically closer to TSAR2 than to TSAR1 (Tamura et al. 2018). Therefore, the mechanisms that regulate triterpenoid biosynthesis are not conserved in *M. truncatula* and *G. uralensis*. Analysis of *L. japonicus* will facilitate the identification of orthologous TFs and evaluation of the evolution of the mechanisms that regulate triterpenoid biosynthesis in legume plants.

Concluding remarks

We evaluated the triterpenoid profile of *L. japonicus* and identified the enzyme functions of *OSC9*, *CYP716A51* and *LjCYP93E1* using yeast heterologous expression systems. Furthermore, we assessed the in planta roles of CYP716A51 and LjCYP93E1 in triterpenoid biosynthesis by performing gene loss-of-function analyses using the CRISPR-Cas9 system and *LORE1* insertion lines. Our experiments proved the involvement of *CYP716A51* and *LjCYP93E1* in the in planta biosynthesis of C-28 oxidized triterpenoids and soyasaponins, respectively. These results increase our knowledge of triterpenoid biosynthesis in legume plants and will facilitate further studies of the biosynthesis and physiological functions of triterpenoids in *L. japonicus*.

Materials and Methods

Plant materials

Lotus japonicus Gifu B-129 WT and *LORE1* insertion lines (Fukai et al. 2012, Urbański et al. 2012) were provided by Miyazaki University, Japan and Aarhus University, Denmark, through the National BioResource Project (NBRP).

Chemicals

β-Amyrin, α-amyrin, lupeol, erythrodiol, uvaol, oleanolic acid, ursolic acid and asiatic acid were purchased from Extrasynthese (Lyon, France). Betulin was purchased from Sigma-Aldrich (St. Louis, MO, USA). Betulinic acid was purchased from Tokyo Chemical Industry (Tokyo, Japan). Soyasapogenol B and soyasapogenol A were purchased from Tokiwa Phytochemical (Chiba, Japan). Sophoradiol, 24-hydroxy-β-amyrin and soyasapogenol E were kindly provided by Dr. Kiyoshi Ohyama.

Cloning of *OSC9*, *CYP716A51* and *LjCYP93E1* and plasmid construction

The coding sequence (CDS) of OSC9 and the backbone of pENTR1A vector were amplified from *L. japonicus* cDNA and purified pENTR1A plasmids, respectively, using primer nos. 1–4. The PCR products were purified using Monarch DNA Gel Extraction Kit (New England Biolabs, Beverly, MA, USA). The Gateway entry clone (pENTR1A-OSC9) was generated using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). The expression vector (pYES3-ADH-OSC9) was generated using Gateway LR Clonase II Enzyme Mix (Thermo Fisher Scientific, Waltham, MA, USA).

The cDNA clone of *CYP716A51* (MWL008g03) was provided by the NBRP (Asamizu et al. 2000, Sato et al. 2001, Asamizu et al. 2004). The CDS of *CYP716A51* was amplified using primer nos. 5 and 6 (Supplementary Table S4). The PCR products were cloned into the pENTR/D-TOPO vector using pENTR Directional TOPO Cloning Kits (Thermo Fisher Scientific) following the manufacturer's instructions. The CDS of *LjCYP93E1* was amplified from *L. japonicus* cDNA using primer nos. 7 and 8 and was cloned into the Gateway pDONR221 vector using the Gateway BP Clonase II Enzyme Mix (Thermo Fisher Scientific) following the manufacturer's instructions. The CDSs of *CYP716A51* and *LjCYP93E1* were transferred into the pELC-GW (Seki et al. 2008) vector using the Gateway LR Clonase II Enzyme Mix (Thermo Fisher Scientific) to generate pELC-CYP716A51 and pELC-LjCYP93E1, respectively. The clone MWL008g03 is derived from the *L. japonicus* Miyakojima MG-20 accession. Sequencing analysis confirmed that the nucleotide sequences from Miyakojima MG-20 and Gifu B-129 accessions were identical.

Yeast in vivo assay

We performed yeast in vivo assay as described previously with minor modifications (Fukushima et al. 2011, Suzuki et al. 2018). *Saccharomyces cerevisiae* INVSc1 (Invitrogen, Carlsbad, CA, USA) harboring pYES3-ADH-bAS (LjOSC1; AB181244), pYES3-ADH-aAS (aAS from *O. europaea*; AB291240) or pYES3-ADH-LUS (LUS from *G. uralensis*; AB663343) was transformed with the above-described expression clones to generate bAS/CPR (CPR from *L. japonicus*; AB433810)/CYP716A51-expressing yeast, aAS/CPR/CYP716A51-expressing yeast, LUS/CPR/CYP716A51-expressing yeast and bAS/CPR/LjCYP93E1-expressing yeast. The empty pELC vector was transformed into OSC-expressing yeast strains to generate control strains (OSC/CPR-expressing yeasts). Transgenic yeasts were inoculated into 2.5 ml of synthetic complete (SC) + 2% glucose medium without specific amino acids. After incubation overnight at 30°C with shaking at 200 rpm, cells were collected by centrifugation and resuspended in 5 ml of SC + 2% galactose medium without specific amino acids to induce CPR and P450s expression. Cells were cultured for 3 d at 30°C with shaking at 180 rpm. The obtained cultures were stored at −30°C until extraction. Next, ethyl acetate (3 ml) was added to the cultures. Cells were sonicated for 30 min and centrifuged, and the ethyl acetate layer was collected. This extraction step was repeated three times. The ethyl acetate was allowed to completely evaporate and the pellet was resuspended in 500 µl of MeOH: chloroform (1:1). One hundred microliters (for functional analysis of LjCYP93E1) or 400 µl (for functional analysis of CYP716A51) of the solution were evaporated and trimethylsilylated using a mixture of 50 µl of *N*,*N*-dimethylformamide (Kishida Chemical Co., Ltd., Osaka, Japan) and 50 µl of BSTFA:TMCS (99:1) (TCI) at 80°C for 30 min.

To analyze OSC9 activity, yeasts harboring pYES3-ADH-GW empty vector were analyzed as a negative control. The culture and extraction methods were similar to those described above, but 2% glucose was used instead of 2% galactose because OSC9 expression is driven by a constitutive promoter.

Quantitative real-time PCR

Total RNA was extracted from leaves, stems and roots of 2.5-month-old *L. japonicus* (Gifu B-129) grown on soil using PureLink Plant RNA Reagent (Thermo Fischer Scientific) in triplicate. The RNA obtained was purified using the RNeasy Plant Mini Kit (Qiagen, USA) after digesting contaminated genomic DNA with recombinant DNase I (RNase-free) (Takara Bio, Shiga, Japan). First-strand cDNA was synthesized from purified total RNA by PrimeScript RT Master Mix (Perfect Real Time) (Takara Bio). We performed qPCR analysis with Light Cycler Nano (Roche, Germany) and FastStart Essential DNA Green Master (Roche). The primers used are listed in Supplementary Table S4 (primer nos.9–20). The expression of the *Ubiquitin* (*UBQ*) gene was analyzed as a reference gene following Delis et al. (2011).

Cyp716a51 loss-of-function mutant hairy root lines

The multiplex guide RNA (gRNA)-expressing CRISPR-Cas9 vector, pMgP237-2A-GFP (Hashimoto et al. 2018, Nakayasu et al. 2018), was used for genome editing of *L. japonicus*. The target sequences of the gRNAs (Supplementary Table S2) were selected from *CYP716A51* using the web-based tool, CRISPRdirect (https://crispr.dbcls.jp/) (Naito et al. 2015). Two gRNA target sequences were simultaneously transferred into the pMgP237-2A-GFP vector as described previously (Nakayasu et al. 2018), generating the T1/T2-pMgP237 vector. *Agrobacterium rhizogenes* ATCC15834 was transformed with the pMgP237 empty vector or the T1/T2-pMgP237 vector.

Induction of hairy roots was performed as reported previously (Hirota et al. 2012), with slight modifications. Seeds of *L. japonicus* were sterilized and placed onto a 0.8% agar plate. The seeds were allowed to germinate at 23°C for 5 d under dark conditions and for 2 d under 16 h light conditions. *Agrobacterium rhizogenes* strains were cultured on YEB plates for 2 d and suspended in sterilized water. The roots of 7-day-old seedlings were cutoff, and *A. rhizogenes* was infected into the cross-sections of hypocotyls. After co-cultivation for 4 d and hairy root elongation (HRE) for 2 weeks, hairy roots were dissected and cultured on fresh cefotaxime-containing HRE agar medium for 2 weeks. After dissection, hairy roots were cultured under dark conditions. Hairy roots exhibiting green fluorescent protein (GFP) fluorescence were transferred to 5 ml of HRE liquid medium without antibiotics. Isolated hairy roots were cultured for 2 months at room temperature with subculturing every 3–4 weeks. Finally, hairy roots were cultured in 100 ml of HRE liquid medium at 25°C with shaking at 90 rpm for 4 weeks.

Genomic DNA was extracted from the hairy root lines using the NucleoSpin Plant II Kit (Macherey-Nagel, Düren, Germany) following the manufacturer's instructions. Mutagenesis was confirmed by PCR with primer nos. 25 and 26 (Supplementary Table S4), and electrophoresis using an MCE-202 MultiNA microchip electrophoresis system (Shimadzu, Kyoto, Japan) following the manufacturer's instructions. The target sequences amplified from putative mutants were cloned into the pJET1.2/blunt vector (CloneJET PCR Cloning Kit; Thermo Fisher Scientific). Insertion and deletion mutations were confirmed by sequencing of several randomly selected clones.

Ljcyp93e1 loss-of-function mutant lines

Two independent mutant lines were cultivated on soil and crude genomic DNA was extracted from their leaves. Homozygous and heterozygous *Ljcyp93e1* mutants were screened by PCR using *LjCYP93E1*-specific primers and a *LORE1*-specific primer (primer nos. 27–41 in Supplementary Table S4). Crude genomic DNA extraction and PCR were performed using KOD FX Neo following the manufacturer's instructions (Toyobo, Osaka, Japan).

Triterpenoids were extracted from the homozygous (leaves, stems and roots) and heterozygous (leaves) mutants and were analyzed as described below.

The expressions of *LjCYP93E1*, *Ljcyp93e1* and *OSC1* transcripts in WT and the progeny of the #X1 line (2-week-old plants) were analyzed by qPCR using primer nos. 15, 16, 19–22, as described above.

Extraction of triterpenoids from *L. japonicus* WT and mutant plants

Plants at the flowering stage and hairy roots were lyophilized and powdered using a multibead shocker (Yasui Kikai, Osaka, Japan).

GC-MS analysis. Powdered tissues (10.00 ± 0.15 mg) were extracted three times with 1 ml of methanol by a sonication-assisted method. Completely dry extracts were resuspended in 2 ml of 50% MeOH and 2 M HCl. The extracts were incubated at 80°C for 1 h to remove the sugar moieties of triterpenoid saponins. The hydrolyzed products were extracted three times with hexane: EtOAc (1:1) and dried completely. The obtained pellet was resuspended in 300 µl of MeOH: chloroform (1:1). A portion of the solution was dried in a GC-MS vial. Derivatization was performed as described above. For quantitative analysis, an asiatic acid authentic standard was applied to the plant tissue powder before extraction. To analyze nonhydrolyzed samples, the heating step was omitted.

LC-MS analysis. Powdered tissues (10.00 ± 0.15 mg) were extracted in 1 ml of methanol with sonication for 1 h. The samples were centrifuged, and the supernatants were filtered through a GL Chromato Disk 4 A (0.2 µm; GL Sciences, Tokyo, Japan).

GC-MS analysis

GC-MS analyses were performed on a 5977A MSD mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) connected to a 7890B gas chromatograph (Agilent Technologies) with a DB-1MS (30 m × 0.25 mm, 0.25 µm film thickness; Agilent Technologies) capillary column for quantitative analysis and with an HP-5MS UI (30 m × 0.25 mm, 0.25 µm film thickness; Agilent Technologies) capillary column for qualitative analysis. The injection temperature was set at 250°C. The column temperature program was as follows: 80°C for 1 min, an increase to 300°C at a rate of 20°C·min^{-1}, and hold for 28 min. The carrier gas was helium at a flow rate of 1.0 ml·min^{-1}. The ion source temperature was 230°C and the quadrupole temperature was 150°C. One microliter of the derivatized sample was injected in splitless injection mode. Peaks were identified by comparing their Rt and mass spectra with those of authentic standards. Samples were analyzed in SIM mode for quantitation (Supplementary Table S5), and in scan mode for qualitative analysis (m/z = 50–850).

LC-MS analysis

Ultra-performance LC-MS (UPLC-MS) analyses were performed using the Acquity Ultra-Performance LC-system with a tandem quadruple detector (Waters, Milford, MA, USA). The sample room was maintained at 15 ± 5°C. Five microliters of each sample were injected for analysis. To separate compounds, an Acquity UPLC BEH C18 column (150 × 2.1 mm, 1.7 µm particle size; Waters) maintained at 60°C was used. The mobile phase consisted of water with 0.05% formic acid (solvent A) and 100% acetonitrile (solvent B). The flow rate was set at 0.56 ml·min^{-1}. The mobile-phase gradient was as follows: 95% A at 0 min, 30% A at 30 min, 5% A at 33 min, 5% A at 36 min, 95% A at 36.1 min and 95% A at 39 min. Compounds were ionized by the electrospray ionization method in negative ion mode. Mass spectra were recorded in the range of 350–2,000 m/z. The settings of the mass spectrometer were as follows: capillary voltage, +3.0 keV; cone voltage, 80 V; source temperature, 150°C; desolvation temperature, 450°C; cone gas flow, 50 l·h^{-1}; and desolvation gas flow, 850 l·h^{-1}. MassLynx (ver. 4.1) software (Waters) was used for data acquisition and analysis.

Phylogenetic analysis

Full-length amino acid sequences were collected from GenBank and miyakogusa.jp (http://www.kazusa.or.jp/lotus/) and were aligned using MUSCLE software (Edgar 2004). A phylogenetic tree was generated by the neighbor-joining method with 1,000 replicates using MEGA6 software (Tamura et al. 2013).

Supplementary Data

Supplementary data are available at PCP online.

Funding

The Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan [JSPS KAKENHI JP17H05442 to H.Se., JP19H02921 to T.M., JP19J10245 to H.Su.; the Frontier Research Base for Global Young Researchers, Osaka University, to E.O.F.].

Acknowledgments

We thank Dr. Kiyoshi Ohyama (Tokyo Institute of Technology) for providing sophoradiol, 24-hydroxy-β-amyrin and soyasapogenol E standards. We also thank Miyazaki University for providing seeds of *L. japonicus* Gifu B-129 accession (WT) and the cDNA clone (for cloning CYP716A51), and Aarhus University for providing seeds of *L. japonicus* Gifu B-129 accession (*LORE1* insertion lines) through the Japanese National BioResource Project.

Disclosures

The authors have no conflicts of interest to declare.

References

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. (2000) Generation of 7137 non-redundant expressed sequence tags from a legume, *Lotus japonicus*. *DNA Res*. 7: 127–130.

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. (2004) Characteristics of the *Lotus japonicus* gene repertoire deduced from large-scale expressed sequence tag (EST) analysis. *Plant Mol. Biol*. 54: 405–414.

Biazzi, E., Carelli, M., Tava, A., Abbruscato, P., Losini, I., Avato, P., et al. (2015) CYP72A67 catalyzes a key oxidative step in *Medicago truncatula* hemolytic saponin biosynthesis. *Mol. Plant* 8: 1493–1506.

Brendolise, C., Yauk, Y.K., Eberhard, E.D., Wang, M., Chagne, D., Andre, C., et al. (2011) An unusual plant triterpene synthase with predominant α-amyrin-producing activity identified by characterizing oxidosqualene cyclases from *Malus* × *domestica*. *FEBS J*. 278: 2485–2499.

Carelli, M., Biazzi, E., Panara, F., Tava, A., Scaramelli, L., Porceddu, A., et al. (2011) *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. *Plant Cell* 23: 3070–3081.

Chabaud, M., Ratet, P., de Sousa Araújo, S., Roldão Lopes Amaral Duque, A.S., Harrison, M. and Barker, D.G. (2007) *Agrobacterium tumefaciens*-mediated transformation and in vitro plant regeneration of *M. truncatula*. Medicago Truncatula Handbook. https://www.noble.org/globalassets/docs/medicago-handbook/agrobacterium-tumefaciens.pdf (5 August 2019, date last accessed).

Confalonieri, M., Cammareri, M., Biazzi, E., Pecchia, P., Fevereiro, M.P.S., Balestrazzi, A., et al. (2009) Enhanced triterpene saponin biosynthesis and root nodulation in transgenic barrel medic (*Medicago truncatula* Gaertn.) expressing a novel β-amyrin synthase (*AsOXA1*) gene. *Plant Biotechnol. J*. 7: 172–182.

Delis, C., Krokida, A., Georgiou, S., Peña-Rodríguez, L.M., Kavroulakis, N., Ioannou, E., et al. (2011) Role of lupeol synthase in *Lotus japonicus* nodule formation. *New Phytol*. 189: 335–346.

Doma, M.K. and Parker, R. (2007) RNA quality control in eukaryotes. *Cell* 131: 660–668.

Du, H., Ran, F., Dong, H.L., Wen, J., Li, J.N. and Liang, Z. (2016) Genome-wide analysis, classification, evolution, and expression analysis of the cytochrome P450 93 family in land plants. *PLoS One* 11: e0165020.

Ebizuka, Y., Shibuya, M. and Wakita, E. (2011) C-22 hydroxylase. US Patent Application No. 12/690,732. http://appft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch-bool.html&r=16&f=G&l=50&co1=AND&d=PG01&s1=Ebizuka&OS=Ebizuka&RS=Ebizuka (5 August 2019, date last accessed).

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and highthroughput. *Nucleic Acids Res*. 32: 1792–1797.

Fenwick, G.R., Price, K.R., Tsukamoto, C. and Okubo, K. (1991) Saponins. *In* Toxic Substances in Crop Plants. Edited by D'Mello, J.P.F., Duffus, C.M. and Duffus, J.H. pp. 285–327. The Royal Society of Chemistry, Cambridge.

Finer, J.J. and McMullen, M.D. (1991) Transformation of soybean via particle bombardment of embryogenic suspension culture tissue. *In Vitro Cell. Dev. Biol. Plant* 27: 175–182.

Fukai, E., Soyano, T., Umehara, Y., Nakayama, S., Hirakawa, H., Tabata, S., et al. (2012) Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon *LORE1*. *Plant J*. 69: 720–730.

Fukushima, E.O., Seki, H., Ohyama, K., Ono, E., Umemoto, N., Mizutani, M., et al. (2011) CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. *Plant Cell Physiol*. 52: 2050–2061.

Fukushima, E.O., Seki, H., Sawai, S., Suzuki, M., Ohyama, K., Saito, K., et al. (2013) Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. *Plant Cell Physiol*. 54: 740–749.

Ghosh, S. (2017) Triterpene structural diversification by plant cytochrome P450 enzymes. *Front. Plant Sci*. 8: 1886.

Hashimoto, R., Ueta, R., Abe, C., Osakabe, Y. and Osakabe, K. (2018) Efficient multiplex genome editing induces precise, and self-ligated type mutations in tomato plants. *Front. Plant Sci*. 9: 916.

Hirota, A., Shimoda, Y. and Imaizumi-Anraku, H. (2012) Protocol for *Lotus japonicus* transformation. *In* Protocols for Plant Transformation. Edited by Tabei, Y. pp. 348–361. Kagaku Dojin, Kyoto, Japan.

Huhman, D.V., Berhow, M.A. and Sumner, L.W. (2005) Quantification of saponins in aerial and subterranean tissues of *Medicago truncatula*. *J. Agric. Food Chem*. 53: 1914–1920.

Iturbe-Ormaetxe, I., Haralampidis, K., Papadopoulou, K. and Osbourn, A. (2003) Molecular cloning and characterization of triterpene synthases from *Medicago truncatula* and *Lotus japonicus*. *Plant Mol. Biol*. 51: 731–743.

Krishnamurthy, P., Fujisawa, Y., Takahashi, Y., Abe, H., Yamane, K., Mukaiyama, K., et al. (2019) High throughput screening and characterization of a high-density soybean mutant library elucidate the biosynthesis pathway of triterpenoid saponins. *Plant Cell Physiol*. 60: 1082–1097.

Krokida, A., Delis, C., Geisler, K., Garagounis, C., Tsikou, D., Peña-Rodríguez, L.M., et al. (2013) A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. *New Phytol*. 200: 675–690.

Lee, S.B. and Suh, M.C. (2015) Advances in the understanding of cuticular waxes in *Arabidopsis thaliana* and crop species. *Plant Cell Rep*. 34: 557–572.

Mertens, J., Pollier, J., Bossche, R.V., Lopez-Vidriero, I., Franco-Zorrilla, J.M. and Goossens, A. (2016) The bHLH transcription factors TSAR1 and TSAR2 regulate triterpene saponin biosynthesis in *Medicago truncatula*. *Plant Physiol*. 170: 194–210.

Miettinen, K., Pollier, J., Buyst, D., Arendt, P., Csuk, R., Sommerwerk, S., et al. (2017) The ancient CYP716 family is a major contributor to the diversification of eudicot triterpenoid biosynthesis. *Nat. Commun*. 8: 14153.

Misra, R.C., Sharma, S., Garg, A., Chanotiya, C.S. and Ghosh, S. (2017) Two CYP716A subfamily cytochrome P450 monooxygenases of sweet basil play similar but nonredundant roles in ursane- and oleanane- type pentacyclic triterpene biosynthesis. *New Phytol*. 214: 706–720.

Moses, T., Pollier, J., Almagro, L., Buyst, D., Van Montagu, M., Pedreño, M.A., et al. (2014a) Combinatorial biosynthesis of sapogenins and saponins in *Saccharomyces cerevisiae* using a C-16α hydroxylase from *Bupleurum falcatum*. *Proc. Natl. Acad. Sci. USA* 111: 1634–1639.

Moses, T., Pollier, J., Shen, Q., Soetaert, S., Reed, J., Erffelinck, M.L., et al. (2015) OSC2 and CYP716A14v2 catalyze the biosynthesis of triterpenoids for the cuticle of aerial organs of *Artemisia annua*. *Plant Cell* 27: 286–301.

Moses, T., Thevelein, J.M., Goossens, A. and Pollier, J. (2014b) Comparative analysis of CYP93E proteins for improved microbial synthesis of plant triterpenoids. *Phytochemistry* 108: 47–56.

Naito, Y., Hino, K., Bono, H. and Ui-Tei, K. (2015) CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* 31: 1120–1123.

Nakamura, M., Linh, T.M., Lien, L.Q., Suzuki, H., Mai, N.C., Giang, V.H., et al. (2018) Transcriptome sequencing and identification of cytochrome P450 monooxygenases involved in the biosynthesis of maslinic acid and corosolic acid in *Avicennia marina*. *Plant Biotechnol*. 35: 341–348.

Nakayasu, M., Akiyama, R., Lee, H.J., Osakabe, K., Osakabe, Y., Watanabe, B., et al. (2018) Generation of α-solanine-free hairy roots of potato by CRISPR/Cas9 mediated genome editing of the *St16DOX* gene. *Plant Physiol. Biochem*. 131: 70–77.

Naoumkina, M.A., Modolo, L.V., Huhman, D.V., Urbanczyk-Wochniak, E., Tang, Y., Sumner, L.W., et al. (2010) Genomic and coexpression analyses

predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. *Plant Cell* 22: 850–866.

Pütter, K.M., van Deenen, N., Müller, B., Fuchs, L., Vorwerk, K., Unland, K., et al. (2019) The enzymes OSC1 and CYP716A263 produce a high variety of triterpenoids in the latex of *Taraxacum koksaghyz*. *Sci. Rep.* 9: 5942.

Rehman, H.M., Nawaz, M.A., Shah, Z.H., Yang, S.H. and Chung, G. (2018) Functional characterization of naturally occurring wild soybean mutant (*sg-5*) lacking astringent saponins using whole genome sequencing approach. *Plant Sci*. 267: 148–156.

Sandeep, Misra, R.C., Chanotiya, C.S., Mukhopadhyay, P. and Ghosh, S. (2019) Oxidosqualene cyclase and CYP716 enzymes contribute to triterpene structural diversity in the medicinal tree banaba. *New Phytol*. 222: 408–424.

Sander, J.D. and Joung, J.K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32: 347.

Sato, S., Kaneko, T., Nakamura, Y., Asamizu, E., Kato, T. and Tabata, S. (2001) Structural analysis of a *Lotus japonicus* genome. I. Sequence features and mapping of fifty-six TAC clones which cover the 5.4 Mb regions of the genome. *DNA Res*. 8: 311–318.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., et al. (2008) Genome structure of the legume, *Lotus japonicus*. *DNA Res.* 15: 227–239.

Sawai, S., Akashi, T., Sakurai, N., Suzuki, H., Shibata, D., Ayabe, S., et al. (2006a) Plant lanosterol synthase: divergence of the sterol and triterpene biosynthetic pathways in eukaryotes. *Plant Cell Physiol*. 47: 673–677.

Sawai, S., Shindo, T., Sato, S., Kaneko, T., Tabata, S., Ayabe, S., et al. (2006b) Functional and structural analysis of genes encoding oxidosqualene cyclases of *Lotus japonicus*. *Plant Sci*. 170: 247–257.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178.

Seki, H., Ohyama, K., Sawai, S., Mizutani, M., Ohnishi, T., Sudo, H., et al. (2008) Licorice β-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc. Natl. Acad. Sci. USA* 105: 14204–14209.

Seki, H., Tamura, K. and Muranaka, T. (2015) P450s and UGTs: key players in the structural diversity of triterpenoid saponins. *Plant Cell Physiol.* 56: 1463–1471.

Shibuya, M., Hoshino, M., Katsube, Y., Hayashi, H., Kushiro, T. and Ebizuka, Y. (2006) Identification of beta-amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. *FEBS J*. 273: 948–959.

Srisawat, P., Fukushima, E.O., Yasumoto, S., Robertlee, J., Suzuki, H., Seki, H., et al. (2019) Identification of oxidosqualene cyclases from the medicinal legume tree *Bauhinia forficata*: a step toward discovering preponderant α-amyrin-producing activity. *New Phytol*. 224: 352–366.

Suzuki, H., Fukushima, E.O., Umemoto, N., Ohyama, K., Seki, H. and Muranaka, T. (2018) Comparative analysis of CYP716A subfamily enzymes for the heterologous production of C-28 oxidized triterpenoids in transgenic yeast. *Plant Biotechnol*. 35: 131–139.

Tadege, M., Wen, J., He, J., Tu, H., Kwak, Y., Eschstruth, A., et al. (2008) Large-scale insertional mutagenesis using the *Tnt1* retrotransposon in the model legume *Medicago truncatula*. *Plant J.* 54: 335–347.

Takagi, K., Nishizawa, K., Hirose, A., Kita, A. and Ishimoto, M. (2011) Manipulation of saponin biosynthesis by RNA interference-mediated silencing of β-amyrin synthase gene expression in soybean. *Plant Cell Rep*. 30: 1835–1846.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*. 30: 2725–2729.

Tamura, K., Teranishi, Y., Ueda, S., Suzuki, H., Kawano, N., Yoshimatsu, K., et al. (2017) Cytochrome P450 monooxygenase CYP716A141 is a unique β-amyrin C-16β oxidase involved in triterpenoid saponin biosynthesis in *Platycodon grandiflorus*. *Plant Cell Physiol*. 58: 874–884.

Tamura, K., Yoshida, K., Hiraoka, Y., Sakaguchi, D., Chikugo, A., Mochida, K., et al. (2018) The basic helix-loop-helix transcription factor GubHLH3 positively regulates soyasaponin biosynthetic genes in *Glycyrrhiza uralensis*. *Plant Cell Physiol*. 59: 783–796.

Tava, A. and Avato, P. (2006) Chemical and biological activity of triterpene saponins from *Medicago* species. *Nat. Prod. Commun*. 1: 1159–1180.

Tava, A., Biazzi, E., Mella, M., Quadrelli, P. and Avato, P. (2017) Artefact formation during acid hydrolysis of saponins from *Medicago* spp. *Phytochemistry* 138: 116–127.

Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P. and Osbourn, A. (2014) Triterpene biosynthesis in plants. *Annu. Rev. Plant Biol.* 65: 225–257.

Tsuno, Y., Fujimatsu, T., Endo, K., Sugiyama, A. and Yazaki, K. (2018) Soyasaponins, a new class of root exudates in soybean (*Glycine max*). *Plant Cell Physiol.* 59: 366–375.

Urbański, D.F., Małolepszy, A., Stougaard, J. and Andersen, S.U. (2012) Genome-wide *LORE1* retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. *Plant J*. 69: 731–741.

Vo, N.N., Fukushima, E.O. and Muranaka, T. (2017) Structure and hemolytic activity relationships of triterpenoid saponins and sapogenins. *J. Nat. Med.* 71: 50–58.

Wang, L., Wang, L., Tan, Q., Fan, Q., Zhu, H., Hong, Z., et al. (2016) Efficient inactivation of symbiotic nitrogen fixation related genes in *Lotus japonicus* using CRISPR-Cas9. *Front. Plant Sci*. 7: 1333.

Yano, R., Takagi, K., Takada, Y., Mukaiyama, K., Tsukamoto, C., Sayama, T., et al. (2017) Metabolic switching of astringent and beneficial triterpenoid saponins in soybean is achieved by a loss-of-function mutation in cytochrome P450 72A69. *Plant J*. 89: 527–539.

Yasumoto, S., Fukushima, E.O., Seki, H. and Muranaka, T. (2016) Novel triterpene oxidizing activity of *Arabidopsis thaliana* CYP716A subfamily enzymes. *FEBS Lett.* 590: 533–540.

Yasumoto, S., Seki, H., Shimizu, Y., Fukushima, E.O. and Muranaka, T. (2017) Functional characterization of CYP716 family P450 enzymes in triterpenoid biosynthesis in tomato. *Front. Plant Sci*. 8: 21.

Young, N.D., Debellé, F., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K., et al. (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480: 520–524.