Plant Biotechnology 38, 127–135 (2021)
DOI: 10.5511/plantbiotechnology.21.0112a

Original Paper —Invited Paper—

Development of selection method for *Glycyrrhiza uralensis* superior clones with high-glycyrrhizic acid contents using DNA sequence polymorphisms in glycyrrhizic acid biosynthetic genes

Takayuki Inui[1,a], Noriaki Kawano[1], Daisuke Araho[2], Yukiyoshi Tamura[2], Nobuo Kawahara[1], Kayo Yoshimatsu[1,*]

[1] Tsukuba Division, Research Center for Medicinal Plant Resources, National Institutes of Biomedical Innovation, Health and Nutrition, 1-2 Hachimandai, Tsukuba, Ibaraki 305-0843, Japan; [2] Research Center, Maruzen Pharmaceuticals Co., Ltd., 1089-8 Sagata, Shinnichi-Cho, Fukuyama, Hiroshima 729-3102, Japan
* E-mail: yoshimat@nibiohn.go.jp Tel: +81-29-838-0573 Fax: +81-29-838-0575

Received December 11, 2020; accepted January 12, 2021 (Edited by T. Kobayashi)

Abstract *Glycyrrhiza* plants are important resources for sweeteners and medicines, because underground parts of them contain glycyrrhizic acid (GL), which has sweet taste and various pharmacological activities (ex. anti-inflammatory, antiallergy, antiviral activity, etc.). Although such importance of them, their supply still depends principally on the collection of wild plants. Therefore, it is an important issue to develop stable and efficient production system of *Glycyrrhiza* plants. To overcome this problem, we established the hydroponic cultivation system of *Glycyrrhiza uralensis* and selected superior *G. uralensis* clones with high-GL contents in the containment greenhouse. In this study, we aimed to develop a method of selecting these superior *G. uralensis* clones by DNA sequence polymorphisms in biosynthetic genes. Among the DNA sequences of GL biosynthetic key enzyme gene (*CYP88D6*), we found *Glycyrrhiza* species and clone-specific polymorphisms in intronic regions. By using these polymorphisms, discrimination among *Glycyrrhiza* species and *G. uralensis* clones became possible. Furthermore, the appearance frequency of superior clone-specific alleles in cloned *CYP88D6* sequences was correlated with GL contents in crude drugs collected from the Japanese market. We also observed the tendency that *G. uralensis* seedlings having superior clone-specific alleles of *CYP88D6* gene showed higher secondary metabolite productivity than those without the alleles. These results indicated that superior clone-specific alleles of *CYP88D6* gene could be applied as DNA markers for selecting *G. uralensis* clones accumulating high secondary metabolites.

Key words: DNA sequence polymorphisms, *Glycyrrhiza uralensis*, glycyrrhizic acid, selection, superior clone.

Introduction

Glycyrrhiza is a perennial plant belonging to *Leguminosae*. Extract of *Glycyrrhiza* species and their constituents have wide bioactivities, such as anti-inflammatory, antioxidative, antimicrobial, antiviral, antitumor, hepatoprotective and neuroprotective effects, etc (Asl and Hosseinzadeh 2008; Hosseinzadeh and Asl 2015). Glycyrrhizae Radix (the dried roots and stolons of *Glycyrrhiza uralensis* Fisher or *G. glabra* Linne) (The Ministry of Health, Labour and Welfare 2016) is the most frequently prescribed crude drug in Japanese traditional medicine (Kampo formulations). Furthermore, its extracts are also used as cosmetics and food additives because their major components, glycyrrhizic acid (GL) has anti-inflammatory effect and sweetness (Hayashi and Sudo 2009).

In Japan, supply of Glycyrrhizae Radix mostly depends on imported material, which has been mainly prepared from wild plants. With increasing demand of Glycyrrhizae Radix, there is a concern of depletion of superior resources and environmental destruction due to over-exploitation of licorice plants. Therefore, it is an important issue to develop stable and efficient production system of *Glycyrrhiza* plants. To overcome this problem, numerous studies have been performed on the selection

Abbreviations: GC, glycycoumarin; GL, glycyrrhizic acid; IL, isoliquiritin; ITS, internal transcribed spacer on nuclear ribosomal DNA; LQ, liquiritin; MPDB, Comprehensive Medicinal Plant Database; MS, Murashige and Skoog; NIBIOHN, National Institutes of Biomedical Innovation, Health and Nutrition; RCMPR, Research Center for Medicinal Plant Resources.

[a] Present address: Hokkaido Division, Research Center for Medicinal Plant Resources, National Institutes of Biomedical Innovation, Health and Nutrition, 108-4 Ohashi, Nayoro, Hokkaido 096-0065, Japan

This article can be found at http://www.jspcmb.jp/
Published online March 18, 2021

of superior clones and the method of cultivation and proliferation (Kojoma et al. 2011; Ozaki et al. 2010; Toda et al. 2012). We also established the hydroponic cultivation system of *G. uralensis* and selected superior *G. uralensis* clones with high-GL contents in the containment greenhouse (Yoshimatsu et al. 2015, 2017, 2018).

To protect these superior clones from unauthorized uses, it is important to distinguish them from other clones. In *Glycyrrhiza*, it is reported that species identification using DNA polymorphism in the internal transcribed spacer on nuclear ribosomal DNA (ITS), matK, and other regions (Kondo et al. 2007). However, these regions are conserved and might be unsuitable for discrimination of superior clones. On the other hand, it is expected that various polymorphisms useful for discrimination are accumulated in intron regions of biosynthetic genes and several genes involved in GL biosynthesis were isolated from *G. uralensis* plants (Lu et al. 2008; Seki et al. 2008, 2011). In this study, we accumulated polymorphisms in intronic regions of GL biosynthetic genes and attempted to discriminate among *Glycyrrhiza* species and *G. uralensis* clones. Furthermore, we examined the correlation between appearance frequency of alleles and GL contents in crude drugs, aimed to develop a method of high-GL superior *G. uralensis* clones by DNA sequence polymorphisms.

Materials and methods

Plant material

The leaves of *G. uralensis* Fisch., *G. glabra* L., *G. echinata* L. and *G. pallidiflora* Maxim. were collected in the field of the Tsukuba Division, Research Center for Medicinal Plant Resources (RCMPR), National Institutes of Biomedical Innovation, Health and Nutrition (NIBIOHN) and the leaves of *G. inflata* Bat. were provided from Dr. Susumu Isoda, Showa University. Accession numbers of field-grown *Glycyrrhiza* plants (*G. uralensis*, *G. glabra, G. echinata* and *G. pallidiflora*) are TS0125-93, TS0469-79, TS0451-79 and TS0330-80, respectively (Table 1). Glycyrrhizae Radixes traded in the Japanese market were collected with the cooperation of Japan Kampo Medicines Manufacturers Association. Their accession numbers are listed in the Table 1 and their detailed information (origin, ITS sequence, etc.) is disclosed at our Comprehensive Medicinal Plant Database (MPDB) (Research Center for Medicinal Plant Resources 2013).

Clonal propagation of G. uralensis

G. uralensis (accession number TS71-08) seeds were collected from field-grown plants in Hokkaido division of RCMPR, NIBIOHN. The seeds were surface-sterilized in 75% (v/v) ethanol for 1 min, washed with sterile water and sterilized in a 2% (v/v) sodium hypochlorite solution containing 0.1% (v/v) Tween-20 for 10 min, followed by three washes with sterile water. The surface-sterilized seeds were germinated aseptically at 23°C under a 14-h light photoperiod (*ca.* 70 μmol m^{-2} s^{-1}) on a half strength macro elements Murashige and Skoog medium (Murashige and Skoog 1962) (MS; Duchefa M0233) containing 2% (w/v) sucrose and solidified with 0.25% (w/v) gelrite (San-Ei Gen FFI). Nodal segments (*ca.* 1 cm) of germinated plants were transferred into a MS medium containing 1% sucrose, 0.5gl^{-1} 2-(N-morpholino) ethanesulfonic acid 0.1 mg l^{-1} indole-3-butyric acid and solidified with 0.25% (w/v) gelrite. Nodal segments were cut and subcultured to fresh medium every 2 months under the same culture condition mentioned above.

Hydroponic cultivation of G. uralensis clones

G. uralensis clones propagated by tissue culture were hydroponically cultivated using Otsuka-A nutrient solution (OAT Agrio Co., Ltd.) diluted to 4 times, under a 16-h photoperiod (450–650 μmol m^{-2} s^{-1}) at 25°C, 60% relative humidity in the growth chamber. Nutrient solution was supplied through an underground irrigation system using pumice (Extra Small, Ohe Chemicals Inc.).

DNA amplification and Sequencing

Genomic DNA was extracted from crude drugs and leaves of field or hydroponically cultivated *Glycyrrhiza* plants, using DNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions.

PCR reaction mix consist of 0.5 μl of KOD plus (TOYOBO), 5 μl of 10×KOD buffer, 5 μl of 2 mM each dNTPs, 2 μl of 25 mM $MgSO_4$, 1 μl of 10 μM forward primer, 1 μl of 10 μM reverse primer, 34.5 μl of distilled water and 1 μl of genomic DNA (10 ng μl^{-1}). PCR condition for exon 4 to exon 6 region of *CYP88D6* homologous genes was as follows: 94°C for 2 min, then 30 cycles of 94°C for 30 s, 58°C for 30 s, 68°C for 60 s, and finally 68°C for 5 min. forward primer was CYP88D6-S1006: 5′-aggacaagattttcctcgcag-3′ and reverse primer was CYP88D6-A1148: 5′-agcactctccatccctttgg-3′. PCR condition for exon 6 to exon 8 region of *CYP88D6* homologous genes was as follows: 94°C for 2 min, then 40 cycles of 94°C for 30 s, 58°C for 30 s, 68°C for 90 s, and finally 68°C for 5 min. forward primer was CYP88D6-S1179: 5′-gtgctaatttgggcaagagc-3′ and reverse primer was CYP88D6-A1420: 5′-agctggtaacgtgacattctgg-3′. The presence of PCR products was confirmed by 1% agarose gel electrophoresis.

Amplified DNA fragments were ligated into the pT7Blue T-vector (Novagen) with DNA ligation kit ver 2.1 (TaKaRa BIO) after the addition of 3′-A overhangs by 10×A-Attachment Mix (TOYOBO). Then, the plasmids were transformed into DH5α *E. coli* cells (Z-Competent ™ Cells, Zymo Research). Transformed *E. coli* cells were grown in LB medium containing 100 μg ml^{-1} ampicillin at 37°C and plasmid DNA was isolated using Illustra Plasmid Prep Mini Spin kit (GE Healthcare Bioscience). The inserted fragments were amplified from plasmids using the vector-specific primers (U-19: 5′-gttttcccagtcacgacgt-3′ and R-20: 5′-cagctatgaccatgattacg-3′, Novagen), and sequenced with the same primers using a Big Dye Terminator Cycle Sequencing Kit 3.1 on ABI PRISM 3100-Avant genetic Analyzer

Table 1. Polymorphisms in partial sequences of intron 7 (85-302) of *CYP88D6* homologous genes.

		Accession number	*CYP88D6* intron 7																														Genotype
			112	116	128–130	134	139	141	142	147	157	163	179	181	186	206	223	224	227	234	236	241	243	244	261	262	273	293	295	296	301–302	
Glycyrrhiza plants	*G. uralensis* Fisch.	TS0125-93	A	G	AAT	G	T	T	A	A	G	T	A	T	C	G	T	G	T	C	T	−	A	G	C	G	G	T	T	G	−−	u
	G. glabra L.	TS0469-79	*	*	***	−	C	C	*	*	C	*	G	*	*	*	C	*	C	T	*	T	G	*	*	A	*	*	*	*	**	g
	G. inflata Bat.		*	*	***	*/−	*/C	*/C	*	*/G	*/C	*	*/G	*	*	*	*/C	*	*/C	*/T	*	*/T	*/G	*	*	*	*/A	*	*/C	*	**	u, g
	G. pallidiflora Maxim.	TS0330-80	*	*	−−−/***	*	*	*	*	*	C	G	*	*	*	*	*	A	*	*	*	*	*	*	*	*	A	*	C	*	**	p
	G. echinata L.	TS0451-79	*	*	***	*	*	*	*	*	C	G	*	*	*	*	*	*	*	*	*	*	*	*	*	*	A	C	C	*	**	e
G. uralensis superior clones	Gu2-3-2		*	*	***	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**	u
	GuIV1		*	*	***	*	*	*	*	*	*	*	*/G	*	*	*/T	*	*	*	*	*	*	*	*	*/T	*	*/A	*	*/C	*	**/TA	IV1
	GuIV2		*	*	***	*	*	*	*	*	*	*	*/G	*	*/A	*	*	*	*	*	*	*	*	*/C	*	*	*/A	*	*/C	*/A	**	IV2
Crude drugs		NIB004	*/C	*	***	*/−	*/C	*/C	*	*	*/C	*	*/G	*	*/A	*/T	*/C	*	*/C	*/T	*	*/T	*/G	*/C	*/T	*/A	*/A	*	*/C	*/A	**/TA	IV1, IV2, g
		NIB005	*	*/A	***	*/−	*	*/C	*/G	*	*/C	*	*/G	*/A	*/A	*	*/C	*	*/C	*/T	*	*/T	*	*/C	*	*	*/A	*	*/C	*/A	**	IV2, g
		NIB006	*	*/A	***	*/−	*	*/C	*/G	*	*/C	*	*/G	*	*/A	*/T	*/C	*	*/C	*/T	*	*/T	*	*/C	*/T	*	*/A	*	*/C	*/A	**/TA	IV1, IV2, g
		NIB007	*	*	***	*	*	*	*	*	*	*	*/G	*/A	*/A	*/T	*	*	*	*	*	*	*	*/C	*/T	*	*/A	*	*/C	*/A	**/TA	IV1, IV2
		NIB074	*	*/A	***	*/−	*	*/C	*/G	*	*/C	*	*/G	*	*	*/T	*/C	*	*/C	*/T	*	*/T	*	*	*/T	*	*/A	*	*/C	*	**/TA	IV1, g
		NIB107	*	*	***	*	*	*	*	*	*	*	*/G	*	*/A	*/T	*	*	*	*	*/C	*	*	*/C	*/T	*	*/A	*	*/C	*/A	**/TA	IV1, IV2
		NIB146	*	*	***	*	*	*	*	*	*	*	*/G	*	*/A	*/T	*	*	*	*	*	*	*	*/C	*/T	*	*/A	*	*/C	*/A	**/TA	IV1, IV2
		NIB176	*	*	***	*	*	*	*	*	*	*	*/G	*/A	*/A	*/T	*	*	*	*	*	*	*	*/C	*/T	*	*/A	*	*/C	*/A	**/TA	IV1, IV2

Asterisk (*) means identical to *G. uralensis* (TS0125-93) sequence and minus (−) means deletion.

or 3130 Genetic Analyzer (Applied Biosystems) following the manufacturer's instructions.

PCR conditions for partial sequence of intron 7 of *CYP88D6* homologous genes were as follows: 94°C for 2 min, 30 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 45 s and finally 72°C for 5 min. PCR reaction mix consist of 0.1 μl of ExTaq (TaKaRa BIO), 2 μl of 10×ExTaq buffer, 1.6 μl of 2.5 mM each dNTPs, 1 μl of 10 μM forward primer, 1 μl of 10 μM reverse primer, 12.3 μl of distilled water and 2 μl of genomic DNA. Amplification primers were CYP88D6i7 Fw: 5′-tagtgcctttaagcacatgg-3′, CYP88D6i7 Rv: 5′-tcatcggtataattgtagcactc-3′ and CYP88D6i7 Rv2: 5′-agagatcaatcaggtagctagagag-3′.

Amplified DNA sequences were obtained by direct and/or cloning sequencing. For direct sequencing, the PCR products purified with Illustra ExoProSTAR (GE Healthcare) were directly sequenced using amplification primers. For cloning, amplified DNA fragments were cloned in the same method as the above. A single *E. coli* colony was dipped into a PCR master mix consists of 5 μl of GoTaq Green Master Mix (Promega), 1 μl of 10 μM R20 primer, 1 μl of 10 μM U19 primer and 3 μl of distilled water. PCR conditions were as follows: 94°C for 2 min, then 30 cycles of 94°C for 15 s, 60°C for 30 s, 72°C for 30 s and finally 72°C for 5 min. Purified PCR products with Illustra ExoProSTAR were directly sequenced using R20 or R19 primers.

Appearance frequency (%) of *CYP88D6* intron 7 each allele in the commercial *Glycyrrhiza* crude drugs obtained from Japanese market was calculated according to the following equation.

$$\text{Appearance frequency (\%)} = \frac{\text{Number of } \textit{E.coli} \text{ clones with the allele}}{\text{Number of } \textit{E.coli} \text{ clones sequenced (8–16)}} \times 100$$

Quantitative analysis of secondary metabolites

GL, glycycoumarin (GC), liquiritin (LQ) and isoliquiritin (IL) contents of crude drugs of the Japanese market were measured as follows. A 50 mg of each powdered sample was extracted by sonication with 7.0 ml 50% ethanol for 30 min. The extract solution was filtered through an Ultrafree-MC (0.45 μm filter unit; Millipore, Bedford, MA, USA) and a 5–20 μl aliquot was analyzed by HPLC using a Waters Alliance HT HPLC system (2795 separation module and 2996 photodiode array detector; Waters). The HPLC column was a TSKgel ODS-100 V column (4.6 mm i.d. ×250 mm, 5 μm, TOSOH) and the column temperature was maintained at 40°C. The mobile phase consisted of acetonitrile (solvent A) and 1% acetic acid (solvent B). The flow rate was 1.0 ml min^{-1} and the gradient program was as follows: 0–21 min 20–76% A, 21–22 min 76–100% A, 22–24 min 100% A, 24–35 min 100–20% A. Contents were quantified by the peak area at 254 nm.

For GL, LQ and IL quantification, hydroponically cultivated plants were extracted as follows. A 200 mg of each powdered sample was extracted with 5.0 ml 50% ethanol and the extract solution was filtered through 0.2 μm membrane filter unit. For GC quantification, the plants were extracted with 5.0 ml ethanol and filtered in the same way. A 2 μl aliquot was analyzed by ACQUITY UPLC (Waters). The UPLC column was an ACQUITY UPLC BEH C18 (2.0 mm i.d. ×50 mm, 1.7 μm, Waters) and the column temperature was set at 40°C. The mobile phase consisted of 0.1% formic acid (solvent A) and acetonitrile (solvent B), and the gradient program was as follows: 0–0.45 min 20% B, 0.45–6.50 min 20–70% B, 6.50–6.75 min 70–100% B, 6.75–7.00 min 20% B at flow rate 0.8 ml min^{-1}. GL contents were quantified at 254 nm, GC at 350 nm, LQ and IL at 316 nm.

Statistical analysis

The mean±standard deviations is shown in Figures, and statistical differences in means were determined by Tukey–Kramer multiple comparison test using the statistical analysis system "R" software version 3.4.2. Different letters over the tops of columns in Figures indicate significant differences ($p<0.05$) by Tukey–Kramer's test. In multiple regression analysis, R was used for statistical computing and visualization (R Core Team 2017).

Results and discussion

Accumulation of DNA sequence polymorphisms in GL biosynthetic genes among G. uralensis superior clones

To establish the method for discrimination among *G. uralensis* superior clones, we focused on DNA sequence polymorphisms in GL biosynthetic genes, especially intronic region of *CYP88D6* gene. CYP88D6 catalyzes the oxidation at the 11-position carbon of β-amyrin and it is one of the key enzymes of GL biosynthesis (Seki et al. 2008). Generally, the exon-intron structure tends to be widely conserved. Therefore, we predicted the exon-intron structures of *G. uralensis CYP88D6* (mRNA: AB433179) gene, based on multiple alignment of *Medicago truncatula* (Mtr) and *Lotus japonica* (Lja) genomic DNA sequences and their *CYP88D6* homologous genes. Their GenBank accession numbers are AC144538 (Mtr genomic sequence), AP010409, AP007265, AP010465 (Lja genomic sequences), AB433175 (Mtr*CYP88D2* mRNA), AB433176 (Mtr*CYP88D3* mRNA), AB433177 (Lja*CYP88D4*

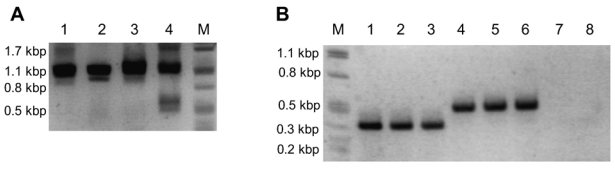

Figure 1. Discrimination among *G. uralensis* superior clones by polymorphism in *CYP88D6* intron 7. (A) PCR-RFLP; lane M: DNA size marker λ/*Pst*I; l, 4: GuIV2; 2, 3: Gu2-3-2; 1, 2: non-digested PCR products; 3, 4: *Hinc*II-digested PCR products; (B) PCR products with GuIV1 and IV2 specific primers; lane M: λ/*Pst*I; l-3: GuIV2; 4-6: GuIV1; 7-8: Gu2-3-2.

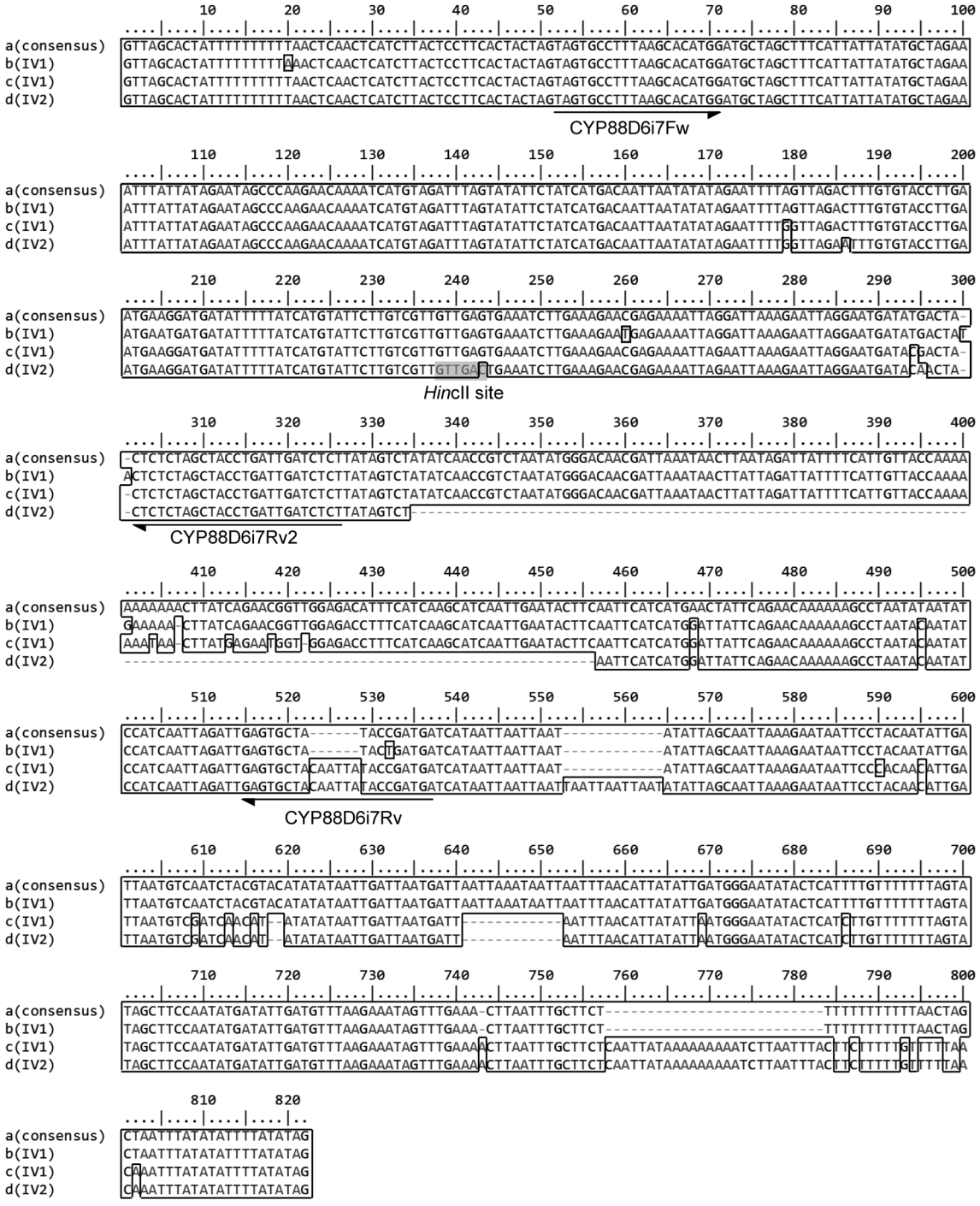

Figure 2. Alignment of *CYP88D6* intron 7 sequences amplified from *G. uralensis* superior clones. The deletions are marked by a minus sign. Identical nucleotides with consensus sequence (a) are boxed. GuIV2 specific *Hinc*II site is highlighted on a gray background.

mRNA) and AB433178 (Lja*CYP88D5* mRNA). These analyses revealed that *CYP88D6* gene consisted of 8 exons and 7 introns. We designed primers that amplified fourth and fifth intronic regions (CYP88D6-S1006 and CYP88D6-A1184) or sixth and seventh intronic regions (CYP88D6-S1179 and CYP88D6-A1420).

Firstly, we amplified intronic region of *CYP88D6* gene from *G. uralensis* superior clones (Gu2-3-2, GuIV1 and GuIV2, high GL contents more than 2.5% in their roots were successfully achieved after only one year's hydroponic cultivation, Yoshimatsu et al. 2015). PCR using primer set of CYP88D6-S1006 and A1184 amplified 0.6 kbp products and 1.1 kbp amplification products were generated by PCR using CYP88D6-S1179 and A1420 from all clones (Figure 1A, Gu2-3-2 and GuIV2). Amplification products were cloned into pT7Blue vector and analyzed the nucleotide sequence. In fourth, fifth and sixth intronic region, we found several

alleles, but they were not clone specific (date not shown). However, we found that the seventh intronic region of *CYP88D6* gene had an abundance of the clone-specific polymorphisms. Namely, Gu2-3-2 had only a common allele (Gu: Figure 2-a, 775 bp), whereas GuIV1 had the common allele (a) and two specific alleles (776 bp IV1-1: Figure 2-b and 792 bp IV1-2: Figure 2-c) and GuIV2 had one specific allele (684 bp IV2: Figure 2-d) in addition to the common allele (a). GuIV2 specific allele (d) had a G to C substitution at position-244, which generated a unique *Hinc*II restriction site. Therefore, digestion of PCR products of GuIV2 with *Hinc*II yield *ca.* 500 bp and 600 bp fragment, whereas Gu2-3-2 PCR products did not (Figure 1A, lane 3, 4). Thus, GuIV2 and Gu2-3-2 could be clearly discriminated by PCR-RFLP. Furthermore, to distinguish briefly among these superior clones, forward primer (CYP88D6i7 Fw) was designed at consensus region of intron 7 and allele (c) and (d) specific reverse primer (CYP88D6i7 Rv) was designed at specific insertion site of intron 7 (Figure 2). As the result of PCR using these primers, specific products were obtained in only GuIV1 and GuIV2 (Figure 1B). In addition, it was possible to distinguish GuIV1 from GuIV2 according to the products size (Figure 1B) because allele (d) had 122 bp deletions compared with allele (c) (Figure 2) (Yoshimatsu et al. 2017).

Accumulation of DNA sequence polymorphisms in GL biosynthetic genes among Glycyrrhiza plants and crude drugs

For further accumulation of DNA sequence polymorphisms in intron 7 of *CYP88D6* homologous genes, we extracted genomic DNA from the leaves of field cultivated *Glycyrrhiza* plants (*G. uralensis, G. glabra, G. inflata, G. pallidiflora, G. echinata*) and commercial crude drugs in Japanese market. Intron 7 partial sequences of *CYP88D6* homolog were amplified using primers designed against consensus sequence (CYP88D6i7Fw and Rv2, Figure 2) or a primer pair amplifying exon 6–8 region (CYP88D6-S1179 and A1420). PCR direct and/or cloning sequencing of *CYP88D6* homologous genes revealed that *CYP88D6* intron 7 sequence of field cultivated *G. glabra* (TS0469-79), *G. pallidiflora* (TS0330-80), *G. echinata* (TS0451-79) had species specific sequence (genotype g, p, e, respectively, in Table 1) and intron 7 sequence of *G. uralensis* (TS0125-93) (genotype u in Table 1) was identical to that of Gu2-3-2 (Table 1). These results demonstrated that *G. uralensis*, *G. glabra*, *G. pallidiflora*, and *G. echinata* could be briefly discriminated from each other by PCR direct sequencing of intron 7 of *CYP88D6* homolog. On the other hand, it was difficult to distinguish *G. inflata* from *G. uralensis* and *G. glabra* because beside *G. uralensis* similar alleles, *G. glabra* similar alleles were also detected in *G. inflata* (Table 1).

Commercial Glycyrrhizae Radix crude drug samples were firstly subjected to direct sequence analysis for partial *CYP88D6* intron 7 sequence (Table 1). Among the crude drug samples, we could detect similar alleles to GuIV1 or GuIV2 specific alleles (IV: b and c, IV2: d) in addition to the common allele (a). Glycyrrhiza Radix samples classified into three types based on the existence of allele (IV1, IV2); A: both IV1 and IV2; B: IV1; C: IV2. Furthermore, we found new alleles similar to *G. glabra* homolog (g) from some samples (Table 1). These samples were also assumed to be *G. uralensis* by the ITS sequences as well as other samples (Dr. Hayashi's data in MPDB) (Research Center for Medicinal Plant Resources 2013).

It might be suggested that ratio of each allele in the sample reflects ratio of the signal intensity (appearance frequency) of each allele in the direct sequence analysis. Therefore, appearance frequency of each allele was examined by cloning sequences. As with direct sequence, five alleles (a, b, c, d, and g) were detected, and their frequency were quite different from samples (Table 2).

Correlation of frequency of CYP88D6 intron 7 alleles and secondary metabolite contents

HPLC analysis of 50% ethanol extract revealed that GL content of commercial crude drug samples was 2.537–7.244 Dry weight % (DW%) (mean±SD=4.354±1.765 DW%). There was also significant variation with GL content of Glycyrrhizae Radix. Therefore, we examined

Table 2. Appearance frequency of *CYP88D6* intron 7 each allele in commercial *Glycyrrhiza* crude drugs obtain from Japanese market.

Accession number	Frequency of *CYP88D6* intron 7 each allele (%)					
	a Gu	b IV1-1	c IV1-2	d IV2	g	b+c+d IV1 and IV2
NIB004	56.3	6.3	0.0	25.0	12.5	31.3
NIB005	35.3	0.0	0.0	41.2	23.5	41.2
NIB006	26.7	13.3	6.7	33.3	20.0	53.3
NIB007	20.0	46.7	0.0	33.3	0.0	80.0
NIB074	0.0	33.3	33.3	5.6	27.8	72.2
NIB107	31.3	43.8	18.8	6.3	0.0	68.8
NIB146	56.3	12.5	0.0	31.3	0.0	43.8
NIB176	6.7	33.3	20.0	40.0	0.0	93.3

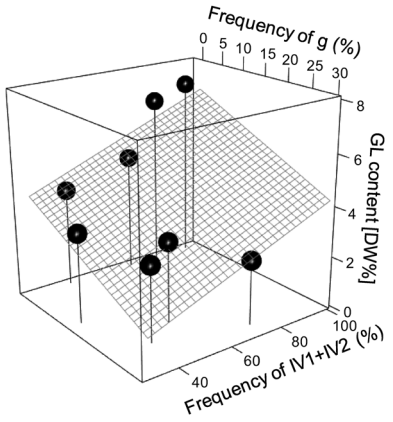

Figure 3. Correlation analysis between glycyrrhizin contents and appearance frequency of *CYP88D6* intron 7 alleles. g, *G. glabra* similar alleles of intron 7 of *CYP88D6* homologous genes; IV1, alleles similar to GuIV1 specific alleles (b and c in Figure 2); IV2, alleles similar with to GuIV2 specific allele (d in Figure 2). Grid lines indicated the regression plane [F=0.01, GL=2.813−0.090×(g)+0.041×(IV1+IV2)].

whether there was correlation between appearance frequency of the alleles and GL content. A multiple linear regression was calculated to predict GL contents based on the appearance frequency of intron 7 sequence of *CYP88D6* homolog [frequency of alleles (g) and total frequency of all alleles specially found in superior *G. uralensis* clones (IV1 and IV2)]. A significant regression equation was found (F=13.160, p=0.010, R^2=0.840) and predicted GL content is equal to 2.813−0.090×(g)+0.041×(IV1 and IV2). Both (g) and (IV1 and IV2) were significant predictors of GL content (g: p=0.024, IV1 and IV2: p=0.047) (Figure 3). These results indicate that these sequences could be used as a molecular marker for selection of *G. uralensis* superior clones that showed high GL content.

Selection of novel G. uralensis superior clones using CYP88D6 intron 7 sequence as a DNA marker

To evaluate usefulness of *CYP88D6* intron 7 sequences as a molecular marker for selection of superior *G. uralensis* clones, sterilized *G. uralensis* seeds were aseptically sowed and 17 clones were propagated by tissue culture (for example, Gu71#1, #22, #23, and #31, Figure 4A). These plantlets were hydroponically cultivated with pumice using diluted Otsuka nutrient solution in the growth chamber (Figure 4B).

Genomic DNA was extracted from leaves of hydroponically cultivated plants and the presence of alleles specially found in *G. uralensis* superior clone was confirmed by PCR. We could classify *G. uralensis* clones into 4 types: (1) specific primers amplified nothing (lane 11 and 16), (2) only allele c (lane 3, 6, 7, 8, 9, 12 and 15), (3) only allele d (lane 1, 4, 5, 13 and 14), and (4) both allele c and d (lane 2 and 10) (Figure 5A). The quality of genomic DNA was verified by PCR using consensus

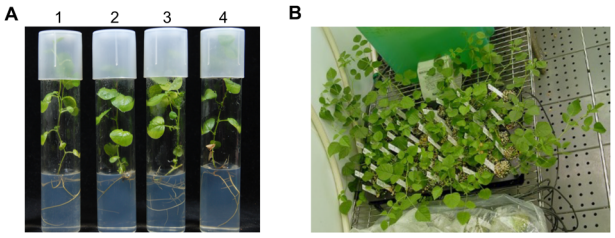

Figure 4. Hydroponic cultivation of *G. uralensis* clones propagated by tissue culture. (A) In vitro propagated *G. uralensis* clones. 1: Gu71#1, 2: Gu71#22, 3: Gu71#23, 4: Gu71#31. (B) *G. uralensis* clones hydroponically cultivated for 27–28 days.

primers (Figure 5B).

G. uralensis plants were hydroponically cultivated for 84–96 days and harvested roots were dried at 40°C. Dry root weight of each type showed no significant variations (type 1: 0.835±0.783 g, type 2: 0.679±0.471 g, type 3: 0.710±0.145 g, and type 4: 0.643±0.216 g). UPLC analysis of 50% ethanol extracts of hydroponically cultivated *G. uralensis* clones (Figure 6) revealed that secondary metabolite content was significantly increased in type 2–4 compared with type 1 [GL: type 1, 0.379±0.224 DW%; type 2, 1.007±0.259 DW%; type 3, 0.803±0.169 DW%; and type 4, 1.175±0.106 DW%; liquiritin (LQ): type 1, 0.099±0.056 DW%; type 2, 0.222±0.126 DW%; type 3, 0.219±0.051 DW%; and type 4, 0.510±0.028 DW%; isoliquiritin (IL): type 1, 0.033±0.015 DW%; type 2, 0.071±0.047 DW%; type 3, 0.056±0.015 DW%; and type 4, 0.110±0.007 DW%; and glycycoumarin (GC): type 1, 0.017±0.018 DW%; type 2, 0.037±0.020 DW%; type 3, 0.041±0.026 DW%; and type 4, 0.044±0.020 DW%]. In particular, LQ and IL contents of type 4 were significantly higher than those of other types (Figure 7).

Among aseptically germinated clones, we selected Gu71#1 (IV2 type), Gu71#12 (IV1 type), Gu71#22 (IV1), and Gu71#31 (IV1 and IV2 type) that were grown well in hydroponics and showed relatively high GL content (Yoshimatsu et al. 2017). Allele (g) was not detected from these selected clones by PCR direct sequence. The above selected clones propagated by cuttings or tissue culture also demonstrate high GL productivity in the field cultivation (detailed data will be presented elsewhere). Gu2-3-2 clone (Figure 1) is one of the superior clones we selected because of its high GL content (over 2.5%) in the root after 1-year hydroponic cultivation (Yoshimatsu et al. 2015), however, it lacks both of IV1 and IV2 type alleles and demonstrated much lower GL content (0.335±0.124%) in the roots compared with those in type 2, 3 and 4 clones after 73 days hydroponic cultivation similarly to the type 1 clones.

It is noteworthy that GL contents in the *G. uralensis* seedling clones seemed to be correlated with the contents of other secondary metabolites such as LQ and IL (Figure 7). Coordinated regulation of apparently separate secondary metabolic pathways has so far been

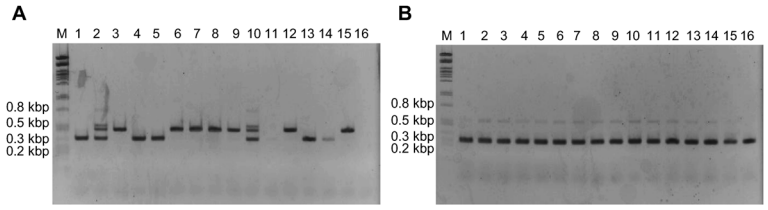

Figure 5. PCR analysis of *CYP88D6* intron 7 genotype of hydroponically cultivated *G. uralensis* clones. (A) PCR with the primers against IV1 and IV2 type specific sequences. (B) PCR with the primers against consensus sequence. Lane M: DNA size marker λ/*Pst*I, 1–15: hydroponically cultivated Gu clones, 16: hydroponically cultivated Gu2-3-2.

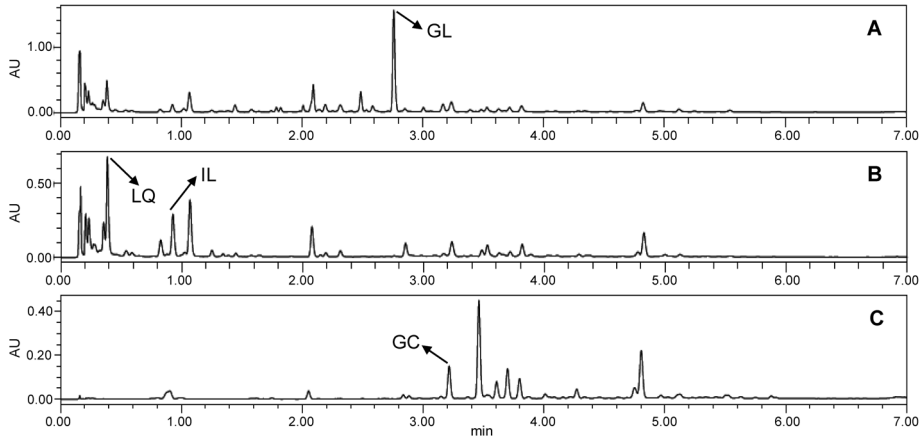

Figure 6. HPLC chromatogram of hydroponically cultivated *G. uralensis* clone (Gu71#31). (A) 254 nm chromatogram, (B) 316 nm chromatogram, (C) 350 nm chromatogram. GL: glycyrrhizic acid, LQ: liquiritin, IL: isoliquiritin, GC: glycycoumarin.

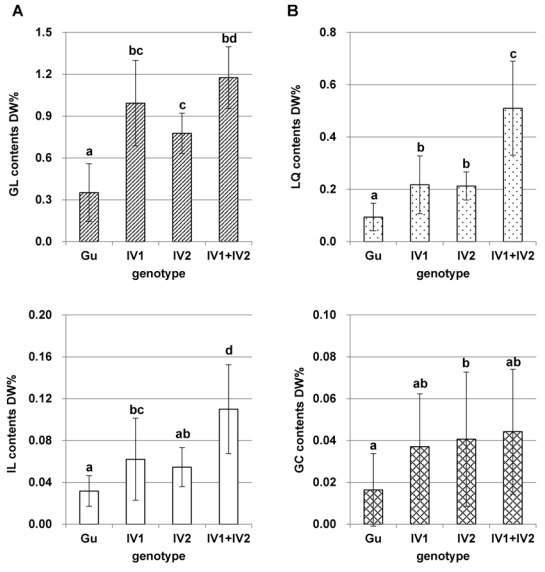

Figure 7. Secondary metabolite contents and *CYP88D6* intron 7 genotype of hydroponically cultivated *G. uralensis* clones. (A) glycyrrhizic acid (GL), (B) liquiritin (LQ), (C) isoliquiritin (IL), (D) glycycoumarin (GC). Different letters over the tops of columns indicate significant differences ($p<0.05$) by Tukey–Kramer's test.

reported (Dudareva et al. 2003; Kang et al. 2014; Zvi et al. 2012). Dudareva et al. (2003) revealed the coordinated regulation of phenylpropanoid and isoprenoid scent production through expression analyses of monoterpene synthase genes in snapdragon flowers. Zvi et al. (2012) overexpressed *Arabidopsis* production of anthocyanin pigment (PAP1) transcription factor in *Rosa* hybrida and demonstrated enhanced production of not only phenylpropanoid but also terpenoid volatiles in rose flowers. Kang et al. (2014) revealed that the defect of the flavonoid biosynthetic enzyme chalcone isomerase (CHI) can be a cause for lack of both flavonoids and terpenoids in the glandular trichomes of the anthocyanin free mutant of cultivated tomato (*Solanum lycopersicum*). These findings may imply the existence of coordinated regulation mechanism between the terpenoid and phenylpropanoid biosynthetic pathways in *G. uralensis*. Therefore, not only *G. uralensis* superior clones but also low GL producing ones may be useful to elucidate sophisticated regulation mechanism of secondary metabolites biosynthesis in *G. uralensis*.

In conclusion, we found that polymorphisms of intron 7 of *CYP88D6* homologous genes were useful for distinguishing among *Glycyrrhiza* species or *G. uralensis* superior clones. Although we need to accumulate much more genetic resources and combine to several regions in order to discriminate between superior clones and other clones with more precision, these results indicate that intronic region of species-specific biosynthetic genes is more effective than that of widely conserved genes, for example, ITS, rbcL gene and so on.

In addition, we found frequency of alleles [(g), (IV1) and (IV2)] in commercial crude drugs and selected

superior clones was well correlated to GL content. In practical selection, *G. uralensis* clones that had both IV1 and IV2 specific alleles showed tendency of high secondary metabolite content. Although the mechanism of correlation with secondary metabolite content and polymorphisms in intronic region of biosynthetic genes is unknown, by confirming the presence of these alleles, we will be able to narrow down superior clone candidates in early stage of cultivation without harvesting grown roots. Thus, intronic regions of biosynthetic genes are expected to be useful for not only discrimination among superior clones but also the first screening markers in the selection of high GL content clones. Recently, genome sequences of *Glycyrrhiza* plants are becoming elucidated (Mochida et al. 2017; Saito 2018). In the future, development of DNA makers is expected to further promote the breeding superior varieties of *Glycyrrhiza* plants.

Acknowledgements

This work is partially supported by Health Labour Sciences Research Grant, Research on New Drug Development from Japanese Ministry of Health, Labour and Welfare, and AMED under Grant Number JP17ak0101034. We thank Dr. Susumu Isoda for supplying the leaves of *G. inflata* and Japan Kampo Medicines Manufacturers Association for collecting crude drugs of the Japanese market.

References

Asl MN, Hosseinzadeh H (2008) Review of pharmacological effects of *Glycyrrhiza* sp. and its bioactive compounds. *Phytother Res* 22: 709–724

Dudareva N, Martin D, Kish CM, Kolosova N, Gorenstein N, Fäldt J, Miller B, Bohlmann J (2003) (E)-beta-ocimene and myrcene synthase genes of floral scent biosynthesis in snapdragon: Function and expression of three terpene synthase genes of a new terpene synthase subfamily. *Plant Cell* 15: 1227–1241

Hayashi H, Sudo H (2009) Economic importance of licorice. *Plant Biotechnol* 26: 101–104

Hosseinzadeh H, Asl MN (2015) Pharmacological effects of *Glycyrrhiza* spp. and its bioactive constituents: Update and review. *Phytother Res* 29: 1868–1886

Kang JH, McRoberts J, Shi F, Moreno JE, Jones AD, Howe GA (2014) The flavonoid biosynthetic enzyme chalcone isomerase modulates terpenoid production in glandular trichomes of tomato. *Plant Physiol* 164: 1161–1174

Kojoma M, Hayashi S, Shibata T, Yamamoto Y, Sekizaki H (2011) Variation of glycyrrhizin and liquiritin contents within a population of 5-year-old licorice (*Glycyrrhiza uralensis*) plants cultivated under the same conditions. *Biol Pharm Bull* 34: 1334–1337

Kondo K, Shiba M, Yamaji H, Morota T, Zhengmin C, Huixia P, Shoyama Y (2007) Species identification of licorice using nr DNA and cpDNA genetic markers. *Biol Pharm Bull* 30: 1497–1502

Lu HY, Liu JM, Zhang HC, Yin T, Gao SL (2008) Ri-mediated transformation of *Glycyrrhiza uralensis* with a squalene synthase gene (GuSQS1) for production of glycyrrhizin. *Plant Mol Biol Report* 26: 1–11

Mochida K, Sakurai T, Seki H, Yoshida T, Takahagi K, Sawai S, Uchiyama H, Muranaka T, Saito K (2017) Draft genome assembly and annotation of *Glycyrrhiza uralensis*, a medicinal legume. *Plant J* 89: 181–194

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio-assays with tobacco tissue cultures. *Physiol Plant* 15: 473–497

Ozaki K, Shibano M, Kusano G, Watanabe H (2010) Aim for production of Glycyrrhizae radix in Japan (2): Selection of pharmaceutically fine strains from *Glycyrrhiza uralensis* Fisher. *Jpn J Pharmacol (Shoyakugaku Zasshi)* 64: 76–82 (in Japanese)

R Core Team (2017) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. https://www.R-project.org/ (accessed on Dec. 2, 2020)

Research Center for Medicinal Plant Resources (2013) Comprehensive medicinal plant database, National institutes of biomedical innovation, health and nutrition, Japan. http://mpdb.nibiohn.go.jp/readme.html (accessed on Dec. 2, 2020)

Saito K (2018) Development of plant metabolomics and medicinal plant genomics. *Yakugaku Zasshi* 138: 1–18 (in Japanese)

Seki H, Ohyama K, Sawai S, Mizutani M, Ohnishi T, Sudo H, Akashi T, Aoki T, Saito K, Muranaka T (2008) Licorice beta-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, Sawai S, Ohyama K, Mizutani M, Ohnishi T, Sudo H, Fukushima EO, Akashi T, Aoki T, Saito K, et al. (2011) Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell* 23: 4112–4123

The Ministry of Health, Labour and Welfare (2016) Glycyrrhiza. In: *The Japanese Pharmacopoeia 17th Edition*: *Official from April 1, 2016, English Version*. Pharmaceutical and Medical Device Regulatory Science Society of Japan, Japan, pp 1862

Toda N, Sasaki S, Takeda O, Wei S, Wang W, Wang Y, Shi H, Li G (2012) Cultivation of licorice (*Glycyrrhiza uralensis*) in China (Part 1): Glycyrrhizin contents in root influenced by stem cutting. *Jpn J Pharmacol (Shoyakugaku Zasshi)* 66: 65–70 (in Japanese)

Yoshimatsu K, Kawano N, Inui T (2017) Genetic identification of high glycyrrhizic acid-producing *Glycyrrhiza* and discrimination method of the plant. Japan Patent 6233793B2, https://www.j-platpat.inpit.go.jp/p0200 (accessed on Dec. 2, 2020)

Yoshimatsu K, Kawano N, Inui T (2018) Cultivation method, hydroponic apparatus, and propagation method of *Glycyrrhiza* plants. Japan Patent 6391127B2, https://www.j-platpat.inpit.go.jp/p0200 (accessed on Dec. 2, 2020)

Yoshimatsu K, Kawano N, Inui T, Chida H (2015) A *Glycyrrhiza* plant strain and *Glycyrrhiza* plant multiplication. Japan Patent 5800227B2, https://www.j-platpat.inpit.go.jp/p0200 (accessed on Dec. 2, 2020)

Zvi MM, Shklarman E, Masci T, Kalev H, Debener T, Shafir S, Ovadis M, Vainstein A (2012) PAP1 transcription factor enhances production of phenylpropanoid and terpenoid scent compounds in rose flowers. *New Phytol* 195: 335–345