Journal of Experimental Botany, Vol. 56, No. 419, pp. 2573–2585, September 2005
doi:10.1093/jxb/eri251 Advance Access publication 8 August, 2005
This paper is available online free of all access charges (see http://jxb.oupjournals.org/open_access.html for further details)

Journal of
Experimental
Botany

RESEARCH PAPER

A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome

Norimoto Shimada[1,*], Ryohsuke Sasaki[1,*], Shusei Sato[2], Takakazu Kaneko[2], Satoshi Tabata[2], Toshio Aoki[1] and Shin-ichi Ayabe[1,†]

[1] *Department of Applied Biological Sciences, Nihon University, Fujisawa, Kanagawa 252-8510, Japan*
[2] *Kazusa DNA Research Institute, Kisarazu, Chiba 292-0812, Japan*

Received 21 February 2005; Accepted 21 June 2005

Abstract

Dihydroflavonol 4-reductase (DFR) is the first committed enzyme of the anthocyanin and condensed tannin pathways. Several *DFR* cDNAs have been cloned, and different specificities of DFR isozymes in the substrate hydroxylation patterns have been reported, but only fragmentary knowledge of *DFR* gene organization is available. Reported here is a comprehensive analysis of *DFR*s of a model legume, *Lotus japonicus*. A total of five *DFR* genes were found to form a cluster within a 38 kb region in the *L. japonicus* genome, whereas six cDNAs, including two splicing variants resulting from a transversion at a splicing acceptor site, were cloned. All the genes were expressed, with different organ specificities, in the mature plant. Three of the DFR proteins heterologously expressed in *Escherichia coli* showed catalytic activity, and their substrate preferences agreed with the variation of a specific active site residue (Asp or Asn) reported to control the specificity. The hydroxylation patterns of anthocyanidins and condensed tannin units in the stems did not reflect the substrate specificity of the expressed isozymes, implying complex regulation mechanisms in the biosynthesis. The two splicing variants and one DFR with Ser at the specificity-controlling position failed to show the activity, but a revertant protein replacing the unusual splicing restored the activity. The phylogenetic tree, constructed with known *DFR* sequences, showed evolutionary divergence of some of the *DFR* genes prior to the plant speciation. This work affords the basis for genetic and biochemical studies on the diversity of DFR and the flavonoid products.

Key words: Anthocyanin, condensed tannin, dihydroflavonol 4-reductase, flavanone 4-reductase, gene duplication, *Lotus japonicus*.

Introduction

Anthocyanins and condensed tannins (CTs) are two of the major flavonoid classes in higher plants. Anthocyanins are responsible for the brick red, red, and blue colours of flowers, and attract pollinators. CTs, also known as proanthocyanidins, are characterized by their ability to form stable tannin–protein complexes (Spencer *et al.*, 1988). These characteristics are due to the structures produced by the polymeric condensation of flavan-3-ols rich in free hydroxyl groups. CTs have anti-herbivore activity by causing reduced efficiency of nutrient absorption in the herbivore gut. In some plant species, accumulation of CTs has been observed after herbivore attack or wounding, suggesting that CTs play roles in defence responses (Peters and Constabel, 2002). CTs in leguminous plants have been reported to affect root nodulation (Pankhurst and Jones, 1979), and, in forage plants, CTs are also important as anti-bloat agents.

Dihydroflavonol 4-reductase (DFR; EC 1.1.1.219) is a pivotal enzyme of the flavonoid pathway leading to common anthocyanins and CTs. This enzyme yields flavan-3,4-diols by the reduction of corresponding dihydroflavonols, which are also intermediates of flavonol biosynthesis through the flavonol synthase reaction (Fig. 1). A number

* These authors contributed equally to this paper.
† To whom correspondence should be addressed. Fax: +81 466 84 3353; E-mail: ayabe@brs.nihon-u.ac.jp

© The Author [2005]. Published by Oxford University Press [on behalf of the Society for Experimental Biology]. All rights reserved.
The online version of this article has been published under an Open Access model. Users are entitled to use, reproduce, disseminate, or display the Open Access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and the Society for Experimental Biology are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact: journals.permissions@oupjournals.org

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

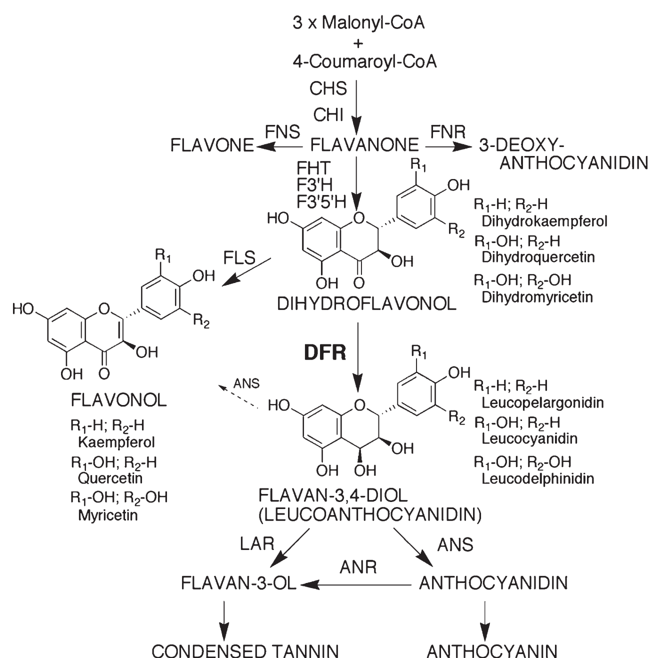

Fig. 1. The biosynthetic pathway of anthocyanins and condensed tannins in *Lotus japonicus*. DFR catalyses the reduction of dihydroflavonols to flavan-3,4-diols. ANR, anthocyanidin reductase; ANS, anthocyanidin synthase; CHI, chalcone isomerase; CHS, chalcone synthase; FHT, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F3′5′H, flavonoid 3′,5′-hydroxylase; FLS, flavonol synthase; FNR, flavanone 4-reductase; FNS, flavone synthase; LAR, leucoanthocyanidin reductase.

of *DFR* cDNAs have been isolated from various plants, and a single or multiple gene(s) encoding DFR protein(s) in a few plant genomes have been reported (Beld *et al.*, 1989; Helariutta *et al.*, 1993; Bongue-Bartelsman *et al.*, 1994; Charrier *et al.*, 1995; Tanaka *et al.*, 1995, 1996; Gong *et al.*, 1997; Bernhardt *et al.*, 1998; Chen *et al.*, 1998; Inagaki *et al.*, 1999; Himi and Noda, 2004). Flower colours in ornamental plants have been modified by controlling the expression levels of *DFR* genes (Aida *et al.*, 2000), and alteration of CT levels has been performed by introducing a *DFR* gene into a forage legume *Lotus corniculatus* (Bavage *et al.*, 1997; Robbins *et al.*, 1998).

DFR proteins of many plants can accept dihydroflavonols with different hydroxylation patterns, i.e. dihydrokaempferol (DHK), dihydroquercetin (DHQ), and dihydromyricetin (DHM), as substrates (Heller *et al.*, 1985; Fischer *et al.*, 1988; Stich *et al.*, 1992; Helariutta *et al.*, 1993; Tanaka *et al.*, 1995; Dellus *et al.*, 1997). On the other hand, DFRs of petunia and *Cymbidium*, which do not produce pelargonidin-derived anthocyanins, were postulated to have low activity to convert DHK to leucopelargonidin (Forkmann and Ruhnau, 1987; Johnson *et al.*, 1999). These data on DFR biochemistry have been acquired mainly from *in vitro* analyses using plant protein extracts or phytochemical analyses of mutant and/or transgenic plants. Heterologous expression of DFR proteins in *Escherichia coli* cells has been performed previously (De Jaeger *et al.*, 1997), but enzyme activities have not been reported. Only recently has DFR activity been successfully measured using an *E. coli* or yeast expression system (Martens *et al.*, 2002; Peters and Constabel, 2002; Fischer *et al.*, 2003; Shimada *et al.*, 2004; Xie *et al.*, 2004). However, sufficient information on the biochemical properties of DFR isozymes is not yet available, and characterization of all the isozymes in a single species has not been performed.

Lotus japonicus is a diploid perennial legume used as a model plant for the studies of classical and molecular genetics of the Fabaceae (Handberg and Stougaard, 1992; Schauser *et al.*, 1999; Nishimura *et al.*, 2002). Expressed sequence tag and genomic sequence databases are being developed (Sato *et al.*, 2001; Nakamura *et al.*, 2002; Asamizu *et al.*, 2003, 2004; Kaneko *et al.*, 2003; Kato *et al.*, 2003). While information on the flavonoid metabolism of *L. japonicus* is limited, cDNAs encoding key enzymes of the isoflavonoid pathway of this plant have been identified (Shimada *et al.*, 2000, 2003; Akashi *et al.*, 2003).

In this paper, the structure and function of all the *DFR* genes in *L. japonicus* are reported. Five *DFR* genes were found to form a cluster within a 38 kb region, and transcripts of all the *DFR*s, including two splicing variants, were detected. The expression patterns of isozymes were different among the organs. Heterologously expressed proteins showed different substrate preferences toward three dihydroflavonols. As far as is known, this is the first comprehensive report on gene structures, expression, and functional identification of all *DFR* paralogues in a single plant species.

Materials and methods

Flavonoid extraction and analysis

Lotus japonicus accession Gifu B-129 was grown in vermiculite with 1000-fold diluted Hyponex® solution (N:P:K=6:10:5; Hyponex-Japan, Osaka, Japan) under greenhouse conditions. Six-week-old plants were used for flavonoid extraction. Red-coloured stems (fresh weight 110 mg) were macerated overnight in 500 μl of 1% (v/v) HCl in 80% (v/v) ethanol at 4 °C. The supernatant and the precipitate fractions were separated. To analyse aglycones of anthocyanins and flavonols, the supernatant was acid-hydrolysed: an equal amount of 3 N HCl in H_2O was added (final concentration 1.5 N HCl in 40% ethanol) and then boiled for 1 h. To analyse CT units, 5% (v/v) HCl in *n*-butanol was added to the precipitate fraction, and hydrolysis was performed in boiling water for 1 h to convert CTs to anthocyanidins. All samples were filtered through a 0.45 μm Millex®-LH filter (Millipore, Bedford, MA, USA). The flavonoid identification was carried out by an HPLC-photodiode array spectrophotometry system (HPLC-PDA) with a CAPCELL PAK C18 MG column (4.6× 150 mm; Shiseido, Tokyo, Japan). Anthocyanidins and flavonols were detected at 525 nm and 375 nm, respectively, and identification was based on retention times and UV–Vis spectra compared with flavonoid standards. The solvent system of HPLC for anthocyanidin analysis was 11.5% (v/v) acetonitrile in phosphoric acid:acetic

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

acid:H_2O (3:8:89; by vol.), and for flavonol analysis 30% (v/v) acetonitrile in H_2O, at the flow rate 1.0 ml min^{-1} at 40 °C.

For 3-deoxyanthocyanidin analysis, 200 mg of red-coloured stems were soaked overnight in 500 μl methanol containing 0.1% (v/v) HCl at 4 °C. The solution was filtered and immediately analysed by HPLC-PDA. HPLC was performed with the solvent system of 11.5% (v/v) acetonitrile in phosphoric acid:acetic acid:H_2O (3:8:89; by vol.) at a flow rate of 1.0 ml min^{-1} at 40 °C, and the eluate was monitored at 471 nm (for apigeninidin) or 485 nm (for luteolinidin).

The flavonoid standards, apigeninidin, cyanidin, delphinidin, kaempferol, luteolinidin, myricetin, pelargonidin, and quercetin, were purchased from Extrasynthèse (Genay, France). For HPLC analysis, these standards were dissolved in methanol (100 mg ml^{-1}).

cDNA cloning

To isolate *DFR* cDNA fragments from *L. japonicus*, degenerate oligonucleotide primers were designed based on the highly conserved amino acid sequences of DFR proteins from several plants. The template cDNA was synthesized from mRNA prepared from a whole plant of *L. japonicus* accession Gifu B-129 accumulating visible anthocyanins, as described previously (Shimada *et al.*, 2000). PCR was performed with the template using the 5′-primer DFR/S1 (5′-AARGCYCCXGARAAYGA-3′) and the 3′-primer DFR/AS1 (5′-RCRAARTACATCCAXSC-3′); the programme was 30 cycles of 94 °C (15 s)/43 °C (10 s)/72 °C (30 s) followed by a final extension for 5 min at 72 °C. cDNA fragments of the predicted size (~220 bp) were cloned into a pT7Blue vector (Novagen, Madison, WI, USA) and sequenced. Rapid amplification of cDNA ends (RACE)-PCR was performed according to the Marathon cDNA amplification kit user manual (Clontech, Palo Alto, CA, USA). Two sets of specific primers, the 5′-primer DFR2as2 (5′-GTTGGTTTTCACATCATCA-GCC-3′) and the 3′-primer DFR2s3 (5′-AGTGTTTCAGAACGCC-ATAAGCAT-3′), and the 5′-primer DFR3as2 (5′-GCTGGTTTTC-ACATCCTCTGCA-3′) and the 3′-primer DFR3s3 (5′-ACCCTCA-ACGTCATAGAGCACCAA-3′), were used for 3′- and 5′-RACE, and then primers for full-length cDNA amplification were re-designed. These primers contained an *Nhe*I or *Xho*I site (underlined) to subclone full-length *DFR* cDNAs into these sites of the expression vector pET28a (Novagen). Primer combinations to clone each full-length cDNA were as follows: GdNhe2 (5′-AGAAGAGAACG-GCTAGCATGGGTTCGG-3′) and Gd2Xho (5′-ACAGGCACCTC-GAGAGACATAAACAAAGCT-3′) for *cDFR1*, LdNhe (5′-AGAG-AGTAGCTAGCATGGGATCAGTAC-3′) and Ld3Xho (5′-TCAC-AGACACCTCGAGATAAACAATGC-3′) for *cDFR2*, GdNhe2 (5′-AGAAGAGAACGGCTAGCATGGGTTCGG-3′) and dfr3Xho (5′-CCACTCGAGAGACACAAAGATATAAAT-3′) for *cDFR3*, GdNhe2 (5′-AGAAGAGAACGGCTAGCATGGGTTCGG-3′) and Gd3Xho (5′-TAAACAAGGCTCGAGTTATTTCTCTG-3′) for *cDFR4*, and dfr5fs1 (5′-AGATCGCTAGCATGAGTTCCGAGTCTGA-3′) and dfr5fas1 (5′-TACCGCTCGAGAAGTAGTAGCAAGTAAC-3′) for *cDFR5*. Primer sets for *cDFR5* were designed from sequence information of the *DFR5* gene (see Results). Three individual cDNA clones for each gene were sequenced.

Genome analysis

Generation of a genome library using the TAC vector (Liu *et al.*, 1999) from the genomic DNA of *L. japonicus* accession Miyakojima MG-20, sequencing strategy, and gene assignment were carried out as described elsewhere (Sato *et al.*, 2001). The library was screened for *DFR* genes by the PCR method using a primer set based on the sequence of *cDFR1* (5′-AGGCAGTCATCTTTACTTTCTTGC-3′) and (5′-AAGTGATCAAGCCTGCCATAAATG-3′). Genetic mapping was carried out with a simple sequence repeat marker found in LjT36H12.

Southern blot analysis

Genome DNA of *L. japonicus* accession Gifu B-129 was isolated using a DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). Genome DNA (10 μg) was digested with *Dra*I, *Eco*RI, *Hin*dIII, or *Xba*I. Electrophoresis, blotting, preparation of digoxigenin-labelled probes, and hybridization were performed according to the DIG application manual (Roche Diagnostics, Mannheim, Germany). The washing was done at 55 °C for 15 min in 1×SSC and 0.1% (w/v) SDS solution, and was repeated twice.

Heterologous expression of recombinant DFR proteins in Escherichia coli

To express *L. japonicus* DFR as N-terminal His-tagged proteins in *E. coli*, each cDNA was subcloned into *Nhe*I–*Xho*I sites of a pET28a vector (Novagen). An empty vector (pET28a only) and expression vectors containing different *DFR* cDNAs were used for transformation of *E. coli* strain BL21(DE3) (Stratagene, La Jolla, CA, USA). DFR proteins were induced by adding 1 mM isopropyl-thio-β-D-galactoside to the culture for 2 h at 28 °C. *Escherichia coli* cells were then harvested by centrifugation at 9400 *g* for 3 min and washed with 0.1 M TRIS–HCl (pH 7.5) containing 10 mM EDTA and 50 mM sucrose. *Escherichia coli* cells were collected as above, frozen with liquid nitrogen, and stored at −80 °C.

The cell pellet (0.3 g) was suspended in 7 ml of 0.1 M potassium phosphate buffer (pH 7.0) containing 10% (w/v) sucrose, disrupted by sonication and then centrifuged at 9400 *g* for 20 min at 4 °C. The supernatant was used to assay DFR.

Enzyme assay

To measure the DFR and flavanone 4-reductase (FNR) activity of His-tagged DFR proteins, the reaction mixtures were acid-treated to convert the direct reaction products, flavan-3,4-diol and flavan-4-ol, to their respective anthocyanidins and analysed by HPLC. For the initial DFR enzyme assay, the reaction mixture (total 500 μl) contained 0.1 M potassium-phosphate buffer (pH 7.0), 10 mM NADPH, and 10 μg of (±)-taxifolin (DHQ; Sigma, St Louis, MO, USA). For FNR assay, the flavanones (10 μg), naringenin and eriodictyol (Extrasynthèse), were used. The reaction was started by the addition of enzymes, and, after 3 h at 25 °C, terminated by extraction with 400 μl of ethyl acetate (twice). The ethyl acetate extracts were evaporated *in vacuo* at room temperature. The dried ethyl acetate extracts were added into 50 μl of 5% (v/v) hydrochloric acid in *n*-butanol, and boiled for 5 min. HPLC was performed using a CAPCELL PAK C18 MG column (4.6×150 mm; Shiseido) with 15% (v/v) (for pelargonidin), 11.5% (v/v) (for apigeninidin and luteolinidin), or 7% (v/v) (for cyanidin and delphinidin) acetonitrile in phosphoric acid:acetic acid:H_2O (3:8:89; by vol.) at a flow rate of 1.0 ml min^{-1} at 40 °C, and the eluate was monitored at 471 nm (for apigeninidin), 485 nm (for luteolinidin), or 525 nm (for cyanidin, delphinidin, and pelargonidin). To examine the relative substrate preferences of DFR2, DFR3, and DFR5, excess amounts (10 μg) of (+)-taxifolin (Extrasynthèse), DHK, and DHM (products from plant materials; Plantech, Reading, UK) were used as the substrates. The assay was performed by the same methods as described above. The enzyme activity was calculated based on the absorbance at 525 nm of converted anthocyanidins measured by NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). The activity was normalized against the value obtained for (+)-DHQ with the same enzyme.

SDS-PAGE analysis

Total protein of *E. coli* cells (50 μg fresh weight) was extracted after the induction of DFR proteins, and SDS-PAGE was carried out using

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

a 12.5% (w/v) acrylamide gel. Gel was stained with 0.1% (w/v) Coomassie blue R-250 in methanol:acetic acid:H_2O (5:1:4; by vol.).

Site-directed mutagenesis of cDFR4a

Mutagenesis of *cDFR4a* was carried out by the unique restriction enzyme site elimination technique (Deng and Nickoloff, 1992). For primers, D4a+7b (5′-TCTGGAATAT*TCTTGAA*CTTAGTGGGGA-C-3′) was used to complement the 7 bp (italicized) deletion resulting in the unusual splicing event (see Results), and 28aER5ko (5′-CGT-TGGTGCGGACATTTCGGTAGTGG-3′) was used to eliminate the *Eco*RV site on the pET28a vector. After sequence verification, the plasmid was used to transform *E. coli* strain BL21(DE3), and expression of the His-tagged DFR4a revertant protein was checked by SDS-PAGE. The enzyme activity towards DHQ was measured as described above (see Enzyme assay above).

Semi-quantitative RT-PCR analysis

Five hundred milligrams each of young seeds and pods, flowers, stems, leaves, roots, and nodules were collected from 8-week-old *L. japonicus* accession Gifu B-129 grown under the same conditions described above. mRNAs were isolated from these samples using Straight A's mRNA isolation system (Novagen) according to the manufacturer's instructions. First-strand cDNAs were synthesized with 200 ng mRNA using a Ready-To-Go T-Primed First-Strand kit (Amersham Biosciences, Piscataway, NJ, USA). The quantity of each template was adjusted to give roughly equal amplification of β-*tubulin* cDNA. RT-PCR was carried out with 0.5 pmol each of specific primers using *ExTaq* DNA polymerase (Takara, Tokyo, Japan) in a final volume of 20 μl according to the manufacturer's protocol. The reaction was performed for 30 cycles of 94 °C (15 s)/60 °C (10 s)/72 °C (30 s) followed by a final extension for 5 min at 72 °C. The products (5 μl) were subjected to electrophoresis on 1.2% (w/v) agarose gel and stained with ethidium bromide. The gene-specific primer pairs used were: 5′-AATTCTGGGAAGTCATCTGCGAC-GAG-3′ and 5′-CTGGTGCACTGAAAGCGTAGCATTAT-3′ for β-*tubulin*, 5′-AGACTTCGTCTCAATCATCCCATCC-3′ and 5′-ACCTTATCCAATCAAGAACAGAACA-3′ for *DFR1*, 5′-AATA-TCCCCACTAAGTTCAAGAATA-3′ and 5′-AACAGAAATCCT-AATCAAGAAACAC-3′ for *DFR2*, 5′-AGAAGAGAACGGCTA-GCATGGGTTCGG-3′ and 5′-CCACTCGAGAGACACAAAGAT-ATAAAT-3′ for *DFR3*, 5′-AGAAGAGAACGGCTAGCATGGG-TTCGG-3′ and 5′-ATTAAACAAGGCCTCGAGATTTCTCTG-3′ for *DFR4*, and 5′-CAAAGAGCACAACATAGACTTCGTC-3′ and 5′-CCATCACTGGCACAAAGATACAGAA-3′ for *DFR5*.

Phylogenetic analysis

A phylogenetic tree of *DFR* genes was constructed based on nucleotide sequences of exon 3 and parts of exons 2 and 4. Exon 3 encodes a region that was postulated to control DFR substrate specificity. A Neighbor–Joining tree was produced from the results of 1000 bootstrap replicates using the CLUSTAL W program (Thompson *et al.*, 1994) of the DNA Data Bank of Japan (Shizuoka, Japan). The phylogenetic tree was displayed by NJplot software (Perriere and Gouy, 1996).

GenBank accession numbers for the nucleotide sequences of *DFR*s are: *Antirrhinum majus*, X15536; *Arabidopsis thaliana*, AB033294; *Bromheadia finlaysoniana*, AF007096; *Callistephus chinensis*, Z67981; *Camellia sinensis*, AB018686; *Cymbidium hybrida*, AF017451; *Daucus carota*, AF184271; *Dianthus caryophyllus*, Z67983; *Forsythia×intermedia*, Y09127; *Fragaria×ananassa*, AF029685; *Gentiana triflora*, D85185; *Gerbera hybrida*, Z17221; *Glycine max DFR1*, AF167556; *Hordeum vulgare*, S69616; *Ipomoea nil DFRA*, *DFRB,* and *DFRC*, AB006793; *Ipomoea purpurea DFRA*, *DFRB,* and *DFRC*, AB011667; *Lilium hybrid*, AF169801; *Lotus corniculatus DFR1,* AF117261; *L. corniculatus DFR2*, AF117262; *L. corniculatus DFR3*, AF117263; *L. corniculatus DFR4*, AF117264; *L. japonicus DFR1*, AB162109; *L. japonicus DFR2*, AB162110; *L. japonicus DFR3*, AB162111; *L. japonicus DFR4a*, AB162112; *L. japonicus DFR5*, AB162114; *Lycopersicon esculentum*, Z18277; *Malus domestica*, AF117268; *Medicago truncatula DFR1*, AY389346; *M. truncatula DFR2*, AY389347; *Oryza sativa*, AB003496; *Perilla frutescens*, AB002817; *Petunia×hybrida*, AF233639; *Rosa hybrida*, D85102; *Solanum tuberosum*, AF449422; *Torenia hybrida*, AB012924; *Triticum monococcum*, AF434703; *Vaccinium macrocarpon DFR1*, AF483835; *Vaccinium macrocarpon DFR2*, AF483836; and *Vitis vinifera*, Y11749.

Results

Flavonoid aglycones and CT units in stems of L. japonicus *accession Gifu B-129*

First, flavonoid aglycones and CT units in red-coloured stems of 6-week-old *L. japonicus* were analysed to complement the limited information so far available (Skadhauge *et al.*, 1997). Two flavonol aglycones, kaempferol (493.6 μg g^{-1} FW) and quercetin (101.8 μg g^{-1} FW), originating from DHK and DHQ, respectively, were detected (Fig. 2B). Myricetin, a flavonol of the DHM derivative, was not detected in this study. Figure 2D shows that the major anthocyanidin is a DHQ derivative, cyanidin (211.18 μg g^{-1} FW), and that another DHQ derivative, peonidin (3′-*O*-methylated cyanidin, 35.5 μg g^{-1} FW), was also detected. By contrast with the presence of a DHK-derived flavonol (kaempferol) and the absence of a DHM-derived flavonol, no accumulation of a DHK-derived anthocyanidin such as pelargonidin and accumulation of a DHM-derived anthocyanidin, delphinidin (33.6 μg g^{-1} FW), were observed (Fig. 2D). CTs are polymers of flavan-3-ol units, which are converted to anthocyanidins by acid-hydrolysis. Cyanidin (75.0 μg g^{-1} FW) and delphinidin (53.2 μg g^{-1} FW) were constituents of the acid-treated *L. japonicus* CT fraction, but peonidin and pelargonidin were not detected (Fig. 2E). The ratio of delphinidin to cyanidin derived from CTs was much higher than that observed in anthocyanin aglycones. Thus, the CTs of *L. japonicus* are polymers of DHQ and DHM derivatives, and DHK derivatives are not present. 3-Deoxyanthocyanidins were also not observed in this study (data not shown).

Cloning of DFR *cDNAs and the structure of* DFR *genes*

Degenerate PCR using the cDNAs synthesized from mRNAs of anthocyanin-accumulating seedlings of *L. japonicus* accession Gifu B-129 as the template gave four cDNA fragments, designated *cDFR1*–*cDFR4*. After RACE-PCR, specific primers for the full-length ORF (open reading frame) of each cDNA were designed. The PCR with the four sets of specific primers yielded five distinct cDNAs (GenBank accession numbers: AB162109 for *cDFR1*, AB162110 for *cDFR2*, AB162111 for *cDFR3*, AB162112 for *cDFR4a*, and AB162113 for *cDFR4b*). Two cDNAs with different lengths,

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

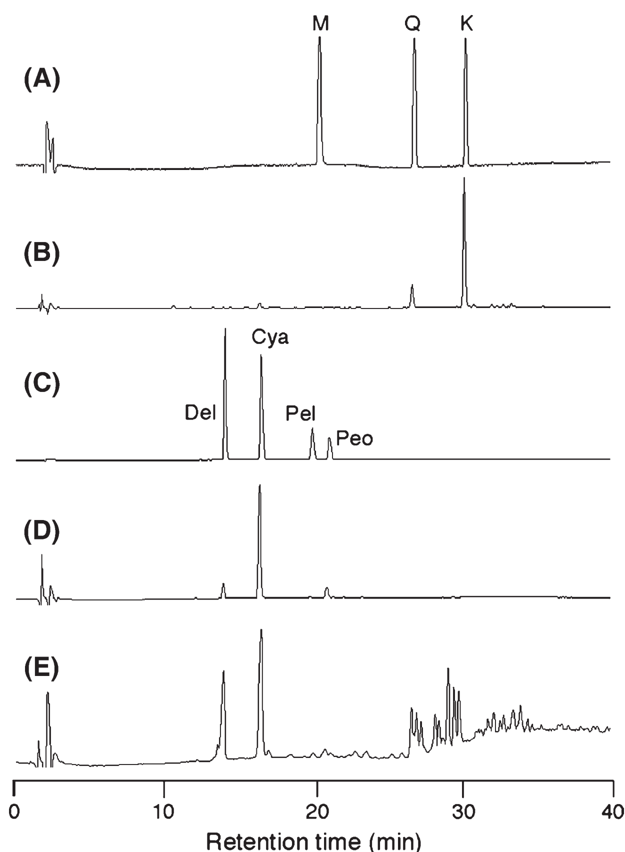

Fig. 2. HPLC chromatograms of flavonoid aglycones and CT units in coloured *L. japonicus* stems. Aglycones of anthocyanins and flavonols, and the CT units after conversion to anthocyanidins in the stems, were identified. (A) Flavonol standard samples: kaempferol (K), myricetin (M), and quercetin (Q). (B) Flavonol aglycones: kaempferol and quercetin were found. (C) Anthocyanidin standard samples: cyanidin (Cya), delphinidin (Del), pelargonidin (Pel), and peonidin (Peo). (D) Anthocyanidins: DHQ-derived anthocyanidins: cyanidin and peonidin, and DHM-derived delphinidin were found, but DHK-derived pelargonidin was not detected. (E) Anthocyanidins derived from CTs: cyanidin and delphinidin were found. Thus, CTs in the coloured stem were polymers of DHM and DHQ derivatives.

cDFR4a and *cDFR4b*, were presumed to be produced due to a mutation at a splicing accepter site of the *DFR4* gene (see below). In addition, another primer set for *DFR5* was designed based on the genome sequence information, and the full-length *DFR5* cDNA (GenBank accession number AB162114) was obtained by PCR.

The identities of nucleotide and deduced amino acid sequences (Fig. 3) among *DFR* paralogues of *L. japonicus* are 96–79% and 94–75%, respectively. Sequence alignment with a number of NADPH- and NADH-dependent dehydrogenases and reductases showed that the N-terminus regions of *L. japonicus* DFRs contain a putative NAD(P)-binding region (Lacombe *et al.*, 1997).

The genome library of *L. japonicus* accession Miyakojima MG-20 has been constructed (Sato *et al.*, 2001). A primer set for screening of *DFR* genes amplified a TAC clone, LjT36H12 (GenBank accession numbers AP006721 and AP006722), which was mapped to the long arm terminal of chromosome V. The sequence of the clone showed that five *DFR*-like genes form a cluster within a 38 kb region (Fig. 4A). Four of these genes correspond to *cDFR1*–*cDFR4a* and *cDFR4b*, and another gene was named *DFR5*. All five genes had the same orientation, and all contained six exons. Although the length of each gene is different from the others, the sizes of the respective exons among these five genes are almost the same (Fig. 4A). Detailed inspection of the LjT36H12 for the *DFR* sequences revealed a transversion, guanine to thymine, at the splicing accepter site of intron 5 of the *DFR4* gene (Fig. 4B). It was considered that *DFR4a* and *DFR4b* mRNAs were produced by an unusual splicing event, which results in a shift of the 3′-splice site to 7 bp and 17 bp downstream, respectively, from the original site.

To investigate the number of *DFR* genes, a Southern blot analysis was performed (Fig. 4C). When the cDNA fragment covering exon 2 of *DFR2* was used as a probe, three bands in the *Dra*I-, *Eco*RI-, *Hin*dIII-, or *Xba*I-digested *L. japonicus* genome were detected. The numbers and sizes of the hybridization signals matched those of the predicted fragments from the restriction map of LjT36H12 well, and no other TAC clone was obtained by PCR-based screening. Thus, in total, five *DFR* genes exist in the *L. japonicus* genome, and these are located in a single locus.

Organ specificity of mRNA accumulation

To examine the expression patterns of *DFR* genes in various organs, semi-quantitative PCR was performed (Fig. 5). The specificity of the primers used in this study was confirmed by PCR using each *cDFR* as a template. mRNAs were isolated from each organ of 8-week-old *L. japonicus* accession Gifu B-129 accumulating visible anthocyanins in stems. No contamination with genomic DNA in the cDNA samples was confirmed by PCR with a primer set designed to overlap one of the exon–intron junctions of *DFR1* (data not shown). Figure 5 shows the different organ specificity of the expression. The transcript of *DFR1* was detected in all the organs tested, and that of *DFR2* in both the aerial and underground parts as well, but not the leaves. *DFR3* was characterized by the specific expression in only stems and leaves. *DFR4* and *DFR5* were expressed mainly in the aerial parts together with weak expression in roots.

Heterologous expression and catalytic activities of DFR proteins

A region determining the substrate preference of DFR was proposed based on the amino acid sequence alignment (Beld *et al.*, 1989), and a mutation analysis showed that an amino acid residue at the position corresponding to the 134th residue of *Gerbera* DFR affected the substrate preference (Johnson *et al.*, 2001). DFRs of many plants

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

```
Gerbera      1 ME--EDSPATVCVTGAAGFIGSWLVMRLLERGYVVHATV-RDPGDLKKVKHLLELPKAQTNLKLWKADLTQEGSFDEAIQGCHGVFHLAT    87
DFR1         1 MG---SAAKTVCVTGSTGFIGSWLVMRLMERGYMVRATVQRDPDNMKKVKHLLELPGAKTNLTIWNADLTEEGSFDEAIKGCSGVFHVAS    87
DFR2         1 MG---SVPETVCVTGAAGFIGSWLVMRLMERGYMVRATV-RDPANMKKVKHLLELPEAKTKLTLWKADLAEEGSFDEAIKGCTGVFHVAT    86
DFR3         1 MG---SAAKTVCVTGSTGFIGSWLVMRLMEGGYTVRATVQRDPDNMKKVKHLLELPGAKTNLTIWNADLTEEGSFDEAINGCSGVFHVAS    87
DFR4a        1 MG---SAAKTVCVTGSTGFIGSWLVMRLMERGYMVRATVQRDPENMKKVKHLLELPGAKTNLTIWNADLTEEGSFDEAINGCSGVFHLAT    87
DFR4b        1 MG---SAAKTVCVTGSTGFIGSWLVMRLMERGYMVRATVQRDPENMKKVKHLLELPGAKTNLTIWNADLTEEGSFDEAINGCSGVFHLAT    87
DFR5         1 MS---SESETVCVTGAAGFIGSWLVMRLIERGYTVRATI-RDPANMKKVKHLLELPDAKTKLSLWKADLAEEGSFDEAIRGCTGVFHVAT    86
Petunia      1 MPLHLRCSATVCVTGAAGFIGSWLVMRLLERGYNVHATV-RDPENKKKVKHLLELPKADTNLTLWKADLTVEGSFDEAIQGCQGVFHVAT    89
                        ******  ***********.*.** *.**. ***  ..********** *.*.*. * ***..******** ** ****.*.

Gerbera     88 PMDFESKDPENEIIKPTIEGVLSIIRSCVKAKTVKKLVFTSSAGTVNGQEKQLHVYDESHWSDLDFIYSKKMTAWMYFVSKTLAEKAAWD   177
DFR1        88 PMDFNSKDPENEVIKPAINGVLDIMKACLKAKTVRRLVFTSSAGILSVSERHKHMLDETCWGDLEFCKKVKMTGWMYFVSKELAEQEALK   177
DFR2        87 PMDFESKDPENEVIKPTINGVLDIMKACQKAKTVRRLVFTSSAGTLNVIEHQKQMFDESCWSDVEFCRRVKMTGWMYFVSKTLAEQEAWK   176
DFR3        88 PMDFNSKDPENEVIKPSINGVLDIMKACQKAKTVRRLVFTSSAGTLNAVEHQKQMCDESCWSDVEFCRRVKMTGWMYFVSKTLAEQEAWK   177
DFR4a       88 PMDFNSKDPENEVIKPTINGVLDIMKACQKAKTVRRLVFTSSAGTLDAVEHQKQMFDESCWSDVEFCRRVKMTGWMYFVSKTLAEQEAWK   177
DFR4b       88 PMDFNSKDPENEVIKPTINGVLDIMKACQKAKTVRRLVFTSSAGTLDAVEHQKQMFDESCWSDVEFCRRVKMTGWMYFVSKTLAEQEAWK   177
DFR5        87 PMDFESKDPENEVIKPTINGLLDILKACEKAKTVRRLVFTSSAGTVDVTEHPKPVIDETCWSDIEFCLRVKMTGWMYFVSKTRAEQEAWK   176
Petunia     90 PMDFESKDPENEVIKPTVRGMLSIIESCAKANTVKRLVFTSSAGTLDVQEQQKLFYDQTSWSDLDFIYAKKMTGWMYFASKILAEKAAME   179
               **** *******.***...*.*.*...* **.**..********..   *... . *...*.* .*. ..***.****.**..**..*..

Gerbera    178 ATKGNNISFISIIPTLVVGPFITSTFPPSLVTALSLITGNEAHYSIIKQGQYVHLDDLCECHIYLYENPKAKGRYICSSHDATIHQLAKI   267
DFR1       178 FAKENNIDFVSIIPSLVVGPFLMPTMPPSLYTALCPITGNEAHYMIMKQSQFVHVDDLCLAHIFLFEHPESEGRYMCSACDANIHDIAKL   267
DFR2       177 FAKEHGIDFITIIPPLVVGSFLMPTMPPSLITALSPITGNEAHYSIIKQGQYVHLDDLCLAHIFLFEHPESEGRYICSASEATIHDIAKL   266
DFR3       178 FAQEHDIDFITIIPSLVVGSFLMPTLPPSLTTALSPITGNEAHYSIIKQGQYVHLDDLCLAHIFLFEHPKSEGRYICSASEATIHDIAKL   267
DFR4a      178 FAKEHDIDFITTIPSLVVGSFLMPTMPPSLITALSPITGNEAHYSIIKQGQFVHLDDLCLAHIFLFEHPESEGRYILSASEATIHDIAKL   267
DFR4b      178 FAKEHDIDFITTIPSLVVGSFLMPTMPPSLITALSPITGNEAHYSIIKQGQFVHLDDLCLAHIFLFEHPESEGRYILSASEATIHDIAKL   267
DFR5       177 YAKEHNIDFVSVIPPLVVGPFLMPTMPPSLITALSLITGNEAHYSIIKQGQYVHLDDLCLAHIFLFENPKAQGRYMCSAYEATIHEVARM   266
Petunia    180 EAKKKNIDFISIIPPLVVGPFITPTFPPSLITALSLITGNEAHYCIIKQGQYVHLDDLCEAHIFLYEHPKADGRFICSSHHAIIYDVAKM   269
               ..... *.*. .** **** *...*.****.***..********.*.**.*.**.****..**.*.*.* ..**...*. .*.*...*..

Gerbera    268 IKDKWPEYYIPTKFPGIDEELPIVSFSSKKLIDTGFEFKYNLEDMFKGAIDTCREKGLLPYSTIKNHINGNHVNGVHHYIKNNDDDHEKG   357
DFR1       268 INTKYPEYNVPTKFKNIPDELELVRFSSKKIKDMGFQFKYTLEDMYTGAIDACREKGLLPKAAETPS-NGIMEK---------------   340
DFR2       267 INSKYPEYNIPTKFKNIPDELELVRFSSKKIKDMGFEFKYSLEDMYTGAIDTCKEKGLLPKAAENPS-NGK------------------   336
DFR3       268 INSKYPEYNVPTKFKNIPDELELVRFSSKKIKDMGFEFKYSLEDMYTGAIDTCKEKGLLPKAAENPS-NGITEK---------------   340
DFR4a      268 INSKYPEYNVPTKIFQMNWSLSDFHQRRSKKWGSNLNTA--------------------------------------------------   306
DFR4b      268 INSKYPEYNVPTK----------------------------------------------------------------------------   280
DFR5       267 INKKYPEFNVPTKFKDIPDELDIIKFSSKKITDLGFKFKYSLEDMYTGAVETCREKGLLPKTAETPATNGTTQK---------------   340
Petunia    270 VREKWPEYYVPTEFKGIDKDLPVVSFSSKKLTDMGFQFKYTLEDMYKGAIDTCRQKQLLPFSPRSAEDNGHNREAIAISAQNYASGKENA   359
               .. *.**...**... .  ..  . .....  . .. ... ..... ...... ......        ..

Gerbera    358 LLCCSKEGQ-----                                                                              366
DFR1       340 --------------                                                                              340
DFR2       336 --------------                                                                              336
DFR3       340 --------------                                                                              340
DFR4a      306 --------------                                                                              306
DFR4b      280 --------------                                                                              280
DFR5       340 --------------                                                                              340
Petunia    360 PVANHTEMLSNVEV                                                                              373
```

Fig. 3. Alignment of deduced amino acid sequences of DFRs. Asterisks indicate conserved residues, and a putative NAD(P)-binding domain is underlined. The boxed region has been postulated to control the substrate specificity of DFR (Beld *et al.*, 1989), and the amino acid residue (indicated by an arrowhead) is especially important for the specificity (Johnson *et al.*, 2001).

acting on three dihydroflavonols have an Asn at the position corresponding to *Gerbera* N134 (referred to as Asn-type DFR), whereas DFRs of petunia and *Cymbidium*, which do not efficiently convert DHK to leucopelargonidin, have an Asp at that position (referred to as Asp-type DFR). An alignment of deduced amino acid sequences of *L. japonicus* DFRs showed that DFR2 and DFR3 are the Asn type and DFR4a, DFR4b, and DFR5 are the Asp type. DFR1 contains Ser at the corresponding position.

For the expression of DFR proteins in *E. coli*, six *DFR* cDNAs were subcloned into *Nhe*I–*Xho*I sites of a pET28a vector. His-tagged DFR proteins were successfully expressed in *E. coli*, and SDS-PAGE of the expressed DFR proteins showed bands around the predicted molecular masses, 40.3 kDa (DFR1+6×His), 39.9 kDa (DFR2+6×His), 40.0 kDa (DFR3+6×His), 36.6 kDa (DFR4a+6×His), 33.5 kDa (DFR4b+6×His), and 40.3 kDa (DFR5+6×His), respectively (Fig. 6A).

Crude extracts of *E. coli*-expressing DFR proteins were used for enzyme assays. DFR activity was first measured using (±)-DHQ as a substrate in the presence of NADPH. The reaction products were extracted with ethyl acetate and then acid-treated to convert the direct DFR reaction products, flavan-3,4-diols, to anthocyanidins. As shown in Fig. 6B, formation of cyanidin was observed in assays with crude DFR2, DFR3, or DFR5 protein. Pelargonidin and delphinidin were produced when DHK and DHM were used as substrates in the assays with these proteins (see below). On the other hand, no detectable DFR activity was observed in the assays with DFR1, DFR4a, or DFR4b protein towards any of the dihydroflavonols above.

Since the expression levels of DFR2, DFR3, and DFR5 proteins in *E. coli* were very different (Fig. 6A), only the relative substrate preferences of these DFR isozymes towards DHK, DHM, and DHQ were compared. Production of anthocyanidins was measured by spectrophotometer

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

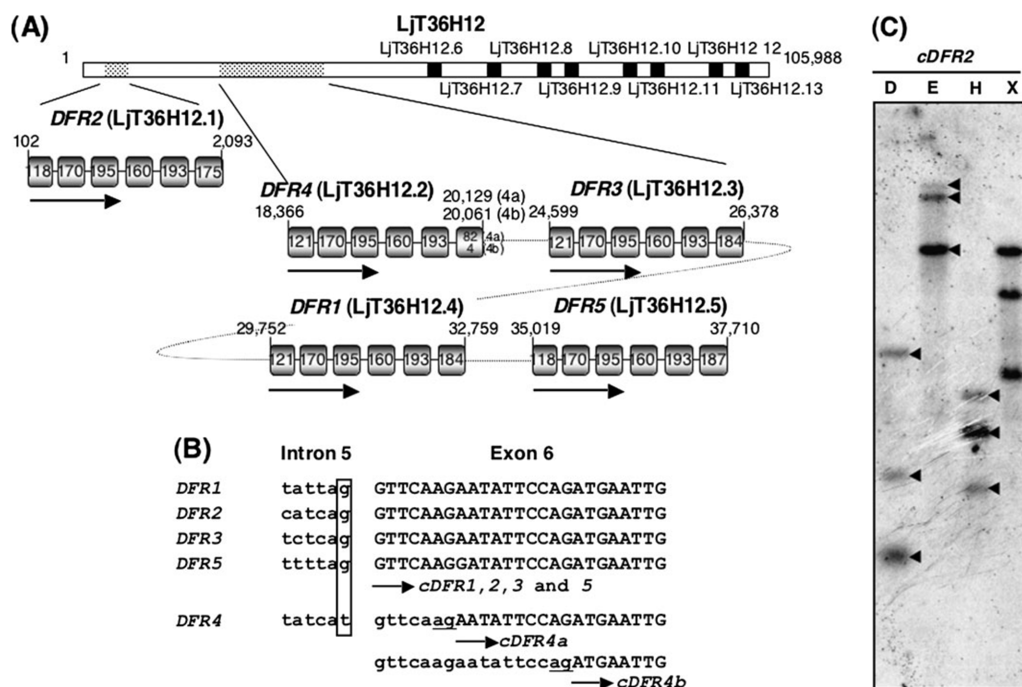

Fig. 4. Structure and organization of *DFR* genes in the *L. japonicus* genome. (A) Physical map of *L. japonicus DFR* genes on the TAC clone LjT36H12. The dotted sections represent the regions of the *DFR* genes. The grey boxes and numbers indicate the exons of the *DFR* genes and the base pair of each exon, respectively. The exact sizes of exon 1 and exon 6 of each *DFR* gene are not known because the start and the end of transcription have not been determined. Thus, the sizes of exon 1 and exon 6 are calculated based on the initiation codon and the stop codon of the open reading frame. The black parts indicate the expected genes. (B) A mutation at the 3′-sequence end of intron 5 of the *DFR4* gene was observed (boxed 't'). This mutation causes an unusual splicing event, which results in a shift of the 3′-splice site to 7 bp and 17 bp downstream (underlined). Capital letters and small letters indicate exon and intron sequences, respectively. (C) Southern blot analysis of *DFR* genes. Genomic DNAs (10 μg) digested with *Dra*I, *Eco*RI, *Hin*dIII, or *Xba*I were hybridized with a DIG-labelled *cDFR2* fragment. Arrowheads indicate the positions of the bands.

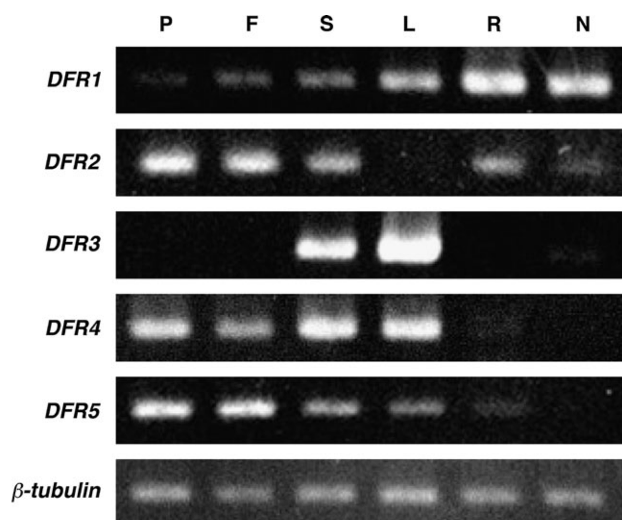

Fig. 5. Accumulation of *DFR* transcripts. Each lane contains equal amounts of cDNA prepared from young seeds and pods (P), flowers (F), stems (S), leaves (L), roots (R), and nodules (N) of 8-week-old *L. japonicus*. The concentration of each cDNA for PCR was adjusted to give an amplification roughly equal to β-*tubulin*. Amplification efficiency was considered to be linear during the PCR (30 cycles for *DFR1*, *DFR2*, *DFR5*, and β-*tubulin*; 40 cycles for others).

and then normalized against the value obtained for DHQ with the same enzyme. The three substrates were presumed to be (+)-enantiomers, but the contamination of stereoisomers was also expected (the optical purity has not been determined), and thus excess substrates were reacted with the enzyme, resulting in the maximum of only 5% yield (data not shown). As shown in Fig. 7, all three enzymes reduced DHM less efficiently than (+)-DHQ [40% to 50% of the activity level to (+)-DHQ]. While DFR2 and DFR3 (Asn type) reduced DHK more efficiently [170% and 350% of the activity level of (+)-DHQ, respectively], DFR5 (Asp type) reduced this substrate less efficiently [70% of the activity level to (+)-DHQ].

Unusual 3-deoxyanthocyanins derived from flavanones accumulate in certain plants, and the FNR activity of DFR proteins was shown to be responsible for their biosynthesis (Halbwirth *et al.*, 2003). To test whether the DFRs of *L. japonicus* have FNR activity, and to analyse the effect of the amino acid residue, Asn or Asp, at the locus described above, the FNR activity of DFR2 (Asn-type DFR) and DFR5 (Asp-type DFR) was measured using the flavanones, naringenin and eriodictyol, as the substrates. Reaction products were extracted with ethyl acetate, and then converted to 3-deoxyanthocyanidins by acid treatment. Apigeninidin and luteolinidin, which are derivatives of naringenin

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

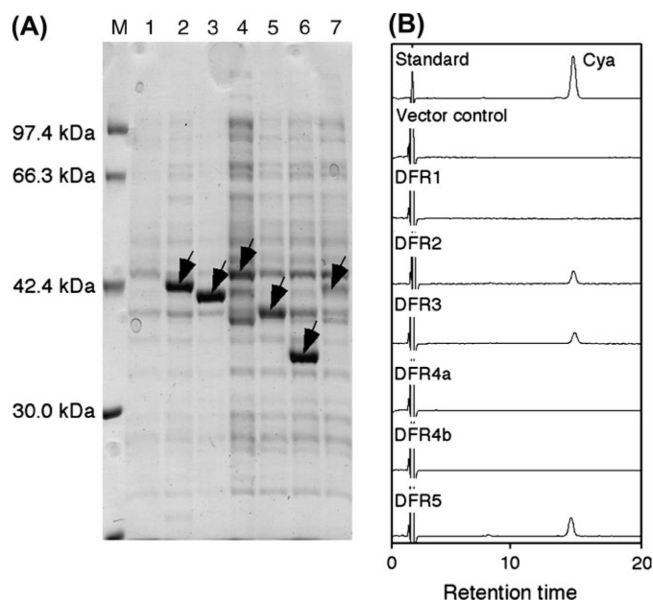

Fig. 6. Heterologous expression of His-tagged six DFR isozymes. *Escherichia coli* strain BL21(DE3) harbouring a pET28a or pET28a–*cDFR* vector construct was incubated with isopropyl-thio-β-D-galactoside before preparing total protein extracts. (A) Protein samples were separated by SDS-PAGE and stained by Coomassie blue: M, molecular mass marker; lane 1, total proteins of the bacteria harbouring the vector only; lanes 2–7, total proteins of the bacteria harbouring *cDFR1* (lane 2), *cDFR2* (lane 3), *cDFR3* (lane 4), *cDFR4a* (lane 5), *cDFR4b* (lane 6), or *cDFR5* (lane 7). Arrowheads indicate each His-tagged DFR protein. (B) HPLC profiles of acid-treated DFR reaction products. Upon acid-treatment, reaction products of DFR2, DFR3, and DFR5 from DHQ as a substrate were converted to cyanidin (Cya), while no product was observed in the reaction with DFR1, DFR4a, or DFR4b.

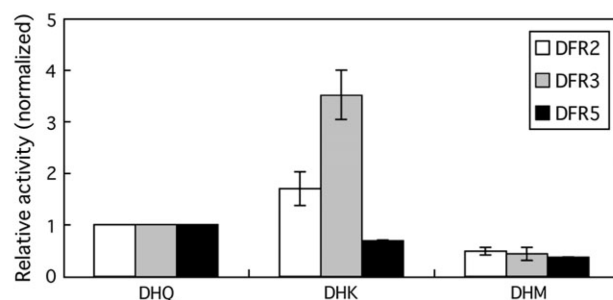

Fig. 7. Relative activity of *L. japonicus* DFRs towards dihydroflavonols. For each enzyme preparation, product amounts were normalized to that obtained from (+)-DHQ to compare the relative activities to each substrate. Thus, the product amount from (+)-DHQ is set at 1.0 for each DFR. The graph represents the average of three independent replicate experiments.

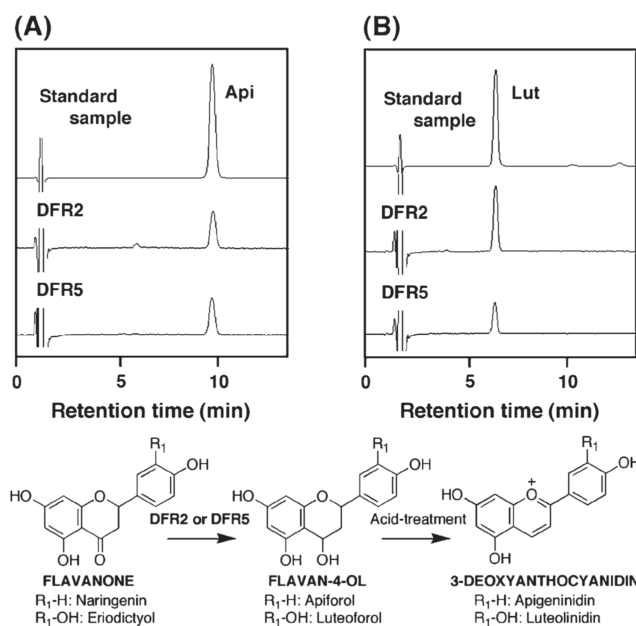

Fig. 8. Flavanone 4-reductase activity of DFRs. After acid-treatment, reaction products of DFRs (DFR2 and DFR5) from naringenin (A) and eriodictyol (B) were converted to apigeninidin (Api) and luteolinidin (Lut), respectively.

and eriodictyol, respectively, were detected by HPLC (Fig. 8A, B).

Enzyme activity of DFR4 *revertant*

A mutation at the splicing accepter site of intron 5 of the *DFR4* gene caused an unusual splicing event, resulting in two splicing variants, *DFR4a* and *DFR4b*, with a frameshift mutation. Mutagenesis of *cDFR4a* using the unique restriction enzyme site elimination technique (Deng and Nickoloff, 1992) was performed to complement the seven nucleotides that had been deleted by the unusual splicing event. After sequence confirmation, the DFR4a revertant protein was expressed in *E. coli*. As shown in Fig. 9A, the molecular mass of DFR4a+6×His (∼37 kDa) shifted to the expected size (∼43 kDa) after the mutagenesis. Enzyme activity of the DFR4a revertant was tested using (±)-DHQ as a substrate (Fig. 9B), and the result clearly showed that DFR4a revertant protein recovered enzyme activity.

Phylogenetic analysis

A phylogenetic tree of *DFR* genes was constructed based on nucleotide sequences of exon 3 and parts of exons 2 and 4. Exon 3 encodes a region that has been postulated to control DFR substrate specificity. The result is shown in Fig. 10 with bootstrap values.

*DFR*s of monocots and eudicots were clearly classified into different branches. Eudicot *DFR*s further diverge into two groups. This phylogenetic tree indicates that Asn-type DFRs are widely distributed in plants, whereas Asp-type DFRs are found in limited plant species that are scattered throughout the eudicots. Moreover, like DFR1 of *L. japonicus*, DFRs other than Asn- or Asp-types are present in a few plant species. *DFR*s of the Fabaceae constitute a monophyletic group, and *L. japonicus*, *L. corniculatus*, and *Medicago truncatula* have both Asn- and Asp-type DFRs. Within this group, some DFRs found in different genera, e.g. *L. japonicus DFR5* and *M. truncatula DFR2*, and *L. japonicus DFR1* and *M. truncatula DFR1*,

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

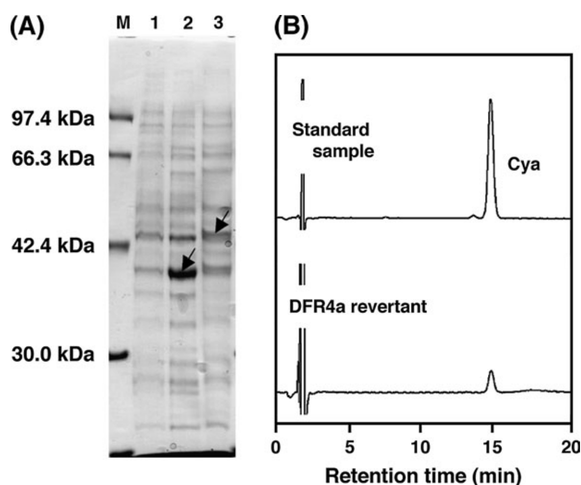

Fig. 9. SDS-PAGE analysis and enzyme assay of His-tagged DFR4a revertant protein. (A) Total protein of the bacteria harbouring the vector only (lane 1), His-tagged DFR4a (lane 2), and His-tagged DFR4a revertant (lane 3). Arrowheads indicate the His-tagged DFR4a and the His-tagged DFR4a revertant proteins, respectively. After mutagenesis, the molecular mass of DFR4a protein had shifted from ~37 kDa to ~43 kDa. M, molecular mass marker. (B) DFR activity of the DFR4a revertant protein. The DFR4a revertant protein reduced DHQ, and cyanidin (Cya) was detected after acid treatment of the reaction products.

are classified into the same branch. On the other hand, *L. japonicus DFR2*, *DFR3*, and *DFR4* have their respective close orthologues in *L. corniculatus*, a tetraploid *Lotus* species.

Discussion

DFR is the first committed enzyme in the biosynthesis of anthocyanins and CTs, general and physiologically significant plant metabolites. Although single or multiple (up to three) *DFR* genes in a few single plant species are known (Beld *et al.*, 1989; Shirley *et al.*, 1992; Bongue-Bartelsman *et al.*, 1994; Sparvoli *et al.*, 1994; Chen *et al.*, 1998; Inagaki *et al.*, 1999), no detailed characterization of all the isozymes in a single plant species has been performed, mainly due to the lack of information on whole genome sequences of most plant species. The *Arabidopsis* genome contains only a single *DFR* gene (Shirley *et al.*, 1992), and the CT and anthocyanin components seem to be simple as well. In the leguminous plants, these metabolites show more complex patterns and more diverse physiological functions. Therefore, a comprehensive analysis of gene structure, expression, and functional identification of all *DFR* paralogues was attempted in a model leguminous plant. As a result, five *DFR* genes, the largest number so far reported, were found to be located in tandem in a single locus of the *L. japonicus* genome. This type of gene structure of isozymes involved in flavonoid biosynthesis is immediately reminiscent of the *L. japonicus chalcone isomerase* (*CHI*) genes: three genes of legume-specific type II *CHI* of the 5-deoxyflavonoid pathway and a type I *CHI* gene of the general 5-hydroxyflavonoid pathway constitute a tandem cluster (Shimada *et al.*, 2003). In addition, genes encoding enzymes involved in the flavonoid pathway, for example, chalcone synthase, chalcone polyketide reductase, and isoflavone reductase, form respective clusters in the *L. japonicus* genome (N Shimada *et al.*, unpublished data).

The DFR proteins translated from the five *L. japonicus DFR* genes displayed subtle but characteristic differences in their amino acid sequence motifs, as well as in the mode of catalytic action. First, substrate preferences of *L. japonicus* DFR2, DFR3, and DFR5 were in agreement with the importance of Asn or Asp at the postulated region controlling the substrate recognition (Beld *et al.*, 1989). DFR2 and DFR3 (Asn-type DFRs) show a higher activity towards DHK than DHQ, whereas DFR5 (Asp-type) reduces DHK less readily than DHQ. In a recent report on the DFRs of another model legume, the activity of *M. truncatula* DFR1 (Asn-type) towards DHK was much higher than the DFR2 (Asp-type) of this plant (Xie *et al.*, 2004). While gene duplication generates functional redundancy and, usually, one of the duplicated genes becomes a pseudogene, duplicated genes can be stably maintained when they acquire different functions. Thus, the presence of both types of DFRs in *L. japonicus* and *M. truncatula* genomes suggests subfunctionalization of *DFR* paralogues in these plants. Second, the DFR1 protein failed to show the DFR activity. In a previous study, the catalytic activity of heterologously expressed DFR was dependent on the expression system (Martens *et al.*, 2002). In this study, however, other DFRs (DFR2, DFR3, and DFR5) expressed in the *E. coli* system possessed the catalytic activity. Thus it is likely that the *DFR1* gene does not encode active DFR. The mutation of the Asn or Asp to Ser in DFR1 could be responsible for the lack of activity. On the other hand, a few DFRs of neither Asn- nor Asp-type have been found in some plant species, although the catalytic activity was not clearly assigned (Fig. 10). Further studies with site-directed mutagenesis are needed to verify the contribution of a specialized amino acid residue to the activity and substrate preference of DFRs. Third, a very striking feature of the *DFR4* gene was a mutation at the splicing accepter site of intron 5, yielding two inactive proteins. Further, interestingly, the catalytic activity was restored in the DFR4a revertant protein, suggesting that the mutation at the splicing accepter site of the *DFR4* gene occurred recently in evolution, because the evolutionary rate of a pseudogene is thought to be much higher than that of a functional gene. Taken together, the organization of the *L. japonicus DFR* gene structures, like those of *CHI*s and other genes of the flavonoid pathway, provides strong support for the idea that (local) gene duplication is an important event in the creation of novel gene functions and, especially, diversity in plant secondary metabolism (Shimada *et al.*, 2003).

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

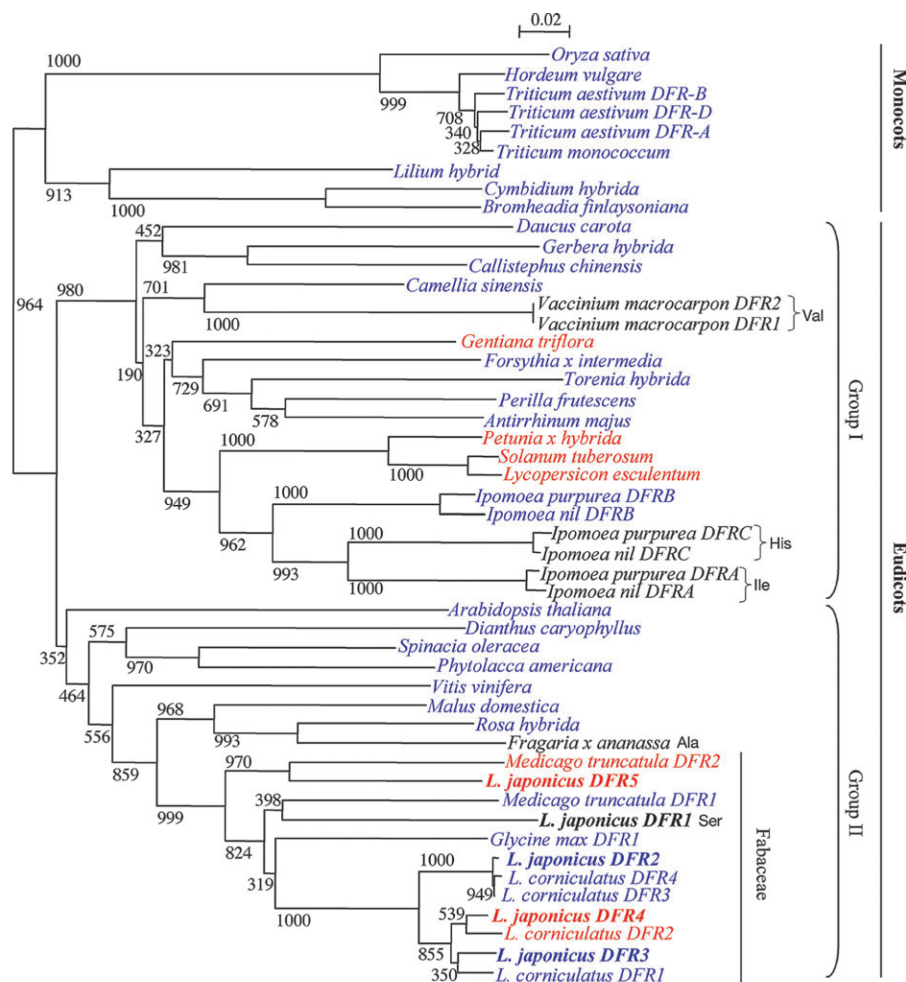

Fig. 10. A phylogenetic tree of *DFR*s constructed by the Neighbor–Joining method. The 321 nucleotides covering a region that has been postulated to control substrate specificity were analysed using the CLUSTAL W program. The bootstrap replicates were 1000 (values are given at the nodes). Asn-type DFRs, Asp-type DFRs, and DFRs of neither Asn- nor Asp-type are shown in blue, red, and black, respectively. For accession numbers of the *DFR* sequences, see Materials and methods.

When the *DFR* genes from broader plant species are compared, another interesting aspect of the evolution of *DFR* genes can be appreciated. The phylogenetic tree (Fig. 10) shows that the majority of the DFRs in the plant kingdom are the Asn type, whereas Asp-type DFRs are distributed only to limited and distant taxa. Therefore, Asp-type DFRs are likely to have evolved convergently from Asn-type DFRs. The Asp-type DFR of *L. japonicus*, DFR5, is similar to that of *M. truncatula*, and thus it is speculated that an ancestral Asp-type DFR might have been present prior to the speciation of *Medicago* and *Lotus*. Also, the *DFR* paralogues of *L. japonicus*, i.e. *DFR2*, *DFR3*, and *DFR4*, are very similar to the corresponding *DFR*s of *L. corniculatus*, suggesting that the duplication process generating these three *DFR*s occurred prior to the speciation of *L. japonicus* and *L. corniculatus*. This kind of relationship was previously reported in *DFR*s of the Japanese morning glory (*Ipomoea nil*) and the common morning glory (*Ipomoea purpurea*) (Inagaki *et al.*, 1999). It will be interesting to examine whether the abnormal splicing takes place in *L. corniculatus DFR2*, the nearest orthologue of *L. japonicus DFR4*.

The anthocyanin and CT compositions of the stems of *L. japonicus* do not simply reflect the substrate preference of the DFRs expressed there. DHQ-derived structures comprise the major, and DHM-derived the minor, anthocyanins, and the CT units are roughly equal amounts of DHQ- and DHM-derived flavan-3-ols. The expression of DFR5 (Asp type) may well correlate with such a metabolite distribution, but Asn-type DFRs (DFR2 and DFR3), with a strong preference for DHK, are also highly expressed. Even inactive DFRs (DFR1, DFR4a, and DFR4b) are expressed in the stem. If the flavonols and anthocyanins/CTs are produced in the same cells and the production of three dihydroflavonols through the actions of flavonoid 3′-hydroxylase and 3′,5′-hydroxylase is nearly equal, the

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

observed distribution of major DHK-derived flavonols and DHQ- and DHM-derived anthocyanins/CTs could be explained by the competition between flavonol synthase and DFR for three dihydroflavonols. However, the flux of substrates towards each pathway may be controlled through the regulation of gene expression of preceding enzymes, and inter- and intracellular compartmentation of the pathway, possibly involving enzyme complexes, should also be important factors determining the flavonoid compositions in the organs. Detailed studies on the constituents of the organs of *L. japonicus* and inspection of the expression patterns of DFR and other isozymes of the flavonoid pathway are expected to elucidate the complex regulation mechanisms of biosynthesis in the future. The framework for such studies at the genomic level is provided in this and related studies (Shimada *et al*., 2003).

An additional puzzling discrepancy between the product accumulation and enzyme expression was that, while leaves of *L. japonicus* accumulate flavonols but do not contain anthocyanins, CTs, or flavan-3,4-diols (Skadhauge *et al.*, 1997), a high level accumulation of *DFR3* transcripts was observed in leaves. Recently, a similar observation was reported that *Spinacia oleracea* and *Phytolacca americana*, both accumulating flavonols but not anthocyanins/CTs, do have DFRs (Shimada *et al.*, 2004). Surprisingly, an *in vitro* reaction using recombinant anthocyanidin synthase (ANS) produced not only anthocyanidin but also a flavonol from flavan-3,4-diol (Turnbull *et al.*, 2003). The expression of *ANS* in *L. japonicus* leaves has not been tested, but the expression of *DFR3* may suggest a possible role of *DFR*s in flavonol biosynthesis through the ANS reaction.

To date, several ethylmethane sulphonate-induced or spontaneous mutant lines of *L. japonicus*, producing neither anthocyanins nor CTs or producing lesser levels of anthocyanins, have been isolated. In some mutants, the expression of *DFR*s was specifically suppressed (N Shimada *et al*., unpublished data). RT-PCR analysis also showed that the expression patterns of all *DFR* paralogues of *L. japonicus* differed among the organs (Fig. 5). Thus, it would be expected that some DFR isozymes might be specialized to anthocyanin synthesis or the CT pathway. The gene duplication has possibly allowed the adaptation of the enzymes for specialized functions and contributed to the divergence of plant metabolisms. The present identification of all the *DFR*s in *L. japonicus* suggests the evolutionary processes of *DFR* genes, and predicts subfunctionalization of *DFR* paralogues. To analyse the function of each DFR isozyme in the flavonoid pathway, detailed biochemical characterization of purified DFR isozymes is in progress.

Acknowledgements

This work was supported by the 21st Century COE programme ‘Bioresource Utilization Based on Microbial Symbiotic Systems’ from the Ministry of Education, Culture, Sports, and Science and the project ‘Development of Fundamental Technologies for Controlling the Material Production Process of Plants’ from the New Energy and Industrial Technology Development Organization. This work was conducted at the Life Science Research Center (Shonan Campus, Nihon University). The authors wish to thank Kentarou Kondoh (Nihon University) for technical assistance.

References

Aida R, Yoshida K, Kondo T, Kishimoto S, Shibata M. 2000. Copigmentation gives bluer flowers on transgenic torenia plants with the antisense dihydroflavonol-4-reductase gene. *Plant Science* **160,** 49–56.

Akashi T, Sawada Y, Shimada N, Sakurai N, Aoki T, Ayabe S. 2003. cDNA cloning and biochemical characterization of *S*-adenosyl-L-methionine: 2,7,4′-trihydroxyisoflavanone 4′-*O*-methyltransferase, a critical enzyme of the legume isoflavonoid phytoalexin pathway. *Plant and Cell Physiology* **44,** 103–112.

Asamizu E, Kato T, Sato S, Nakamura Y, Kaneko T, Tabata S. 2003. Structural analysis of a *Lotus japonicus* genome. IV. Sequence features and mapping of seventy-three TAC clones which cover the 7.5 Mb regions of the genome. *DNA Research* **10,** 115–122.

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

Bavage AD, Davies IG, Robbins MP, Morris P. 1997. Expression of an *Antirrhinum* dihydroflavonol reductase gene results in changes in condensed tannin structure and accumulation in root cultures of *Lotus corniculatus* (bird’s foot trefoil). *Plant Molecular Biology* **35,** 443–458.

Beld M, Martin C, Huits H, Stuitje AR, Gerats AG. 1989. Flavonoid synthesis in *Petunia hybrida*: partial characterization of dihydroflavonol-4-reductase genes. *Plant Molecular Biology* **13,** 491–502.

Bernhardt J, Stich K, Schwarz-Sommer Z, Saedler H, Wienand U. 1998. Molecular analysis of a second functional *A1* gene (*dihydroflavonol 4-reductase*) in *Zea mays*. *The Plant Journal* **14,** 483–488.

Bongue-Bartelsman M, O’Neill SD, Tong Y, Yoder JI. 1994. Characterization of the gene encoding dihydroflavonol 4-reductase in tomato. *Gene* **138,** 153–157.

Charrier B, Coronado C, Kondorosi A, Ratet P. 1995. Molecular characterization and expression of alfalfa (*Medicago sativa* L.) flavanone-3-hydroxylase and dihydroflavonol-4-reductase encoding genes. *Plant Molecular Biology* **29,** 773–786.

Chen MS, SanMiguel P, Bennetzen JL. 1998. Sequence organization and conservation in *sh2/a1*-homologous regions of sorghum and rice. *Genetics* **148,** 435–443.

De Jaeger G, Buys E, Eeckhout D, Bruyns AM, De Neve M, De Wilde C, Gerats T, Van Montagu M, Fischer R, Depicker A. 1997. Use of phage display for isolation and characterization of single-chain variable fragments against dihydroflavonol 4-reductase from *Petunia hybrida*. *FEBS Letters* **403,** 116–122.

Dellus V, Heller W, Sadermann H, Scalbert A. 1997. Dihydroflavonol 4-reductase activity in lignocellulosic tissues. *Phytochemistry* **45,** 1415–1418.

Deng WP, Nickoloff JA. 1992. Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Analytical Biochemistry* **200,** 81–88.

Fischer D, Stich K, Britsch L, Grisebach H. 1988. Purification and characterization of (+)dihydroflavonol (3-hydroxyflavanone)

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

4-reductase from flowers of *Dahlia variabilis*. *Archives of Biochemistry and Biophysics* **264,** 40–47.

Fischer TC, Halbwirth H, Meisel B, Stich K, Forkmann G. 2003. Molecular cloning, substrate specificity of the functionally expressed dihydroflavonol 4-reductases from *Malus domestica* and *Pyrus communis* cultivars and the consequences for flavonoid metabolism. *Archives of Biochemistry and Biophysics* **412,** 223–230.

Forkmann G, Ruhnau B. 1987. Distinct substrate-specificity of dihydroflavonol 4-reductase from flowers of *Petunia hybrida*. *Zeitschrift fur Naturforschung C – a Journal of Biosciences* **42,** 1146–1148.

Gong Z, Yamazaki M, Sugiyama M, Tanaka Y, Saito K. 1997. Cloning and molecular analysis of structural genes involved in anthocyanin biosynthesis and expressed in a forma-specific manner in *Perilla frutescens*. *Plant Molecular Biology* **35,** 915–927.

Halbwirth H, Martens S, Wienand U, Forkmann G, Stich K. 2003. Biochemical formation of anthocyanins in silk tissue of *Zea mays*. *Plant Science* **164,** 489–495.

Handberg K, Stougaard J. 1992. *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. *The Plant Journal* **2,** 487–496.

Helariutta Y, Elomaa P, Kotilainen M, Seppanen P, Teeri TH. 1993. Cloning of cDNA coding for dihydroflavonol-4-reductase (DFR) and characterization of dfr expression in the corollas of *Gerbera hybrida* var. Regina (Compositae). *Plant Molecular Biology* **22,** 183–193.

Heller W, Forkmann G, Britsch L, Grisebach H. 1985. Enzymatic reduction of (+)-dihydroflavonols to flavan-3,4-*cis*-diols with flower extracts from *Matthiola incana* and its role in anthocyanin biosynthesis. *Planta* **165,** 284–287.

Himi E, Noda K. 2004. Isolation and location of three homoeologous *dihydroflavonol-4-reductase* (*DFR*) genes of wheat and their tissue-dependent expression. *Journal of Experimental Botany* **55,** 365–375.

Inagaki Y, Johzuka-Hisatomi Y, Mori T, Takahashi S, Hayakawa Y, Peyachoknagul S, Ozeki Y, Iida S. 1999. Genomic organization of the genes encoding dihydroflavonol 4-reductase for flower pigmentation in the Japanese and common morning glories. *Gene* **226,** 181–188.

Johnson ET, Ryu S, Yi H, Shin B, Cheong H, Choi G. 2001. Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. *The Plant Journal* **25,** 325–333.

Johnson ET, Yi H, Shin B, Oh BJ, Cheong H, Choi G. 1999. *Cymbidium hybrida* dihydroflavonol 4-reductase does not efficiently reduce dihydrokaempferol to produce orange pelargonidin-type anthocyanins. *The Plant Journal* **19,** 81–85.

Kaneko T, Asamizu E, Kato T, Sato S, Nakamura Y, Tabata S. 2003. Structural analysis of a *Lotus japonicus* genome. III. Sequence features and mapping of sixty-two TAC clones which cover the 6.7 Mb regions of the genome. *DNA Research* **10,** 27–33.

Kato T, Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. 2003. Structural analysis of a *Lotus japonicus* genome. V. Sequence features and mapping of sixty-four TAC clones which cover the 6.4 Mb regions of the genome. *DNA Research* **10,** 277–285.

Lacombe E, Hawkins S, Van Doorsselaere J, Piquemal J, Goffner D, Poeydomenge O, Boudet AM, Grima-Pettenati J. 1997. Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *The Plant Journal* **11,** 429–441.

Liu YG, Shirano Y, Fukaki H, Yanai Y, Tasaka M, Tabata S, Shibata D. 1999. Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. *Proceedings of the National Academy of Sciences, USA* **96,** 6535–6540.

Martens S, Teeri T, Forkmann G. 2002. Heterologous expression of dihydroflavonol 4-reductases from various plants. *FEBS Letters* **531,** 453–458.

Nakamura Y, Kaneko T, Asamizu E, Kato T, Sato S, Tabata S. 2002. Structural analysis of a *Lotus japonicus* genome. II. Sequence features and mapping of sixty-five TAC clones which cover the 6.5 Mb regions of the genome. *DNA Research* **9,** 63–70.

Nishimura R, Hayashi M, Wu GJ, *et al.* 2002. HAR1 mediates systemic regulation of symbiotic organ development. *Nature* **420,** 426–429.

Pankhurst CE, Jones WT. 1979. Effectiveness of *Lotus* root nodules. 2. Relationship between root nodule effectiveness and '*in vitro*' sensitivity of fast-growing *Lotus* rhizobia to flavolans. *Journal of Experimental Botany* **30,** 1095–1107.

Perriere G, Gouy M. 1996. WWW-query: an on-line retrieval system for biological sequence banks. *Biochimie* **78,** 364–369.

Peters DJ, Constabel CP. 2002. Molecular analysis of herbivore-induced condensed tannin synthesis: cloning and expression of dihydroflavonol reductase from trembling aspen (*Populus tremuloides*). *The Plant Journal* **32,** 701–712.

Robbins MP, Bavage AD, Strudwicke C, Morris P. 1998. Genetic manipulation of condensed tannins in higher plants. II. Analysis of birdsfoot trefoil plants harboring antisense dihydroflavonol reductase constructs. *Plant Physiology* **116,** 1133–1144.

Sato S, Kaneko T, Nakamura Y, Asamizu E, Kato T, 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 Research* **8,** 311–318.

Schauser L, Roussis A, Stiller J, Stougaard J. 1999. A plant regulator controlling development of symbiotic root nodules. *Nature* **402,** 191–195.

Shimada N, Akashi T, Aoki T, Ayabe S. 2000. Induction of isoflavonoid pathway in the model legume *Lotus japonicus*: molecular characterization of enzymes involved in phytoalexin biosynthesis. *Plant Science* **160,** 37–47.

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

Shimada S, Takahashi K, Sato Y, Sakuta M. 2004. Dihydroflavonol 4-reductase cDNA from non-anthocyanin-producing species in the Caryophyllales. *Plant and Cell Physiology* **45,** 1290–1298.

Shirley BW, Hanley S, Goodman HM. 1992. Effects of ionizing radiation on a plant genome: analysis of two Arabidopsis *transparent testa* mutations. *The Plant Cell* **4,** 333–347.

Skadhauge B, Gruber MY, Thomsen KK, von Wettstein D. 1997. Leucocyanidin reductase activity and accumulation of proanthocyanidins in developing legume tissues. *American Journal of Botany* **84,** 494–503.

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C. 1994. Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). *Plant Molecular Biology* **24,** 743–755.

Spencer CM, Cai Y, Martin R, Gaffney SH, Goulding PN, Magnolato D, Lilley TH, Haslam E. 1988. Polyphenol complexation – some thoughts and observations. *Phytochemistry* **27,** 2397–2409.

Stich K, Eidenberger T, Wurst F, Forkmann G. 1992. Enzymatic conversion of dihydroflavonols to flavan-3,4-diols using flower extracts of *Dianthus caryophyllus* L (carnation). *Planta* **187,** 103–108.

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024

Tanaka Y, Fukui Y, Fukuchi-Mizutani M, Holton TA, Higgins E, Kusumi T. 1995. Molecular cloning and characterization of *Rosa hybrida* dihydroflavonol 4-reductase gene. *Plant and Cell Physiology* **36,** 1023–1031.

Tanaka Y, Yonekura K, Fukuchi-Mizutani M, Fukui Y, Fujiwara H, Ashikari T, Kusumi T. 1996. Molecular and biochemical characterization of three anthocyanin synthetic enzymes from *Gentiana triflora*. *Plant and Cell Physiology* **37,** 711–716.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22,** 4673–4680.

Turnbull JJ, Nagle MJ, Seibel JF, Welford RW, Grant GH, Schofield CJ. 2003. The C-4 stereochemistry of leucocyanidin substrates for anthocyanidin synthase affects product selectivity. *Bioorganic and Medicinal Chemistry Letters* **13,** 3853–3857.

Xie DY, Jackson LA, Cooper JD, Ferreira D, Paiva NL. 2004. Molecular and biochemical analysis of two cDNA clones encoding dihydroflavonol-4-reductase from *Medicago truncatula*. *Plant Physiology* **134,** 979–994.

Downloaded from https://academic.oup.com/jxb/article/56/419/2573/532069 by guest on 23 April 2024