# Molecular Cloning of the Genes Involved in Anthocyanin Biosynthesis in Camellia japonica

Tateishi, Nobumine
Laboratory of Horticultural Science, Division of Agricultural Botany, Department of Plant Resources, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University

Ozaki, Yukio
Laboratory of Agricultural Ecology, Division of Agricultural Ecology, Department of Plant Resources, Faculty of Agriculture, Kyushu University

Okubo, Hiroshi
Laboratory of Agricultural Ecology, Division of Agricultural Ecology, Department of Plant Resources, Faculty of Agriculture, Kyushu University

https://doi.org/10.5109/17796

出版情報：九州大学大学院農学研究院紀要. 55 (1), pp.21-28, 2010-02-26. Faculty of Agriculture, Kyushu University
バージョン：
権利関係：

J. Fac. Agr., Kyushu Univ., **55** (1), 21–28 (2010)

# Molecular Cloning of the Genes Involved in Anthocyanin Biosynthesis in *Camellia japonica*

**Nobumine TATEISHI[1]*, Yukio OZAKI[2] and Hiroshi OKUBO**

Laboratory of Horticultural Science, Division of Agricultural Botany,
Department of Plant Resources, Faculty of Agriculture,
Kyushu University, Fukuoka 812–8581, Japan
(*Received October 26, 2009 and accepted November 19, 2009*)

Partial cDNA sequences of three anthocyanin biosynthetic genes (*F3H*, *flavanone 3–hydroxylase*; *DFR*, *dihydroflavonol 4–reductase*; *ANS*, *anthocyanidin synthase*) were isolated from the petals of *Camellia japonica*. Their deduced partial amino acid sequences shared high homologies with those of woody plant species (CjF3Ha, 98.0%, CjF3Hb, 91.2% and CjDFR, 99.0% with *Camellia sinensis*; CjANS, 90.3% with *Rhododendron* × *pulchrum*). Some important amino acid residues for enzymatic activities were also conserved in the isolated clones, suggesting that the genes we identified in this study were the homologues of *C. japonica*. Gene–specific primer pairs were designed based on each partial cDNA sequence. The application of these primer pairs to RT–PCR analyses was tested.

## INTRODUCTION

The primary reason of the bright flower colours of plants is to be a signal to attract insects and birds for successful pollination by emphasising their floral tissue against a background of vegetation (Glover, 2007). The major pigments that cause flower colour are carotenoids, flavonoids and betalains (Davies, 2009). Flavonoids also have some additional important functions, like defence against pathogens and predators, prevention from damaging by UV light, action as antioxidants, and involvement in pollen development and germination (Glover, 2007). Anthocyanins are one of the plant–derived flavonoid compounds and are responsible for colours ranging from pink and red to purple and deep blue (Deroles, 2009). Most detailed studies about anthocyanin biosynthesis have been achieved in *Antirrhinum* and *Petunia* as models and many genes involved in the biosynthetic pathway have been isolated so far (Martin *et al*., 1991; Holton and Cornish, 1995).

*Camellia japonica* has a long history as a representative woody ornamental plant in Japan (Tuyama, 1968a). Wild type *C. japonica* has a single–petalled red flowers, and two major anthocyanins were identified in the petals, namely cyanidin 3–glucoside and cyanidin 3–galactoside (Sakata *et al*., 1986, 1987). However, molecular mechanism dominating their flower pigmentation has not been reported, and the breeding program of camellia plants still depends on the selection of the seedlings that appear by chance. Molecular information controlling flower characters of camellia must contribute to an efficient breeding system of the plant.

In this paper, we isolated partial cDNA sequences of three anthocyanin biosynthetic genes in *C. japonica*, *flavanone 3–hydroxylase* (*F3H*), *dihydroflavonol 4–reductase* (*DFR*) and *anthocyanidin synthase* (*ANS*), and confirmed that their gene–specific primer pairs are available for RT–PCR analyses.

## MATERIALS AND METHODS

### Plant materials

Fully expanded flowers of wild *C. japonica* accession (collected in Kurose, Fukue Island, Nagasaki Prefecture, Japan), 'Tamanoura' and 'Hatsu–arashi' were picked immediately at the beginning of anthesis. The petals of 'Tamanoura' were separated into red and white marginal picotee parts, and only the red parts were used in this study. The petals were frozen promptly using liquid nitrogen and stored at −80 °C until use for RNA extraction.

### RNA extraction

RNA extraction was carried out following Kiefer *et al*. (2000)'s manner with some modifications. Frozen tissues (100–150 mg) were ground to fine powder with mortar and pestle using liquid nitrogen and 800 μl pre–warmed (65 °C) extraction buffer [100 mM Tris–HCl (pH 8.0), 25 mM EDTA, 2 M NaCl, 2% CTAB (w/v), 2% polyvinylpyrrolidone (w/v), 0.5% spermidine (w/v) and 2% ß–mercaptoethanol (v/v)] were added followed by incubation at 65 °C for 10 min. One hundred μl of Nucleon PhytoPure DNA extraction resin (GE Healthcare, England) and 400 μl of chloroform/isoamylalcohol (24:1) were added and the sample tubes were kept on a shaker for 10 min at room temperature. After centrifugation at 9,000 × g for 10 min at 4 °C, aqueous phase was washed with 500 μl of chloroform/isoamylalcohol (24:1) at least three times. The aqueous phase, to which 500 μl of isopropanol were added, were incubated on ice for 1 h, followed by centrifugation at 9,000 × g for 10 min at 4 °C. RNA pellets were dissolved in 100 μl DEPC water with

[1] Laboratory of Horticultural Science, Division of Agricultural Botany, Department of Plant Resources, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Fukuoka 812–8581, Japan
[2] Laboratory of Agricultural Ecology, Division of Agricultural Ecology, Department of Plant Resources, Faculty of Agriculture, Kyushu University, Fukuoka 811–2307, Japan
* Corresponding author (E–mail: nobumine2006@yahoo.co.jp)

2 M of LiCl and incubated at 4 °C overnight. After centrifugation at 9,000 × g for 30 min at 4 °C, the RNA pellets were treated with DNase I (Roche Diagnostics, Germany) at 37 °C for 20 min. Equal volume of isopropanol were added for RNA precipitation, followed by centrifugation at 9,000 × g for 2 min at 4 °C. After two times wash with 70% EtOH, the pellets were dried and dissolved in DEPC water. The RNA concentration and purity were evaluated with the absorbance at 260 nm.

**RT–PCR analyses**

cDNA synthesis and RT–PCR were carried out using TaKaRa RNA PCR Kit (AMV) Ver. 3.0 (Takara, Japan) following the manufacturer's instructions. We mixed two kinds of primers in the ratio of Oligo dT–adaptor primer:Random 9 mers = 4:1 and used for reverse tran-

**Table 1.** Primer pairs designed and used in this study

| Primer | Gene | | Sequence |
|---|---|---|---|
| Degenerate primer | *F3H* | Forward | 5'–ATG TCC GGT GGB AAR AAR GG–3' |
| | | Reverse | 5'–TTG CTC ATC TTC YTC YKG TAC–3' |
| | *DFR* | Forward | 5'–RAG GAY CCY GAG AAT GAR G–3' |
| | | Reverse | 5'–GCT GTA YTT GAA YTY GAA YCC–3' |
| | *ANS* | Forward | 5'–TTG AGT GGS AGG ATT AYT TYT TYC–3' |
| | | Reverse | 5'–GGT TCG CAG AAV ATH GCC C–3' |
| Gene–specific primer | *CjF3Ha* | Forward | 5'–ACG GAG ACC TAC AGC GAG AA–3' |
| | | Reverse | 5'–ATG ATC CGC ATT CTT GAA CC–3' |
| | *CjDFR* | Forward | 5'–AAC AAC CCA TTT TCG ACG AG–3' |
| | | Reverse | 5'–TTG TAC TCG GGC CAT TTC TC–3' |
| | *CjANS* | Forward | 5'–ACG CAA AGC AAC TAC GAG G–3' |
| | | Reverse | 5'–CTA CCG TGT CAC CAA TGT GC–3' |

B=C+G+T; R=A+G; Y=C+T; K=G+T; S=C+G; V=A+C+G; H=A+C+T

```
        10          20          30          40          50          60
TGGATTCATC GTTTCCAGTC ATCTCCAGGG AGAAGCAGTG CAAGACTGGA GAGAAATAGT

        70          80          90         100         110         120
GACCTACTTC TCATACCCGA TCCGGGCCCG GGACTATTCA AGATGGCCCG ACAAGCCCGA

       130         140         150         160         170         180
AGGGTGGAGG GCTGTGACGG AGACCTACAG CGAGAAATTG ATGGACTTGG CTTGCAAGTT
                  CjF3Ha F
       190         200         210         220         230         240
GCTGGAGGTG TTGTCTGAGG CCATGGACCT TGAGAAGGAG GCTCTTACAA AAGCCTGTGT

       250         260         270         280         290         300
TGATATGGAT CAGAAGGTGG TTGTAAATTT CTACCCGAAA TGCCCACAAC CCGACCTCAC

       310         320         330         340         350         360
GCTCGGACTC AAGCGACACA CGGATCCGGG TTCCATCACC CTGCTCCTCC AGGACCAGGT

       370         380         390         400         410         420
TGGTGGGCTC CAGGCCACTA GAGATGGGGG CAAGACCTGG ATCACGGTTC AGCCCGTGGA

       430         440         450         460         470         480
GGGAGCTTTT GTTGTTAATC TGGGTGACCA TGGTCATTAT CTAAGCAATG GGAGGTTCAA

       490         500         510         520         530         540
GAATGCGGAT CATCAGGCAG TAGTGAACTC CAACTGCAGC CGACTATCAA TCGCTACATT
CjF3Ha R
       550         560         570         580         590         600
CCAGAACCCA GCTCCCGAGG CGACAGTATA CCCACTGAAG ATTAGGGAGG GAGAGAAGCC

       610         620         630         640
GGTTCTTGAA GAGCCAATCA CGTTCGCCGA TAT.......
```

**Fig. 1.** Nucleotide sequence of partial cDNA of *CjF3Ha*. The shadowed boxes represent the primer regions of *CjF3Ha* F and R used in this study.

scriptional reaction. PCR amplification was performed in a total volume of 50 $\mu$l containing 45 ng template cDNA, 0.2 $\mu$M of each primer, 10 $\mu$l of 5 × PCR Buffer, 0.2 mM of each dNTP, 1.25 Unit *TaKaRa Ex Taq* HS polymerase. In case of using degenerate primer pairs, their volumes were calculated according to its degeneracy (Table 1). Finally, RNase Free $H_2O$ was added up to 50 $\mu$l. Amplification was conducted using TaKaRa PCR Thermal Cycler Dice TP–600 (TaKaRa, Japan) with one cycle of 3 min at 94 ºC, followed by 30 cycles of 20 sec at 94 ºC, 20 sec at 56 ºC and 1 min at 72 ºC, finally one cycle of 10 min at 72 ºC. PCR products were separated by electrophoresis in 1.5% (w/v) agarose gels and visualised under UV illumination after staining with ethidium bromide.

### Molecular cloning of *F3H*, *DFR* and *ANS* in *C. japonica*

Degenerate RT–PCR products from wild *C. japonica* were subcloned into pGEM–T Easy Vector (Promega, USA) and transformed into Competent high *Escherichia coli* DH5 $\alpha$ (Toyobo, Japan). After culture on the LB plates containing 100 $\mu$g/ml of ampicillin, 100 $\mu$g/ml of X–Gal and 23.83 $\mu$g/ml of IPTG, only white colonies were selected, and then the plasmids containing inserts were extracted using LaboPass Plasmid Mini Purification Kit (Hokkaido System Science, Japan). BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, USA) and ABI PRISM 310 genetic analyzer (Applied Biosystems, USA) were employed for sequence analyses. The deduced amino acid sequence alignments and phylogenetic trees were obtained by using DNASIS–Mac v3.2 (Hitachi Software Engineering, Japan) and PHYLIP version 3.6 (Felsenstein, 2004), respectively.

## RESULTS

### Molecular cloning of *F3H*, *DFR* and *ANS* in *C. japonica*

We identified 633 bp fragment of partial *CjF3H*. Sequence analysis revealed that it contained multi–gene family. The major one was named *CjF3Ha* (GenBank accession no. AB524883) and minor one was *CjF3Hb*

```
                      10         20         30         40         50
CjF3Ha              1 GFIVSSHLQG EAVQDWREIV TYFSYPIRAR DYSRWPDKPE GWRAVTETYS      50
CjF3Hb              1 GFIVSSHLQG EAVQDWREIV TYFSYPIRAR DYSRWPDKPE EWRAVTEKYS      50
Vitis               1 GFIVSSHLQG EAVQDWREIV TYFSYPLRTR DYSRWPDKPE GWRSVTQEYS      50
Gentiana            1 GFIVSSHLQG EAVRDWREIV TYFSYPIKSR DYSRWPDKPE GWKSVTEKYS      50
Ipomoea nil         1 GFIVSSHLQG EAVKDWREIV TYFSYPVRAR DYSRWPDKPE GWRAVTEKYS      50
Arabidopsis         1 GFIVSSHLQG EAVQDWREIV TYFSYPVRNR DYSRWPDKPE GWVKVTEEYS      50
Oryza sativa        1 GFIVSSHLQG EAVKDWREIV TYFSYPVKSR DYSRWPDKPA GWRAVVEQYS      50
                      ********** *** ****** ******   * *********   *  *   **
                      60         70         80         90         100
CjF3Ha             51 EKLMDLACKL LEVLSEAMDL EKEALTKACV DMDQKVVVNF YPKCPQPDLT     100
CjF3Hb             51 SSLMELACKL LEVLSEAMGL EKEALTNACV DMDQKVVVNY YPKCPQPDLT     100
Vitis              51 EKLMGLACKL LEVLSEAMDL DKDALTNACV DMDQKVVVNF YPQCPQPDLT     100
Gentiana           51 EQLMNLACKL LEVLSESMRL EKEALTKACV DMDQKIVVNF YPKCPQPDLT     100
Ipomoea nil        51 EKLMDLACKL LEVLSEAMGL EKEALSKACV ELDQKLVVNF YPKCPEPDLT     100
Arabidopsis        51 ERLMSLACKL LEVLSEAMGL EKESLTNACV DMDQKIVVNY YPKCPQPDLT     100
Oryza sativa       51 ERLMGLACKL LGVLSEAMGL DTNALADACV DMDQKVVVNF YPKCPQPDLT     100
                        ** ***** * **** * *     *  ***   *** ***  ** ** ****
                      110        120        130        140        150
CjF3Ha            101 LGLKRHTDPG SITLLLQDQV GGLQATRDGG KTWITVQPVE GAFVVNLGDH     150
CjF3Hb            101 LGLKRHTDPG TITLLLQDQV GGLQATRDGG KTWITVQPVE GAFVVNLGDH     150
Vitis             101 LGLKRHTDPG TITLLLQDQV GGLQATRDGG KTWITVQPVE GAFVVNLGDH     150
Gentiana          101 LGLKRHTDPG TITLLLQDQV GGLQATRDGG KSWITVQPVD GAFVVNLGDH     150
Ipomoea nil       101 LGLKRHTDPG TITLLLQDQV GGLQATKDGG KTWITVQPVD GAFVVNLGDH     150
Arabidopsis       101 LGLKRHTDPG TITLLLQDQV GGLQATRDNG KTWITVQPVE GAFVVNLGDH     150
Oryza sativa      101 LGLKRHTDPG TITLLLQDLV GGLQATRDAG KTWITVQPIP GSFVVNLGDH     150
                      **********  ******* * ****** * * * ******   * ********
                      160        170        180        190        200
CjF3Ha            151 GH-------- ---YLSNGRF KNADHQAVVN SNCSRLSIAT FQNPAPEATV     200
CjF3Hb            151 GH-------- ---YLSNGRF KNADHQAVVN SNSSRLSIAT FQNPAPEAIV     200
Vitis             151 GH-------- ---YLSNGRF KNADHQAVVN SNHSRLSIAT FQNPAPEATV     200
Gentiana          151 GH-------- ---YLSNGRF KNADHQAVVN SNYSRLSIAT FQNPAPEATV     200
Ipomoea nil       151 GH-------- ---FLSNGRF KNADHQAVVN SEHSRMSIAT FQNPAPEAKV     200
Arabidopsis       151 GH-------- ---FLSNGRF KNADHQAVVN SNSSRLSIAT FQNPAPDATV     200
Oryza sativa      151 AHIMHLLGNV NLQYLSNGRF KNADHQAVVN SDCCRLSIAT FQNPAPDAMV     200
                       *            ****** ********** *   * **** ****** * *
                      210        220        230        240        250
CjF3Ha            201 YPLKIREGEK PVLEEPITFA D......... .......... ..........     250
CjF3Hb            201 YPLKIREGEK SIMEEPITFP D......... .......... ..........     250
Vitis             201 YPLKIREGEK AVLEGPITFA E......... .......... ..........     250
Gentiana          201 YPLAIRDGEK PVLDEPITFA E......... .......... ..........     250
Ipomoea nil       201 YPLKVREGEK PILEEPITFA E......... .......... ..........     250
Arabidopsis       201 YPLKVREGEK AILEEPITFA E......... .......... ..........     250
Oryza sativa      201 YPLAVRDGEE PILEEPITFA E......... .......... ..........     250
                      ***  * **       ****
```

**Fig. 2.** Comparison of the deduced amino acid sequences of partial *CjF3Hs* and those of corresponding parts of other plant species. Asterisks indicate completely identical amino acid residues among six species. The shadowed amino acid residues (His106, Asp108 and His175) represent absolute conservation important for catalytic activity. Species names and their GenBank accession numbers are as follows; Arabidopsis (*Arabidopsis thaliana*, AF064064); Gentiana (*Gentiana triflora*, AB193311); Ipomoea nil (*Ipomoea nil*, D83041); Oryza sativa (*Oryza sativa*, NM001060692); Vitis (*Vitis vinifera*, EF192467).

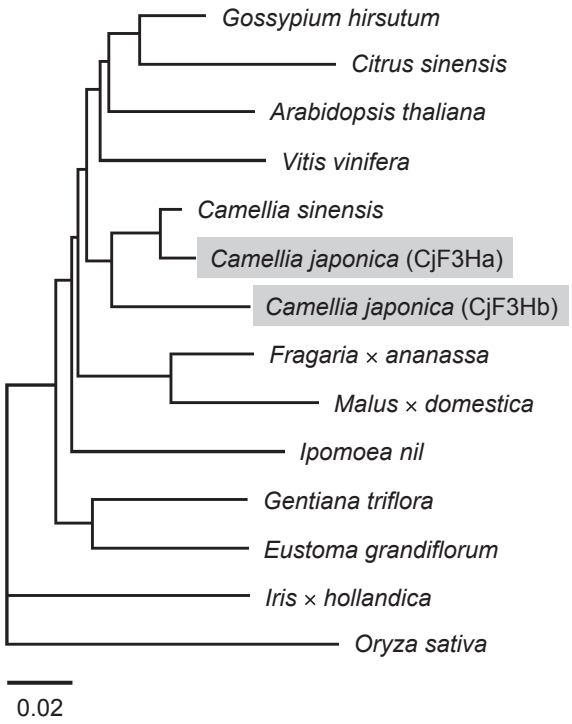


**Fig. 3.** A phylogenetic tree for F3H constructed by neighbor–joining methods using partial deduced amino acid sequences, including CjF3Ha (AB524883), CjF3Hb (AB524884) identified in this study. Sequence data were collected from GenBank databases [*Arabidopsis thaliana* (AF064064), *Camellia sinensis* (AY641730), *Citrus sinensis* (AB011795), *Eustoma grandiflorum* (AB078956), *Fragaria* × *ananassa* (AB201760), *Gentiana triflora* (AB193311), *Gossypium hirsutum* (EF187440), *Ipomoea nil* (D83041), *Iris* × *hollandica* (AB183826), *Malus* × *domestica* (AB074486), *Oryza sativa* (NM001060692), *Vitis vinifera* (EF192467)].

```
        10          20          30          40          50          60
TAATCAAGCC GACAATCAAC GGTGTGTTGA GCATCATAAG GTCATGCACC AAAGCTAAGA

        70          80          90         100         110         120
CAGTGAAGAG GCTGGTGTTC ACATCCTCTG CTGGAACTGT TAATGTCCAG GAACACCAAC

       130         140         150         160         170         180
AACCCATTTT CGACGAGAAC AATTGGAGTG ACTTGGATTT CATCAATAAG AAGAAGATGA
    CjDFR F
       190         200         210         220         230         240
CTGGCTGGAT GTATTTTGTT TCAAAAACAT TGGCAGAGAA AGCAGCATGG GAAGCAGCAA

       250         260         270         280         290         300
AAGAGAACAA CATTGATTTC ATTAGTATCA TTCCTACATT AGTTGTAGGA CCTTTCATCA

       310         320         330         340         350         360
TGCCAACATT CCCACCAAGC CTAATCACCG CTCTCTCCCC CATCACTAGG AATGAAGGAC

       370         380         390         400         410         420
ACTATTCGAT CATAAAGCAA GGGCAGTTTG TGCACCTTGA TGATCTCTGT GAATCTCATA

       430         440         450         460         470         480
TATTCTTGTA TGAGCATCCT CAGGCTGAGG GCAGATACAT TTGCTCCTCC CATGATGCTA

       490         500         510         520         530         540
CCATCCATGA TTTGGCCAAA CTGATGAGAG AGAAATGGCC CGAGTACAAT GTCCCCACTG
                                                 CjDFR R
       550         560         570         580         590         600
AGTTTAAGGG GATAGACAAG GACTTGCCAG TTGTGTCGTT CTCATCGAAG AAGTTGATAG

       610
GAATG.....
```

**Fig. 4.** Nucleotide sequence of partial cDNA of *CjDFR*. The shadowed boxes represent the primer regions of *CjDFR* F and R used in this study.

```
                          10         20         30         40         50
CjDFR                1 IKPTINGVLS IIRSCTKAKT VKRLVFTSSA GTVNVQEHQQ PIFDENNWSD      50
Vitis                1 IKPTIEGMLG IMKSCAAAKT VRRLVFTSSA GTVNIQEHQL PVYDESCWSD      50
Malus                1 IKPTINGLLD ILKACQKAKT VRKLVFTSSA GTVNVEEHQK PVYDESNWSD      50
Rosa                 1 IKPTINGVLD IMQACLKAKT VRRLVFTSSA GSVNVEETQK PVYNESNWSD      50
Gentiana             1 IKPTIDGFLS IIRSCVKAKT VKKLVFTSSA GTVDVQEQQK PVYDENDWSD      50
Ipomoea nil          1 IKPAINGVLN IINSCVKAKT VKRLVFTSSA GTLNVQPQQK PVYDETCWSD      50
Arabidopsis          1 IKPTVNGMLG IMKACVKAKT VRRFVFTSSA GTVNVEEHQK NVYDENDWSD      50
Oryza sativa         1 VKPTVEGMLS IMRACRDAGT VKRIVFTSSA GTVNIEERQR PSYDHDDWSD      50
                         **   * *  *   *  * * *   ****** *       *         ***
                          60         70         80         90        100
CjDFR               51 LDFINKKKMT GWMYFVSKTL AEKAAWEAAK ENNIDFISII PTLVVGPFIM     100
Vitis               51 MEFCRAKKMT AWMYFVSKTL AEQAAWKYAK ENNIDFITII PTLVVGPFIM     100
Malus               51 VEFCRSVKMT GWMYFVSKTL AEQAAWKYAK ENNIDFITII PTLVIGPFLM     100
Rosa                51 VEFCRRVKMT GWMYFASKTL AEQEAWKFAK ENNIDFITII PTLVIGPFLM     100
Gentiana            51 LDFINSTKMT GWMYFVSKIL AEKAAWEVTK ANDIGFISII PTLVVGPFIT     100
Ipomoea nil         51 LDFIYAKKMT GWMYFASKIL AEKEAWKVTK EKKIDFISII PPLVVGPFIT     100
Arabidopsis         51 LEFIMSKKMT GWMYFVSKSL AEKAAWDFAE EKGLDFISII PTLVVGPFIT     100
Oryza sativa        51 IDFCRRVKMT GWMYFVSKSL AEKAAMEYAR EHGLDLISVI PTLVVGPFIS     100
                        *    ***  **** ** * **  *            *  * * ** ***
                         110        120        130        140        150
CjDFR              101 PTFPPSLITA LSPITRNEGH YSIIKQGQFV HLDDLCESHI FLYEHPQAEG     150
Vitis              101 SSMPPSLITA LSPITGNEAH YSIIRQGQFV HLDDLCNAHI YLFENPKAEG     150
Malus              101 PSMPPSLITG LSPILRNESH YGIIKQGQYV HLDDLCLSHI YLYEHPKAEG     150
Rosa               101 PSMPPSLITG LSPLTGNESH YSIIKQGQFI HLDDLCQSHI YLYEHPKAEG     150
Gentiana           101 TTFPPSLITA LSLITGNEAH YGIIKQGQFV HLDDLCEAHI FLYEHPEAEG     150
Ipomoea nil        101 PTFPPSLITA LSLITGNQAH YSIIKQGQYV HLDDLCEAHI FLYEHPKAEG     150
Arabidopsis        101 TSMPPSLITA LSPITRNEAH YSIIRQGQYV HLDDLCNAHI FLYEQAAAKG     150
Oryza sativa       101 NGMPPSHVTA LALLTGNEAH YSILKQVQFV HLDDLCDAEI FLFESPEARG     150
                          ***  *  *     *  * * *  * *   ******   *  * *   * *
                         160        170        180        190        200
CjDFR              151 RYICSSHDAT IHDLAKLMRE KWPEYNVPTE FKGID-KDLP VVSFSSKKLI     200
Vitis              151 RYICSSHDCI ILDLAKMLRE KYPEYNIPTE FKGVD-ENLK SVCFSSKKLT     200
Malus              151 RYICSSHDAT IHELVKMLRE KYPEYNIPTK FKGID-DNLE PVHFSSKKLR     200
Rosa               151 RYICSSHDAT IHEIAKLLKG KYPEYNVPTT FKGIE-ENLP KVHFSSKKLL     200
Gentiana           151 RYICSSHDTT IHDLAKMIRQ NWPEYYIPTK LKGID-EDIP VVSFSSNKLI     200
Ipomoea nil        151 RFICSSHHTT IHGLADMITQ NWPEYYIPSE FKGIE-KDLP VVYFSSKKLQ     200
Arabidopsis        151 RYICSSHDAT ILTISKFLRP KYPEYNVPST FEGVD-ENLK SIEFSSKKLT     200
Oryza sativa       151 RYVCSSHDAT IHGLATMLAD MFPEYDVPRS FPGIDADHLQ PVHFSSWKLL     200
                        *  ****    *            ***  *     *           *** **
                         210        220        230        240        250
CjDFR              201 GM........ .......... .......... .......... ..........     250
Vitis              201 DL........ .......... .......... .......... ..........     250
Malus              201 EI........ .......... .......... .......... ..........     250
Rosa               201 ET........ .......... .......... .......... ..........     250
Gentiana           201 DL........ .......... .......... .......... ..........     250
Ipomoea nil        201 DM........ .......... .......... .......... ..........     250
Arabidopsis        201 EM........ .......... .......... .......... ..........     250
Oryza sativa       201 AH........ .......... .......... .......... ..........     250
```

**Fig. 5.** Comparison of the deduced amino acid sequence of partial *CjDFR* and those of corresponding parts of other plant species. Asterisks indicate completely identical amino acid residues among eight species. The shadowed amino acid residues (Ser29 and Tyr64) represent absolute conservation important for catalytic activity. Species names and their GenBank accession numbers are as follows; Arabidopsis (*Arabidopsis thaliana*, AB033294); Gentiana (*Gentiana triflora*, D85185); Ipomoea nil (*Ipomoea nil*, AB006792); Malus (*Malus* × *domestica*, AY227728); Oryza sativa (*Oryza sativa*, AB003496); Rosa (*Rosa hybrida*, D85102); Vitis (*Vitis vinifera*, X75964).

(GenBank accession no. AB524884), both encoded 210 amino acid residues (Figs. 1 and 2). Phylogenetic tree was constructed to compare the homology of deduced amino acid sequences of CjF3Ha and CjF3Hb with other related F3Hs (Fig. 3). CjF3Ha and CjF3Hb shared 98.0% and 91.2% identities with F3H from *Camellia sinensis*, respectively.

The length of partial *DFR* clone of *C. japonica* was 605 bp, in which 201 amino acid residues were encoded (Figs. 4 and 5) (GenBank accession no. AB524885). Figure 6 shows the comparison of deduced amino acid sequences of partial DFRs. The deduced amino acid sequence of CjDFR showed high identity with that of *Camellia sinensis* (99.0%).

Partial *CjANS* length we obtained was 468 bp and it encoded 155 amino acid residues (Figs. 7 and 8) (GenBank accession no. AB524886). The deduced amino acid sequence of CjANS clustered with those of woody plants, for instance, *Rhododendron* × *pulchrum* (90.3%) (Fig. 9).

**Gene–specific RT–PCR**

We designed gene–specific primer pairs for each gene (*CjF3Ha*, *CjDFR* and *CjANS*) based on the cDNA sequences we obtained in this study (Figs. 1, 4, 7 and Table 1). Although two sequences in *F3H*, *CjF3Ha* and *CjF3Hb*, were identified, *CjF3Ha* was applied for the design of *CjF3H* gene–specific primer, because of its frequent appearance in cloning procedure. Bands of expected sizes were amplified successfully in all samples, wild type *C. japonica* and 'Tamanoura' (red parts of the petals) and 'Hatsu–arashi' with RT–PCR using these

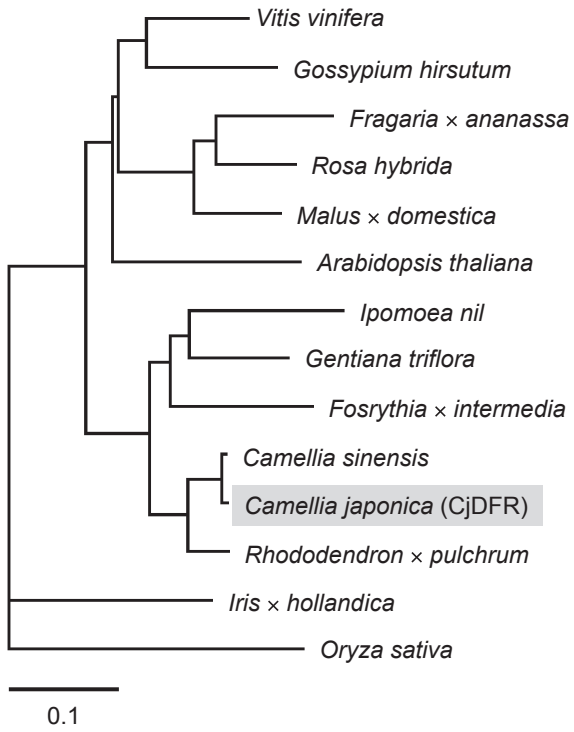


**Fig. 6.** A phylogenetic tree for DFR constructed by neighbor–joining methods using partial deduced amino acid sequences, including CjDFR (AB524885) identified in this study. Sequence data were collected from GenBank databases [*Arabidopsis thaliana* (AB033294), *Camellia sinensis* (AB018685), *Fosrythia* × *intermedia* (Y09127), *Fragaria* × *ananassa* (AY695812), *Gentiana triflora* (D85185), *Gossypium hirsutum* (EF187441), *Ipomoea nil* (AB006792), *Iris* × *hollandica* (AB304917), *Malus* × *domestica* (AY227728), *Oryza sativa* (AB003496), *Rhododendron* × *pulchrum* (AB289595), *Rosa hybrida* (D85102), *Vitis vinifera* (X75964)].

```
        10          20          30          40          50          60
ACCTTGTCTT CCCTGAAGAC AAGCGTGACA TGTCCATTTG GCCTAAGACA CCATCCGACT

        70          80          90         100         110         120
ATATTCCGGC AACAAGCGAG TACGCAAAGC AACTACGAGG TCTAGCAACA AAAGTCCTGT
                                 CjANS F
       130         140         150         160         170         180
CGGCCCTCTC ACTCGGCTTG GGACTAGAAG AAGGCCGACT AGAAAAAGAA GTAGGAGGCA

       190         200         210         220         230         240
TGGAAGAGCT GCATCTCCAA ATGAAAATAA ACTATTACCC AAAATGCCCT CAGCCAGAGC

       250         260         270         280         290         300
TCGCCCTCGG CGTCGAAGCC CACACCGACG TCTCTGCCCT CACCTTCATC CTCCACAACA

       310         320         330         340         350         360
TGGTTCCCGG CCTGCAACTC TTCTACGAGG GCAAATGGAT CACTGCCAAA TGCGTCCCCA

       370         380         390         400         410         420
ACTCCATTAT CATGCACATT GGTGACACGG TAGAAATTCT CAGTAACCGC AAGTACAAGA
                                 CjANS R
       430         440         450         460         470
GCATTCTCCA TCGTGGACTC GTTAATAAGG AAAAAGTGAG GATTTCGT..
```

**Fig. 7.** Nucleotide sequence of partial cDNA of *CjANS*. The shadowed boxes represent the primer regions of *CjANS* F and R used in this study.

primer pairs (*CjF3Ha*, 357bp; *CjDFR*, 412bp; *CjANS*, 312bp) (Fig. 10).

## DISCUSSION

We identified the partial cDNA sequences of *CjF3H*, *CjDFR* and *CjANS*. Their deduced amino acid sequences showed high identities with those of woody plants (Figs. 3, 6 and 9) and some amino acid residues important for their catalytic activities were conserved in the sequences obtained in this study (Figs. 2, 5 and 8). These facts support that the genes we identified were the homologues of *C. japonica*.

Gene–specific primer pairs were constructed to amplify the parts of three anthocyanin biosynthetic genes, *CjF3Ha*, *CjDFR* and *CjANS* (Figs. 1, 4, 7 and Table 1).

```
                        10         20         30         40         50
CjANS             1 LVFPEDKRDM SIWPKTPSDY IPATSEYAKQ LRGLATKVLS ALSLGLGLEE      50
Vitis             1 LIFPEDKRDM TIWPKTPSDY VPATCEYSVK LRSLATKILS VLSLGLGLEE      50
Malus             1 CVYPEDKRDL SIWPQTPADY IEATAEYAKQ LRELATKVLK VLSLGLGLDE      50
Rosa              1 CVYPEDKRDL SIWPQTPSDY IVATSEYAKE LRGLATKILT VLSLGLGLEE      50
Gentiana          1 CIYPERKRDM SIWPKTPHDY IPATIEYAKQ LRDLATKVLA VLSVGLGLEP      50
Ipomoea nil       1 CIFPEDKTDL SIWPKTPSDY IDATKEYAKQ LRALATKVLA VLSLGLGLEE      50
Arabidopsis       1 NYLPSSIRNP SKWPSQPPKI RELIEKYGEE VRKLCERLTE TLSESLGLKP      50
Oryza sativa      1 LVHPDHLADH SLWPANPPEY VPVSRDFGGR VRTLASKLLA ILSLGLGLPE      50
                      *        **  *               * *        **  ***
                        60         70         80         90        100
CjANS            51 GRLEK----- EVGGMEE-LH LQMKINYYPK CPQPELALGV EAHTDVSALT     100
Vitis            51 GRLEK----- EVGGMEE-LL LQKKINYYPK CPQPELALGV EAHTDVSALT     100
Malus            51 GRLEK----- EVGGLEE-LL LQMKINYYPK CPQPELALGV EAHTDVSALT     100
Rosa             51 GRLEK----- EVGGLEE-LV LQMKINYYPK CPQPELALGV EAHTDISALT     100
Gentiana         51 DRLEN----- EVGGMEE-MI LQKKINYYPK CPQPELALGV EAHTDVSALT     100
Ipomoea nil      51 GRLEK----- EVGGMEE-LL LQMKINYYPK CPQPELALGV EAHTDVSALT     100
Arabidopsis      51 NKLMQALGGG DKVGAS---- --LRTNFYPK CPQPQLTLGL SSHSDPGGIT     100
Oryza sativa     51 ETLERRLRGH ELAGVDDDLL LQLKINYYPR CPRPDLAVGV EAHTDVSALS     100
                     *          *          * **  ** * *  *    * *
                       110        120        130        140        150
CjANS           101 FILHN-MVPG LQLFYEGKWI TAKCVPNSII MHIGDTVEIL SNRKYKSILH     150
Vitis           101 FILHN-MVPG LQLFYEGKWV TAKCVPNSII MHIGDTIEIL SNGKYKSILH     150
Malus           101 FILHN-MVPG LQLFYEGKWV TAKCVPNSIV MHIGDTLEIL SNGKYKSILH     150
Rosa            101 FILHN-MVPG LQLFYGGKWV TAKCVPNSIV MHIGDTLEIL SNGKYKSILH     150
Gentiana        101 FILHN-MVPG LQLFYQGKWI TAKCVPDSII MHVGDTLEIL SNGKYKSILH     150
Ipomoea nil     101 FILHN-MVPG LQLFYGGKWV TAKCVPNSII MHVGDTVEIL SNGKYKSILH     150
Arabidopsis     101 ILLPDEKVAG LQVRRGDGWV TIKSVPNALI VNIGDQLQIL SNGIYKSVEH     150
Oryza sativa    101 FILHN-GVPG LQVHHAGSWV TARPEPGTIV VHVGDALEIL TNGRYTSVLH     150
                      *    * * **      *  *    *        **   **  *  * * *
                       160        170        180        190        200
CjANS           151 RGLVNKEKVR IS..... ..... ..... ..... .....     200
Vitis           151 RGLVNKEKVR IS..... ..... ..... ..... .....     200
Malus           151 RGMVNKEKVR IS..... ..... ..... ..... .....     200
Rosa            151 RGLVNKEKVR IS..... ..... ..... ..... .....     200
Gentiana        151 RGLVNKEKVR IS..... ..... ..... ..... .....     200
Ipomoea nil     151 RGVVNREKVR VS..... ..... ..... ..... .....     200
Arabidopsis     151 QVIVNSGMER VS..... ..... ..... ..... .....     200
Oryza sativa    151 RGLVSRDAVR LS..... ..... ..... ..... .....     200
                      *     *  *
```

**Fig. 8.** Comparison of the deduced amino acid sequence of partial *CjANS* and those of corresponding parts of other plant species. Asterisks indicate completely identical amino acid residues among eight species. The shadowed amino acid residues (His93, Asp95 and His150) represent absolute conservation important for catalytic activity. Species names and their GenBank accession numbers are as follows; Arabidopsis (*Arabidopsis thaliana*, AY093302); Gentiana (*Gentiana triflora*, AB193310); Ipomoea nil (*Ipomoea nil*, AB073920); Malus (*Malus* × *domestica*, AB074487); Oryza sativa (*Oryza sativ*a, Y07955); Rosa (*Rosa hybrida*, AB239787); Vitis (*Vitis vinifera*, EF192468).

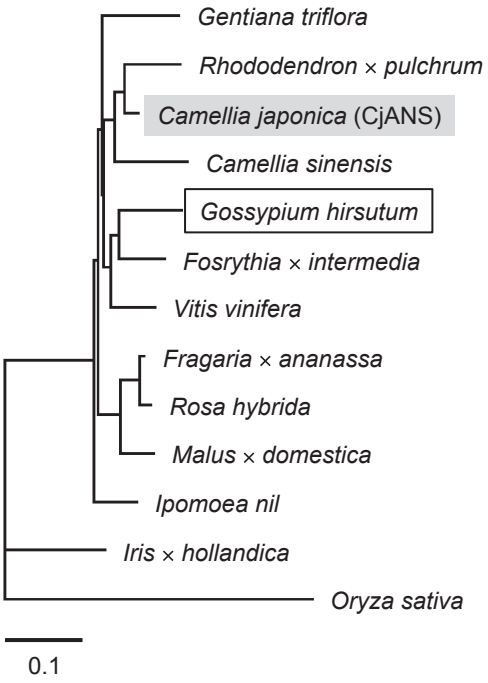


**Fig. 9.** A phylogenetic tree for ANS constructed by neighbor–joining methods using partial deduced amino acid sequences, including CjANS (AB524886) identified in this study. Sequence data were collected from GenBank databases [*Camellia sinensis* (AY830416), *Fosrythia* × *intermedia* (Y12489), *Fragaria* × *ananassa* (AY695817), *Gentiana triflora* (AB193310), *Gossypium hirsutum* (EF187442), *Ipomoea nil* (AB073920), *Iris* × *hollandica* (AB284174), *Malus* × *domestica* (AB074487), *Oryza sativa* (Y07955), *Rhododendron* × *pulchrum* (AB289596), *Rosa hybrida* (AB239787), *Vitis vinifera* (EF192468)].

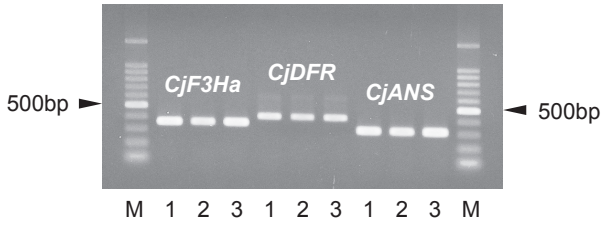


**Fig. 10.** RT–PCR products of *CjF3Ha*, *CjDFR* and *CjANS* using gene–specific primer pairs designed in this study. M indicates 100 bp DNA Ladder Marker. 1, wild type *C. japonica*; 2, 'Tamanoura'; 3, 'Hatsu–arashi'. Three independent flowers were analysed per each accession.

They successfully amplified single PCR products of expected sizes (Fig. 10). It was, therefore, shown that primer pairs designed here are applicable to investigate the expression profiles of these genes on petal pigmentation of *C. japonica*.

Insertion or excision of transposable elements on anthocyanin biosynthetic genes caused flower colour mutation, like fleck and sector in the genus *Ipomoea* (Iida *et al*., 2004). *Camellia japonica* has been domesticated for a long time in Japan (Tuyama, 1968a), and there are many cultivars bearing flowers which imply such involvements of genetic mutations (Tuyama, 1968b). Molecular information we presented here is useful to elucidate the mechanism of such flower colour mutation in *C. japonica*.

*Camellia japonica* 'Hatsu–arashi' lacks an accumulation of anthocyanins and results in generating a white coloured flower (Savige, 1993). However, the expression levels of three genes tested in its petals showed no differences from those of wild type *C. japonica* and 'Tamanoura' (red parts of the petals) (Fig. 10). One possibility is that other anthocyanin biosynthetic genes except *CjF3H*, *CjDFR* and *CjANS* might be down–regulated. An alternative is an involvement of mutations caused by nucleotide substitution or insertion/deletion, if some of the above three genes are responsible. These mutations induce an amino acid substitution or frame shift translation, which are responsible for loss of enzymatic activities, although mRNA transcriptional levels are not suppressed, as in the case of *Matthiola incana* (Hemleben *et al*., 2004).

## ACKNOWLEDGEMENTS

This study was supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists. We also acknowledge a financial support from The Otomo Fund of The International Camellia Society (ICS).

## REFERENCES

Davies, K. M. 2009 Modifying anthocyanin production in flowers. *In* "Anthocyanins", ed. by K. Gould, K. Davies and C. Winefield, Springer, New York, pp. 49–83

Deroles, S. 2009 Anthocyanin biosynthesis in plant cell cultures: a potential source of natural colourants. *In* "Anthocyanins", ed. by K. Gould, K. Davies and C. Winefield, Springer, New York, pp. 107–167

Felsenstein, J. 2004 PHYLIP (Phylogeny Inference Package) version 3.6, Department of Genome Sciences, University of Washington, Seattle, U.S.

Glover, B. J. 2007 *Understanding Flowers and Flowering*. Oxford Univ. Press Inc., New York.

Hemleben, V., A. Dressel, B. Epping, R. Lukacin, S. Martens and M. B. Austin 2004 Characterization and structural features of a chalcone synthase mutation in a white–flowering line of *Matthiola incana* R. Br. (Brassicaceae). *Plant Mol. Biol*., **55**: 455–465

Holton, T. A. and E. C. Cornish 1995 Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell*, **7**: 1071–1083

Iida, S., Y. Morita, J. D. Choi, K. I. Park and A. Hoshino 2004 Genetics and epigenetics in flower pigmentation associated with transposable elements in morning glories. *Adv. Biophys*., **38**: 141–159

Kiefer, E., W. Heller and D. Ernst 2000 A simple and efficient protocol for isolation of functional RNA from plant tissues rich in secondary metabolites. *Plant Mol. Biol. Rep*., **18**: 33–39

Martin, C., A. Prescott, S. Mackay, J. Bartlett and E. Vrijlandt 1991 Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. *Plant J*., **1**: 37–49

Sakata, Y., K. Arisumi and I. Miyajima 1986 Cyanidin 3–galactoside, a new anthocyanin from *Camellia japonica* subsp. *rusticana* (Honda) Kitamura and its occurrence in the garden forms of *Camellia* of Japanese origin. *J. Japan. Soc. Hort. Sci*., **55**: 82–88

Sakata, Y., K. Arisumi and I. Miyajima 1987 Constitution of anthocyanins in flowers of the wild forms of section camellia of Japanese and Formosan origin. *J. Japan. Soc. Hort. Sci*., **56**: 208–214

Savige, T. J. 1993 *The International CAMELLIA REGISTER*. Fine Arts Press Pty. Ltd., Sydney, p. 781

Tuyama, T. 1968a *Camellias of Japan*. Hirokawa Publishing Co. Ltd., Tokyo, pp. 1–2

Tuyama, T. 1968b *Camellias of Japan Color Plates*. Hirokawa Publishing Co. Ltd., Tokyo, pp. 160–200