A bHLH transcription factor, DvIVS, is involved in regulation of anthocyanin synthesis in dahlia (Dahlia variabilis).

Author(s)
Ohno, Sho; Hosokawa, Munetaka; Hoshino, Atsushi; Kitamura, Yoshikuni; Morita, Yasumasa; Park, Kyeung-Ii; Nakashima, Akiko; Deguchi, Ayumi; Tatsuzawa, Fumi; Doi, Motoaki; Iida, Shigeru; Yazawa, Susumu

Citation

Issue Date
2011-10

URL
http://hdl.handle.net/2433/189855

© 2011 The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Type
Journal Article
A bHLH transcription factor, *DvIVS*, is involved in regulation of anthocyanin synthesis in dahlia (*Dahlia variabilis*)

Sho Ohno¹, Munetaka Hosokawa¹,*, Atsushi Hoshino², Yoshikuni Kitamura¹, Yasumasa Morita²†, Kyeung-Il Park²‡, Akiko Nakashima¹, Ayumi Deguchi¹, Fumi Tatsuzawa³, Motoaki Doi¹, Shigeru Iida²§ and Susumu Yazawa¹

¹ Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
² National Institute for Basic Biology, Okazaki 444-8585, Japan
³ Faculty of Agriculture, Iwate University, Morioka 020-8550, Japan
† Present address: National Institute of Floricultural Science, Tsukuba, 305-8519 Japan.
‡ Present address: Department of Environmental Horticultural Science, College of Ecology & Environmental Science, Kyungpook National University, Sangju 742-711, Korea.
§ Present address: Graduate School of Nutritional and Environmental Science and Graduate School of Pharmaceutical Sciences, University of Shizuoka, Shizuoka, 422-8526, Japan.
* To whom correspondence should be addressed. E-mail: mune@kais.kyoto-u.ac.jp

Received 20 March 2011; Revised 6 June 2011; Accepted 8 June 2011

Abstract

Dahlias (*Dahlia variabilis*) exhibit a wide range of flower colours because of accumulation of anthocyanin and other flavonoids in their ray florets. Two lateral mutants were used that spontaneously occurred in ‘Michael J’ (MJW) which has yellow ray florets with orange variegation. MJOr, a bud mutant producing completely orange ray florets, accumulates anthocyanins, flavones, and butein, and MJY, another mutant producing completely yellow ray florets, accumulates flavones and butein. Reverse transcription–PCR analysis showed that expression of *chalcone synthase* (*DvCHS1*), *flavanone 3-hydroxylase* (*DvF3H*), *dihydroflavonol 4-reductase* (*DvDFR*), *anthocyanidin synthase* (*DvANS*), and *DvIVS* encoding a basic helix–loop–helix transcription factor were suppressed, whereas that of *chalcone isomerase* (*DvCHI*) and *DvCHS2*, another *CHS* with 69% nucleotide identity with *DvCHS1*, was not suppressed in the yellow ray florets of MJY. A 5.4 kb CACTA superfamily transposable element, *transposable element of Dahlia variabilis* 1 (*Tdv1*), was found in the fourth intron of the *DvIVS* gene of MJW and MJY, and footprints of *Tdv1* were detected in the variegated flowers of MJW. It is shown that only one type of *DvIVS* gene was expressed in MJOr, whereas these plants are likely to have three types of the *DvIVS* gene. On the basis of these results, the mechanism regulating the formation of orange and yellow ray florets in dahlia is discussed.

Key words: Anthocyanin, bHLH, CACTA superfamily, dahlia, transcription factor.

Introduction

Dahlias (*Dahlia variabilis*) are popular ornamental plants cultivated in many countries. They belong to the Asteraceae and are autoallooctoploids with chromosome number 2n=64 (Lawrence, 1929, 1931b; Lawrence and Scott-Honcrieff, 1935; Gatt et al., 1998). Because of high polyploidy, dahlias exhibit various inflorescence shapes, colours, and sizes, and ~50 000 cultivars have been identified and named in the past century (McClaren, 2009). In particular, dahlias exhibit a wide range of ray floret colours, such as ivory, yellow, pink, red, purple, and black. The pigments accumulated in ray florets are flavonoids, mainly anthocyanins, butein, and flavones and their derivatives that produce red, yellow, and ivory colours, respectively (Bate-Smith and Swain, 1953; Nordström and Swain, 1953, 1956, 1958; Bate-Smith et al., 1955; Harborne et al., 1990; Yamaguchi et al., 1999). In previous studies, four elements were believed to explain the
inheritance of ray floret colours in dahlia: A (pale anthocyanin), B (deep anthocyanin), I (flavone), and Y (yellow) (Lawrence, 1931a; Lawrence and Scott-Honcrieff, 1935; Bate-Smith et al., 1955; Broertjes and Ballegoo, 1967; Singh et al., 1970). Despite its importance, few studies use molecular analysis techniques such as gene isolation and enzyme analysis (Fischer et al., 1988; Wimmer et al., 1998; Yamaguchi et al., 1999; Ogata et al., 2001; Suzuki et al., 2002). Flavonoid synthesis pathway genes have not yet been isolated, except for flavonoid 3'-hydroxylase (Schlangen et al., 2009, 2010).

The flavonoid synthesis pathway is the most well-studied secondary metabolite synthesis pathway in plants. Since flavonoids determine flower colours, regulation of the flavonoid synthesis pathway is very important in ornamental plants. In particular, anthocyanins are found in the flowers of many species and have many functions, such as recruiting pollinators and seed dispersers and scavenging active oxygen species (Yamasaki et al., 1996; Winkel-Shirley, 2001). Anthocyanidin, the aglycone of anthocyanin, is formed in the anthocyanin synthesis pathway by the condensation of three molecules of malonyl-CoA with one molecule of 4-coumaryl-CoA (Fig. 1); the enzymes involved in this pathway are as follows: chalcone synthase (CHS), chalcone isomerase (CHI), flavonoid 3-hydroxylase (F3H), dihydroflavonol 4-reductase (DFR), and anthocyanidin synthase (ANS). The synthesized anthocyanidins are glycosylated by glucosyltransferase and modified through processes such as acylation and malonylation (Tanaka et al., 2008). The genes encoding the above-mentioned enzymes involved in the anthocyanin synthesis pathway are regulated by transcription factors such as basic helix–loop–helix (bHLH), R2R3-MYB, and WD40 repeats (WDRs); these three proteins function as transcription factors by forming complexes or acting alone (Koes et al., 2005). In D. variabilis, first, three molecules of malonyl-CoA and one molecule of 4-coumaryl-CoA, the same substrate as that in chalcone synthesis, are converted to trihydroxylchalcone by CHS and chalcone reductase (Bomati et al., 2005) and then hydroxylated by chalcone 3-hydroxylase or a particular copy of flavonoid 3'-hydroxylase to form butein. Recently, the gene for chalcone 3-hydroxylase, which converts trihydroxylchalcone to butein, was isolated from Cosmos sulphureus (Schlangen et al., 2010b), but whether the anthocyanin and butein synthesis pathways are regulated by the same regulator gene is unknown.

To isolate the genes that impart colour to flowers, mutants formed by transposable elements have been widely used. This method of gene isolation, called transposon tagging, was used to isolate Antirrhinum majus Pallida (DFR) (Martin et al., 1985), Zea mays bronze (UDP-glucose flavonoid 3-O-glucosyltransferase) (Fedoroff et al., 1984), Petunia hybrida phi (An1) (Chuck et al., 1993), and many more. Therefore, to clarify whether anthocyanin and butein synthesis in dahlia are regulated by the same or different mechanisms, two lateral mutants MJOr and MJY that spontaneously occurred in ‘Michael J’ (MJW), which has yellow ray florets with orange variegation, were used. MJOr, a bud mutant producing orange ray florets, accumulates anthocyanins, flavones, and butein, and MJY, another mutant producing yellow ray florets, accumulates flavones and butein. In this study, anthocyanin synthesis pathway genes and transcription factors were isolated and it was shown that DvIVS, belonging to the An1 subgroup bHLH transcription factor family, regulates anthocyanin synthesis in ray florets of dahlia. Finally, a putative model to explain the mechanism regulating expression of orange and yellow flower colours in dahlia is proposed.

**Materials and methods**

**Plant material and developmental stages**

*Dahlia variabilis* cv. ‘Michael J’ (MJW) was obtained from Yukihiro Fukuda (Haibara, Nara, Japan). MJW has bright yellow ray florets with a small number of brilliant orange markings (Fig. 2A). It rarely produces completely orange (Fig. 2B) or completely yellow (Fig. 2C) ray florets as a bud mutation. The completely orange mutant line (MJOr) and the completely yellow ray floret mutant line (MJY) were isolated and obtained from Yukihiro Fukuda. The two lines and wild-type plants were grown under standard greenhouse conditions or in the experimental field of Kyoto University (Kyoto, Japan). To analyse temporal gene expression, the MJOr ray florets were classified into five stages on the basis of the degree of colouration (Fig. 3A): stage 1, an uncoloured ray floret; stage 2, the tip of a ray floret is coloured; stage 3, a ray floret coloured in the centre; stage 4, colouring is completed and the ray floret starts to unfold; and stage 5, a completely unfolded ray floret. The size and stage of MJY ray florets used were almost same.

**HPLC analysis**

Ray florets were soaked overnight in 5 ml of MAW solution (methanol:acetic acid:water, 4:1:5 v/v/v) to extract the pigments.
For pigment hydrolysis, extracted solutions were evaporated, redissolved in 2 ml of 20% hydrochloric acid, boiled, and used as crude aglycones. HPLC analysis was performed using an LC10A system (Shimadzu, Kyoto, Japan) with a C18 column (Nihon Waters K.K., Tokyo, Japan) maintained at 40°C and a photo-diode array detector. The detection wavelength was 350 nm for flavones, 380 nm for chalcones, and 530 nm for anthocyanins. Eluant A was 1.5% phosphate dissolved in water and eluant B was 1.5% phosphate, 20% acetic acid, and 25% MeCN dissolved in water. Analysis was performed at a flow rate of 1 ml min⁻¹ and column temperature of 40°C, using a mobile phase gradient starting at 20% B to 85% B over 40 min with 5 min re-equilibration at 20% B. As standards for the determination of flavonoids, commercially available naringenin, apigenin, and luteolin (Wako Pure Chemical Industries Ltd, Osaka, Japan) as well as HPLC-separated and HPLC-purified hydrolysed cyanidin and pelargonidin from rose petals were used. To obtain hydrolysed standards of butein and isoliquiritigenin, extracts from orange ray florets of a seedling line, HywR7R, were separated by paper chromatography, and each band was eluted with methanol. Each eluate was dried and redissolved in a small amount of methanol. The colour, Rf value, and maximum wavelength of eluted compounds were measured, and the compounds were determined by comparing the data with those of authentic butein (kindly supplied by Dr Norio Saito) and previously reported data (Nordström and Swain, 1956; Saito et al., 1970).

Isolation of nucleic acids and sequencing
To clone the DvIVS genomic region and Tdv1 region, genomic DNA of ray florets was isolated using a modified cetyltrimethyl-ammonium bromide method (Murray and Thompson, 1980) and
purified with MagExtractor™-Plant Genome (Toyobo Co. Ltd, Osaka, Japan). Genomic DNA was also isolated using Qiagen Genomic-tip 100/G (Qiagen, Valencia, CA, USA) for genomic PCR analyses. Total RNA was isolated using the isothiocyanate/cesium chloride centrifugation method and an RNeasy Plant Mini Kit (Qiagen), a modified QuickGene RNA Culture Cell Kit S (FujiFilm, Tokyo, Japan), or a Get PURE RNA Kit (Dojindo, Kumamoto, Japan). All sequence analyses were performed using a BigDye® Terminator v 3.1 Cycle Sequencing Kit and a 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

**Gene isolation**

*DvCHS1*, *DvCHI*, *DvF3H*, *DvDFR*, *DvANS*, *DvIVS*, *DvDEL*, and *DvActin* cDNAs were isolated using degenerate primers designed from conserved regions of each gene shown in Supplementary Table S1 available at JXB online. To isolate *DvMYB1*, *DvMYB2*, *DvWDRI*, and *DvWDR2* cDNAs, dahlia cDNA libraries provided by Dr Yoshikazu Tanaka (Sunotty, Shimomura, Mishima, Osaka, Japan) were screened (Suzuki et al., 2002). *Ipomoea nil* MYB (AB232770) and WDR (AB232779) cDNAs were used as probes, and AlkPhos Direct (GE Healthcare Bio-Sciences Corp., Piscataway, NJ, USA) was used for probe hybridization. To determine the complete sequence of mRNA of all genes, 3′ rapid amplification of cDNA ends (RACE) and 5′ RACE was performed using a GeneRacer™ Kit (Invitrogen, Carlsbad, CA, USA). Primers used are shown in Supplementary Table S2. Each end of the genes was sequenced for at least 16 clones for each gene to avoid single nucleotide polymorphisms (SNPs) because of dahlia’s high ploidy. While performing 5′ RACE of *DvMYB1*, a novel CHS-like fragment was obtained that differed from *DvCHS1*. This new CHS was named *DvCHS2* and each cDNA end was determined by 5′ and 3′ RACE. For 5′ and 3′ RACE of the transcription factor genes, total RNA extracted from *D. variabilis* ‘Matsubirayuhashi’ was used.

**RT-PCR and real-time RT-PCR**

The total RNA of MJOr and MJY was subjected for reverse transcription using an oligo(dT) 20 primer and ReverTra Ace polymerase (Toyobo). RT-PCRs of *DvCHS1*, *DvCHI*, *DvF3H*, and *DvDFR* were performed using KOD Plus polymerase (Toyobo), and PCR products were cloned into pCR®-Blunt II-TOPO® using a Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen). The PCR program was set at 94 °C for 2 min, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 2–3 min.

Total RNA extracted from stage 1–5 ray florets was used. Real-time RT-PCR was performed using SYBR® Premix Ex Taq™ II (Takara Bio Inc., Ohtsu, Japan) according to the manufacturer’s instructions. Real-time PCR was performed using a 7900HT Fast Real-Time PCR System (Applied Biosystems). The primers used are shown in Supplementary Table S4 at JXB online. The PCR program was set at 50 °C for 1 s, 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s, 60 °C for 15 s, 72 °C for 20 s, and the subsequent dissociation steps. Three different colors were used as replicates for each stage, and *DvActin* was used as an internal standard. Because of the low transcription level, samples that did not reach the threshold were calculated as 40 cycles.

To test partial expression of *DvIVS*, three primer sets were used for RT-PCR: IVS-Full-F and IVS-Full-R, 35S-IVS-F and IVS-611R, and IVS-725F and IVS-Full-R (Supplementary Table S5 at JXB online). The first primer set amplifies from the start to the stop codon. The second primer set amplifies from 30 bp after the start codon to the third exon, and the third primer set amplifies from the fifth exon to the stop codon.

**Denaturing gradient gel electrophoresis (DGGE) analysis**

RT-PCR was performed with KOD Dash polymerase (Toyobo), using the cDNA of MJOr and MJY. The PCR program was set at 94 °C for 2 min, followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 2 min. A total of 40 cloned plasmids of both lines with *DvCHS1*, *DvCHI*, *DvF3H*, and *DvDFR* were extracted, except for MJY *DvCHS1* (only 20 plasmids were extracted). To determine the accurate sequence, plasmids with *DvCHS1*, *DvF3H*, and *DvDFR* of both lines were sequenced twice from both the 3′ and 5′ ends. The primers used are shown in Supplementary Table S6 at JXB online. To exclude misreading of sequences or misamplification in PCR, SNPs that were only observed in one clone were omitted.

To analyse the sequence of the genomic region of *DvCHS1* in MJOr, MJY, and MJW, genomic PCR was performed using the primer set shown in Supplementary Table S3 at JXB online. A total of 30 clones of the *DvCHS1* gene were sequenced for each line.

**GenBank accession numbers**

Sequence data of these genes have been deposited in the DDBJ/Genbank data libraries under accession numbers AB576129, AB576130, AB576131, AB576132 (*DvCHS1*), AB576133 (*DvCH2*), AB591827 (*DvCHI*), AB591828 (*DvF3H*), AB591829 (*DvDFR*), AB591830 (*DvANS*), AB601003 (*DvMYB1*), AB601004 (*DvMYB2*), AB601005 (*DvIVS*), AB601006 (*DvDEL*), AB601007
(DvWDR1), AB601008 (DvWDR2), AB601010 (DvIVS genomic region), and AB601009 (Tdv1).

## Results

### Pigment analysis

Wild-type MJW produces yellow ray florets with orange variegation, while MJOr and MJY produce completely orange and completely yellow ray florets, respectively, without any variegation. To clarify the difference in pigments between MJOr and MJY, aglycones were extracted from ray florets and analysed by HPLC. In both lines, peaks of chalcones (butein and isoliquiritigenin) and flavones (apigenin and luteolin) were detected. In addition to these, peaks of anthocyanidins (pelargonidin and cyanidin) were detected. In MJY, peaks of chalcones (butein and isoliquiritigenin) and flavones (apigenin and luteolin) were detected. In MJOr, the relative expression level of DvIVS and DvCHS1 compared with that of DvActin increased until stage 3 or 4 and then decreased, but consistently very low expression levels were maintained in MJY (Fig. 3C). In contrast, expression levels of DvCHS2 were slightly lower in MJY, but the expression pattern of DvCHS2 in MJOr and MJY was the same (Fig. 3C). This result supported the results of RT-PCR analyses and suggested that DvCHS1, but not DvCHS2, is regulated by DvIVS.

DvCHS1 (Supplementary Table S7 at JXB online), DvCHI (Supplementary Table S8), DvF3H (Supplementary Table S9), and DvDFR (Supplementary Table S10) cDNA sequences isolated from MJOr and MJY ray florets were determined and the number of different sequences expressed in both lines was determined on the basis of SNPs. Based on sequence analysis of transcripts, 17 and 15 sequences of DvCHI (Supplementary Table S8), nine and 12 sequences of DvF3H (Supplementary Table S9), and nine and eight sequences of DvDFR (Supplementary Table S10) were observed in MJOr and MJY, respectively. Because drastic differences were not observed in the number of sequences expressed in both lines, whole copies were assumed to be suppressed in MJY. However, 12 different sequences of DvCHS1 were detected in MJOr, while only one sequence was detected in MJY (Supplementary Table S7). At least three major transcripts (DvCHS1 type 1, 2, and 4) were detected in MJOr, while only one transcript (DvCHS1 type 4) was detected in MJY (Supplementary Table S7). The genomic fragments of DvCHS1 were isolated and it was observed that not only MJOr and MJW, but also MJY has DvCHS1 genes that have coding sequences identical to these three major transcripts (Supplementary Tables S11, S12).

### Temporal expression of anthocyanin synthesis pathway genes

To analyse genes associated with anthocyanin synthesis, partial sequences of DvCHS1, DvCHI, DvF3H, DvDFR, DvANS, DvIVS, and DvDEL were isolated using degenerate primers. DvMYB1, DvMYB2, DvWDR1, and DvWDR2 were isolated by cDNA library screening using cDNA fragments of InMYB1 and InWDR1 as probes (Morita et al., 2006). Full cDNA sequences of all genes including DvCHS2, another CHS gene, were isolated by 5’ and 3’ RACE. Partial sequences of DvActin were also isolated with degenerate primers by 3’ RACE.

To examine the temporal expression pattern of the anthocyanin synthesis pathway genes during ray floret development, the MJOr ray florets were divided into five developmental stages based on the degree of colouration and RT-PCR was performed. Although expression levels of DvCHS2 and DvCHI showed no difference, expression levels of DvCHS1, DvF3H, DvDFR, and DvANS were lower in MJY than in MJOr at all developmental stages (Fig. 3B). Down-regulation of multiple genes encoding the anthocyanin synthesis pathway enzymes suggested that transcription factor(s) do not function in MJY.

Temporal expression analysis by RT-PCR was also performed for transcription factors. Although expression levels of DvMYB1, DvMYB2, DvDEL, DvWDR1, and DvWDR2 did not differ significantly, the expression level of DvIVS was lower in MJY than in MJOr, which is similar to that of DvCHS1, DvF3H, DvDFR, and DvANS (Fig. 3B). This result suggested that DvIVS is a transcription factor that activates anthocyanin biosynthesis genes including DvCHS1, DvF3H, DvDFR, and DvANS.

To confirm these results, real-time RT-PCR of DvCHS1 and DvCHS2, which seem to be regulated and not regulated by DvIVS, respectively, was performed using total RNA extracted from stage 1–5 ray florets. In MJOr, the relative expression level of DvIVS and DvCHS1 compared with that of DvActin increased until stage 3 or 4 and then decreased, but consistently very low expression levels were maintained in MJY (Fig. 3C). In contrast, expression levels of DvCHS2 were slightly lower in MJY, but the expression pattern of DvCHS2 in MJOr and MJY was the same (Fig. 3C). This result supported the results of RT-PCR analyses and suggested that DvCHS1, but not DvCHS2, is regulated by DvIVS.

### Characterization of DvIVS

The full-length DvIVS cDNA was 2483 bp containing a 649 amino acid open reading frame (ORF). The DvIVS ORF retained the bHLH domain at 472–530 amino acids (Fig. 4A). BLASTP search (http://blast.ncbi.nlm.nih.gov/)

### Table 1. Aglycones in the petals of MJOr and MJY

<table>
<thead>
<tr>
<th>Line</th>
<th>Anthocyanidin</th>
<th>Flavone</th>
<th>Chalcone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pelargonidin</td>
<td>Cyanidin</td>
<td>Apigenin</td>
</tr>
<tr>
<td>Orange</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Yellow</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>

+, Abundant; –, not detected.
(A) Phylogenetic analysis of DvIVS. (A) Amino acid comparison of the bHLH domain of DvIVS, InIVS, PhAN1, and AtTT8. Numbers (*) indicate amino acids that are fully conserved in each of the proteins. (B) Phylogenetic tree for bHLH transcription factors associated with anthocyanin synthesis pathways. The entire amino acid sequences were aligned using ClustalW, and the tree was constructed by the Neighbor–Joining method. Bootstrap values of 1000 retrials are indicated on each branch, and the scale shows 0.1 amino acid substitutions per site. The abbreviations shown in front of each protein indicate the plant species: Dv, Dahlia variabilis; Am, Antirrhinum majus; At, Arabidopsis thaliana; In, Ipomoea nil; Pf, Perilla frutescens; Ph, Petunia hybrida; Zm, Zea mays. The accession number of each protein is: DvIVS (AB601010), InIVS (AB643516), PhAN1 (AAB32636), AtTT8 (NP_176552), AtGLL3 (NP_680372), AtTT8 (CAC14865), IndEL (BAE94393), InIVS (BAE94394), PfF3G1 (BAC56998), PhAN1 (AAG25927), PhJAF13 (AAC39455), and ZmIN1 (AAB03841).

Fig. 4. Phylogenetic analysis of DvIVS. (A) Amino acid comparison of the bHLH domain of DvIVS, InIVS, PhAN1, and AtTT8. (B) Phylogenetic tree for bHLH transcription factors associated with anthocyanin synthesis pathways. Bootstrap values of 1000 retrials are indicated on each branch, and the scale shows 0.1 amino acid substitutions per site. The abbreviations shown in front of each protein indicate the plant species: Dv, Dahlia variabilis; Am, Antirrhinum majus; At, Arabidopsis thaliana; In, Ipomoea nil; Pf, Perilla frutescens; Ph, Petunia hybrida; Zm, Zea mays. The accession number of each protein is: DvIVS (AB601010), InIVS (AB643516), PhAN1 (AAB32636), AtTT8 (NP_176552), AtGLL3 (NP_680372), AtTT8 (CAC14865), IndEL (BAE94393), InIVS (BAE94394), PfF3G1 (BAC56998), PhAN1 (AAG25927), PhJAF13 (AAC39455), and ZmIN1 (AAB03841).

Analysis of the DvIVS genomic region

The DvIVS genomic region coding for full-length mRNA (AB601010) of MJOr was 8477 bp and contained nine exons and eight introns (Fig. 5A). Genomic PCR was performed to examine whether a transposable element was inserted in the DvIVS gene of MJY. An ~8.2 kb amplified product was obtained from MJOr, and an ~13.6 kb amplified product was obtained from MJY (data not shown). The 5385 bp insertion sequence (AB601009) was found in the fourth intron, 903 bp downstream of the fourth exon of the DvIVS gene of MJY. To characterize the structure of the DvIVS gene of MJW, genomic PCR was performed using the primer set IVS Full-F and IVS-G2869R. Not only the genomic fragment (8.2 kb) that should include the insertion sequence, but also a shorter fragment (2.8 kb) without the insertion sequence were detected in MJW (Fig. 5A).

The inserted element of MJY carried a terminal inverted repeat (TIR) starting with 5′-CACTA-3′ (Table 2), which is characteristic of a CACTA superfamily transposable element. The insertion generated a 3 bp (TGG) target site duplication (TSD), similar to the TSD generated by other CACTA-type transposable elements (Inagaki et al., 1994; Zabala and Vodkin, 2008; Xu et al., 2010). It also contained an ORF for antisense orientation (Fig. 5A) (nucleotides 1735–3555), encoded a 606 amino acid polypeptide, and shared high identity with other CACTA superfamily transposable elements [41% identical to TNP2 in A. majus Tam1 (Nacken et al., 1991), 33% identical to TNP2 in Glycine max Tgm9 (Xu et al., 2010), and 32% identical to Eu-1 in Zea mays (Pereira et al., 1986)]. To test whether the insertion sequence has the ability to excise from the DvIVS gene, 600 bp fragments containing the insertion site in MJW were amplified and cloned into plasmid vectors. A total of 23 clones were obtained with the DvIVS gene fragments with small rearrangements that were presumed to be footprint sequences of the insertion sequence (Table 3, Fig. 5B).

From these results, the insertion sequence was regarded as an active transposable element and named transposable element of D. variabilis 1 (Tdv1). The DvIVS fragments could be obtained from MJOr but not from MJY. In addition to the DvIVS gene fragment, two different gene fragments were obtained. Both of them were shorter than the DvIVS fragments, and showed 97% identity to the DvIVS gene. This suggested that the dahlia lines used here have at least three types of DvIVS genes.

To test whether all three DvIVS genes are expressed in the flowers, DvIVS cDNA fragments from MJOr were amplified, cloned into a plasmid vector, and the cDNA sequence was analysed. Although 20 clones of DvIVS cDNA were sequenced, all of them were derived from the DvIVS gene (data not shown). The result suggested that the anthocyanin pigmentation in the flowers depend on the DvIVS gene and that the two other DvIVS genes are not expressed in the flowers.

To characterize the DvIVS transcripts in MJY, RT-PCR was performed using three primer sets. Successful amplification was achieved only by the primer set amplifying the region upstream of Tdv1 insertion sites in both MJOr and MJY; however, other primer sets did not successfully amplify regions in MJY (Fig. 5C). This suggested that Tdv1 insertion into the fourth intron of DvIVS resulted in truncated DvIVS mRNA in MJY. On analysing DvIVS mRNA expression in MJY by 3′ RACE, truncated DvIVS mRNA containing the fourth exon and a short part of the fourth intron (nucleotides 125–250 downstream of the fourth exon/fourth intron junction) was detected (data not shown). Thus, Tdv1 insertion into the DvIVS gene of MJY resulted in generation of short-sized DvIVS transcripts lacking the bHLH domain, thereby leading to an inability for transcriptional regulation of anthocyanin synthesis genes.
**Discussion**

**DvIVS is a bHLH transcription factor associated with anthocyanin synthesis in dahlia.**

The following transcription regulators of anthocyanin synthesis pathway enzymes are well known: bHLH, R2R3-MYB, and WDR (Koes et al., 2005). To date, many studies have been conducted to determine flavonoid synthesis pathway transcription factors. *Anl* and *JAF13*, *An2*, and *An11* have been isolated as bHLH, R2R3-MYB, and WDR transcription factors, respectively, regulating anthocyanin synthesis in petunia (Quattrocchio et al., 1993, 1998; De Vetten et al., 1997; Spelt et al., 2000, 2002). Similarly, bHLH, R2R3-MYB, and WDR transcription factors control anthocyanin synthesis in the three *Ipomoea* species (Park et al., 2004, 2007; Morita et al., 2006).

bHLH proteins are found in both animals and plants, and constitute a superfamily of transcription factors.
Table 3. Amplified fragment with IVS-G2241F and IVS-G2869R

<table>
<thead>
<tr>
<th>Type</th>
<th>Length (bp)</th>
<th>Identity to the DvIVS genome (%)</th>
<th>MJOr</th>
<th>MJY</th>
<th>MJW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>628</td>
<td>100</td>
<td>10</td>
<td>0</td>
<td>23*</td>
</tr>
<tr>
<td>2</td>
<td>625</td>
<td>97</td>
<td>6</td>
<td>16</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>615</td>
<td>97</td>
<td>1</td>
<td>3</td>
<td>16</td>
</tr>
</tbody>
</table>

* Footprint sequences were not considered in type 1 of MJW.

Arabidopsis has at least 147 bHLH genes, and these are divided into 21 subfamilies (Toledo-Ortiz et al., 2003). The bHLH genes associated with regulation of anthocyanin synthesis were first found in Z. mays (Ludwig et al., 1989), and have subsequently been found in many other species. Recent studies showed that AtTT8 (Nesi et al., 2000), AtGL3 (Payne et al., 2000), and AtEGL3 (Zhang et al., 2003) regulate anthocyanin and proanthocyanidin synthesis in A. thaliana. These bHLH genes regulating anthocyanin synthesis are usually divided into two subgroups on the basis of their phylogenetic tree (Spelt et al., 2000): one subgroup includes AtTT8, Anl, and InIVS, and the other includes AtGL3, AtEGL3, JAF13, and InDEL. Usually, the bHLH domain in the former subgroup is not separated by an intron, while that in the latter group is separated by an intron. For example, the bHLH domains of Anl (Spelt et al., 2000), IpIVS (Park et al., 2007), AtTT8 (Nesi et al., 2000), and GhbHLH1 in Gentiana triflora (Nakatsuka et al., 2008) are encoded within a single exon, while those of JAF13 and ZmInl are encoded within two exons. DvIVS shows high identity with Anl (50%), InIVS (48%), and AtTT8 (48%) in deduced amino acid sequences and is classified into the same subgroup with these bHLHs in the phylogenetic tree (Fig. 4B). DvIVS has a bHLH domain encoded within a single exon (Fig. 5A), suggesting that the structure of the DvIVS gene is similar to that of the Anl gene. These data suggest that DvIVS is an orthologue of AtTT8, Anl, and InIVS, and that DvIVS regulates transcription of DvCHS1, DvF3H, DvDFR, and DvANS, but not of DvCHS2 and DvCHI. Real-time RT-PCR showed that the expression patterns of DvIVS and DvCHSI were highly coordinated, while that of DvCHS2 was not, supporting the suggestion (Fig. 3C). A slight decrease in the expression levels of DvCHS2 was observed in MJY; this may have been caused by the excess production of substrate resulting from an inability to synthesize anthocyanin. In many cases, although the particular set of genes may differ between species and tissues within a species, a subset of the anthocyanin pathway genes is affected by these transcription factors (Gonzalez et al., 2008). Transcription regulation in the three Ipomoea species (Park et al., 2004, 2007; Morita et al., 2006) is similar to that in dahlia.

The cDNA fragment derived from the 5′ end of DvIVS mRNA containing the first to third exons was detected in MJY; however, the cDNA fragment from the 3′ end of DvIVS mRNA was not observed. This truncated mRNA did not contain a bHLH domain in the eighth exon, suggesting that the bHLH domain of DvIVS is important for anthocyanin synthesis in dahlia.

Flower colour variegation is caused by Tdv1

Flower colour variegation is often caused by recurrent excision of transposable elements in a pigment biosynthesis gene. For example, Tpn1, a CACTA superfamily transposable element in DFR in I. nil (Inagaki et al., 1994); Tip100, a hAT superfamily transposable element in CHS-D in I. purpurea (Habu et al., 1998); and Tgm9, a CACTA superfamily transposable element in DFR2 in G. max (Xu et al., 2010) resulted in variegated flower phenotypes. Therefore, an association of transposable elements with floret colour was expected in MJY. It was hypothesized that MJOr was a complete revertant mutant line and MJY was a loss-of-function mutant line, where the transposable element could not transpose from a gene associated with anthocyanin synthesis.

A 5.4 kb insertion sequence, named Tdv1, was found in the fourth intron of the DvIVS gene in MJY. Tdv1 shares a number of CACTA superfamily characteristic features. CACTA superfamily transposable elements are DNA transposons directly transposed from DNA to DNA (Feschotte et al., 2002), and almost all CACTA transposable elements harbour ‘CACTA’ in the outermost region of TIR (Tian, 2006). Tdv1 also has a CACTA superfamily characteristic TIR (CACTACAA) (Table 2) and a 3 bp (TGG) TSD was observed in the MJY genome (Fig. 5B). The 606 amino acid ORF in Tdv1 shared high identity with other CACTA superfamily transposable elements [41% identical to TNP2 in A. majus Tam1 (Nacken et al., 1991), 33% identical to TNP2 in G. max Tgm9 (Xu et al., 2010), and 32% identical to En-1 in Z. mays (Pereira et al., 1986)]. Further analysis of this translated amino acid sequence by NCBI’s CDD (Marchler-Bauer et al., 2011) showed high identity with Transposase_21 (pfa02992). The footprint sequences formed by excision of the transposable element were also found in the MJY genome (Table 3, Fig. 5C), indicating that Tdv1 is an active CACTA superfamily transposable element. Because no PCR fragments with footprint sequences of Tdv1 in MJY could be amplified, transposition of Tdv1 is completely suppressed in MJY.

Anthocyanin and butein synthesis in dahlia

Flavones and butein accumulated in the ray florets of both MJOr and MJY. Thus, to synthesize these compounds, at least one CHS other than DvCHS1 must function in MJY. Indeed, another CHS, DvCHS2, coding for a 389 amino acid protein, was identified whereas DvCHS1 coded for a 389 amino acid protein. The nucleotides from the start to the stop codon and the amino acid sequence in DvCHS1 were 69% and 82% identical, respectively, to those in DvCHS2. A phylogenetic tree was constructed using the deduced amino acid sequences associated with flavonoid synthesis, including DvCHS1 and DvCHS2 (Supplementary Fig. S1 at JXB online). DvCHS1 showed high identity with GhCHS4 (88%) and DvCHS2 showed high identity with GhCHS1 (89%) and GhCHS3 (84%). GhCHS1 and GhCHS3 are associated with anthocyanin and flavonol synthesis in the corolla of gerbera.
and GhCHS4 is expressed in ray florets and carpels and is a typical CHS that synthesizes naringenin chalcone (Laitinen et al., 2008). CHS generally functions as a homodimer (Ferrer et al., 1999); therefore, amino acid residues associated with the formation of active sites are commonly conserved in plants. Cys164, His303, and Asn336, as CoA-binding sites, Val98, Thr132, Ser133, Met137, Gly163, Thr194, Val196, Thr197, Gly211, Gly216, Ile254, Gly256, Ser338, and Pro375 as active site residues, and Phe215 and Phe265 as gatekeepers are conserved in CHS2 of Medicago sativa (Ma et al., 2009). DvCHS1 and DvCHS2 possess all these conserved residues. These observations indicate that DvCHS1 and DvCHS2 encode functional CHS proteins. Therefore, although whether DvCHS1 and DvCHS2 have functional selectivity is still unclear, presumably DvCHS2 but not DvCHS1 contributes to synthesis of flavones or butein in MJY, and regulation of anthocyanin and flavone/butein synthesis is different in dahlia.

Redundant copies regulated by DvIVS

Dahlia is an autoallooctoploid (Lawrence, 1929, 1931b; Lawrence and Scott-Honcrieff, 1935; Gatt et al., 1998). DvCHS1 (Supplementary Table S7 at JXB online), DvCHI (Supplementary Table S8), DvF3H (Supplementary Table S9), and DvDFR (Supplementary Table S10) mRNA sequences expressed in MJOr and MJY ray florets were sequenced and the number of different sequences expressed in both lines was determined on the basis of SNPs. Sequences of DvCHI, DvF3H, and DvDFR cDNA were almost the same in MJOr and MJY. However, a drastic difference was detected in the expression sequence of DvCHS1. At least three different major transcripts were detected in MJOr, but only one transcript was detected in MJY. This difference was confirmed by DGGE analysis (Supplementary Fig. S2 at JXB online). The genomic region of DvCHS1 of MJOr, MJW, and MJY was sequenced (Supplementary Tables S11, S12). Three common sequences were detected among these three lines. Thus, mRNA differences were derived not from genomic differences, but presumably from differences in DvIVS transcriptional regulation. Although the expression level of DvCHS1 was much lower in MJY (Fig. 3C), one type of DvCHS1 mRNA expressed in MJY was also expressed in MJOr, suggesting that DvCHS1 expression is not completely regulated by DvIVS. DvCHS1 can be divided into two groups: the regulation of the first is totally dependent on DvIVS, while the other can also be activated independently of DvIVS. The common sequence belongs to the latter group and the others belong to the former group. This imperfect regulation might be derived from genetic redundancy. Nevertheless, further study is needed in this regard. The manner in which plants with higher polyploidy regulate redundant genes is very interesting.

Finally, a putative model of anthocyanin and butein synthesis in MJOr and MJY was constructed (Fig. 6). In MJOr, DvIVS functions, followed by activation of DvCHS1, DvF3H, DvDFR, and DvANS and synthesis of anthocyanidin, flavone, and butein. In MJY, because of Tdv1 insertion, only partial DvIVS mRNA is expressed and transcriptional regulation cannot function. Although DvCHS1, DvF3H, DvDFR, and DvANS are not activated, DvCHS2 and DvCHI are not affected by DvIVS, resulting in accumulation of flavone and butein.

Fig. 6. The putative model of anthocyanin and butein synthesis in MJOr and MJY. For simplicity, the 3′ hydroxylation pathway is abbreviated. (A) The model for MJOr. DvIVS can function, then DvCHS1, DvF3H, DvDFR, and DvANS are activated, and anthocyanidin, flavone, and butein are synthesized. (B) The model for MJY. Because of Tdv1 insertion, only truncated DvIVS mRNA is expressed and transcriptional regulation cannot function. Although DvCHS1, DvF3H, DvDFR, and DvANS are not activated, DvCHS2 and DvCHI are not affected by DvIVS, resulting in accumulation of flavone and butein.
in accumulation of flavone and butein. Therefore, MJO\textsubscript{r} produces orange ray florets on account of the accumulation of anthocyanin and butein, while MJY produces yellow ray florets on account of the accumulation of butein only.

**Supplementary data**

Supplementary data are available at *JXB* online.

Figure S1. Phylogenetic tree for CHS associated with flavonoid synthesis.

Figure S2. DGGE analysis of *DvCHS1* cDNA from MJO\textsubscript{r} and MJY.

Table S1. Degenerate primers used for isolating the cDNA fragment.

Table S2. Primers used for RACE.

Table S3. Primers used for semi-quantitative RT-PCR.

Table S4. Primers used for real-time RT-PCR.

Table S5. Primers used for analysing the *DvIVS* genomic region and *Tdv1* region.

Table S6. Primers used for confirming the sequence.

Table S7. SNPs of *DvCHS1* cDNA in MJO\textsubscript{r} and MJY.

Table S8. SNPs of *DvCH1* cDNA in MJO\textsubscript{r} and MJY.

Table S9. SNPs of *DvF3H* cDNA in MJO\textsubscript{r} and MJY.

Table S10. SNPs of *DvDFR* cDNA in MJO\textsubscript{r} and MJY.

Table S11. SNPs in the exon of the *DvCHS1* genomic region in MJO\textsubscript{r}, MJY, and MJW.

Table S12. SNPs in the intron of the *DvCHS1* genomic region in MJO\textsubscript{r}, MJY, and MJW.

**Acknowledgements**

We thank Yukihiro Fukuda for supplying ‘Michael J’, and Dr Yoshikazu Tanaka (Suntory, Japan) for supplying the dahlia cDNA library.

**References**


Lawrence WJC. 1929. The genetics and cytology of *Dahlia* species. *Journal of Genetics* 21, 125–159.


Ludwig SR, Habera LF, Dellaporta SL, Wessler SR. 1989. Lc, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region. Proceedings of the National Academy of Sciences, USA 86, 7092–7096.


Nacken WKF, Piotrowiak R, Saedler H, Sommer H. 1991. The transposable element Tam1 from Antirrhinum majus shows structural homology to the maize transposable En/Spm and has no sequence specificity of insertion. Molecular and General Genetics 228, 201–208.


Xu M, Brar HK, Grosic S, Palmer RG, Bhattacharyya MK. 2010. Excision of an active CACTA-like transposable element from DFR2 causes variegated flowers in soybean (Glycine max (L.) Merr.). Genetics 184, 53–63.


