ORIGINAL PAPER

**Takashi Nakatsuka · Masahiro Nishihara
Keiichiro Mishiba · Hiroshi Hirano
Saburo Yamamura**

Two different transposable elements inserted in flavonoid 3′,5′-hydroxylase gene contribute to pink flower coloration in *Gentiana scabra*

Abstract Pink-flowered gentian plants (*Gentiana scabra*) have been bred from spontaneous mutations of blue-flowered gentian plants, but the formation mechanism(s) is unknown so far. To investigate the process, two independent pink-flowered gentian plant lines were analyzed by a molecular biological approach. HPLC analysis showed that petals of the blue-flowered cultivar contained a small amount of cyanidin derivatives and major delphinidin derivatives, whereas pink petals had only a small amount of cyanidin derivatives. To find the causal factor(s) of this change, we focused on flavonoid 3′,5′-hydroxylase (F3′,5′H), which is a key enzyme for delphinidin biosynthesis in the flavonoid biosynthetic pathway. Molecular analyses confirmed that the loss of delphinidin synthesis could be attributed to the insertions of different transposable elements in the *F3′,5′H* gene in each independent pink-flowered gentian plant. Sequence analysis showed that these transposable elements were classified into an *hAT* superfamily and terminal-repeat retrotransposon in miniature (TRIM), by which normal *F3′,5′H* transcripts were interrupted. Southern blot analysis indicated that they belong to high copy number elements and are also found in a related gentian species (*G. triflora*). These results suggest that the transposable elements inserted in *F3′,5′H* are the source of the mutations and may also play a substantial role in the genomic evolution of the genus *Gentiana*.

Introduction

The pigmentations that color most flowers, fruits, and seeds, are anthocyanin secondary metabolites, and their biosynthesis has been well characterized in the flowers of petunia (*Petunia hybrida*) and snapdragon (*Antirrhinum majus*) and in the kernels of maize (*Zea mays*) (see reviews; Holton and Cornish 1995; Mol et al. 1998; Winkel-Shirley 2001; Koes et al. 2005). Among these, transposable elements greatly contribute to advances in the discovery and characterization of new flavonoid biosynthesis-related genes (van Houwelingen et al. 1998).

Transposable elements have been isolated from a number of species including higher plants, and can be divided into two classes according to whether their transposition intermediate is RNA or DNA, designated as retrotransposon (class 1) or DNA transposon (class 2), respectively (Feschotte et al. 2002; Graig 2002). Retrotransposons are the most abundant and widespread class of eukaryotic transposable elements, consisting of either a long terminal repeat (LTR) or non-LTR retrotransposons (Kumar and Bennetzen 1999). LTR-retrotransposons have LTRs as direct repeats and autonomous elements including at least two genes, so-called *gag* and *pol*. LTR-retrotransposons are further subclassified into *Ty-1 copia* and *Ty3-gypsy* groups according to their sequence similarity and order of encoded gene products. In contrast, non-LTR retrotransposons are divided into long interspersed repetitive

T. Nakatsuka · M. Nishihara (✉) · K. Mishiba
H. Hirano · S. Yamamura
Iwate Biotechnology Research Center,
22-174-4, Narita, Kitakami,
024-0003 Iwate, Japan
E-mail: mnishiha@ibrc.or.jp
Tel.: +81-197-682911
Fax: +81-197-683881

Present address: K. Mishiba
Graduate School of Agriculture and Biological Sciences,
Department of Plant Science, Osaka Prefecture University,
1-1, Gakuen-cho, Sakai, 599-8531 Osaka, Japan

elements (LINE) and short interspersed repetitive elements (SINE). Recently, terminal-repeat retrotransposon in miniature (TRIM) that belong to a new group of LTR-retrotransposons have been identified in several plant species including *Arabidopsis* and rice from the analysis of genome sequence data (Witte et al. 2001).

DNA transposable elements can be grouped into families based on both the sequence similarity of the homologue of terminal inverted repeats (TIR) and the specific number of nucleotides comprising the target site duplication (TSD) caused by insertion of the element into the host genomic DNA (Kunze and Weil 2002). Members of the *hAT* family, which include activator/dissociation (*Ac/Ds*) elements of maize (Kunze and Starlinger 1989) and *Tam3* elements of snapdragon (Hehl et al. 1991), cause an 8 bp direct duplication upon insert into a new site, whereas members of the CACTA family, which include enhancer/suppressor-mutator (En/Spm) elements of maize (Pereira et al. 1986) and *Psl* of petunia (Snowden and Napoli 1998), cause a 3 bp of TSD. Besides, miniature inverted repeat transposable elements (MITEs) are found as small transposable elements that occur in high copy number in the host genomes, containing *mPing* elements of *Oryza sativa* activated in cells derived from anther culture (Kikuchi et al. 2003; Nakazaki et al. 2003; Jiang et al. 2003). Most gentian plants (*Gentiana scabra*, *G. triflora* and their hybrids) cultivated in Japan exhibit blue or purple flowers, which contain anthocyanin from delphinidin derivatives, named gentiodelphin (Goto et al. 1982) and albireodelphin (Hosokawa et al. 1997). Recently, pink-flowered gentian cultivars have been bred using mutants that have arisen from natural variations. A pink-flowered gentian is known to accumulate cyanidin derivatives, called gentiocyanin, lacking delphinidin derivatives in petals (Hosokawa et al. 1995). The structural genes involved in anthocyanin biosynthesis have been isolated and characterized from the flowers of *G. triflora* (Tanaka et al. 1996; Kobayashi et al. 1998; Fujiwara et al. 1998; Fukuchi-Mizutani et al. 2003; Nakatsuka et al. 2005). The key enzyme of delphinidin biosynthesis is known to be flavonoid 3′,5′-hydroxylase (F3′,5′H), which catalyzes the hydroxylation at the 5′ position of the B-ring of the anthocyanidin skeleton. F3′,5′H belongs to the cytochrome P450 superfamily, and its cDNAs have been isolated from petunia (Holton et al. 1993) and gentian (Tanaka et al. 1996). Therefore, we hypothesized that pink-flowered gentian plants might have an absence of F3′,5′H activity due to certain mutations in the *F3′,5′H* gene.

In this study, we report on the isolation and characterization of two transposable elements, *hAT* transposable elements and TRIM within the *F3′,5′H* gene in two different pink-flowered gentian plants. The two transposable elements are thought to be nonautonomous and these homologues are abundant in the gentian genome. This is the first-known report on transposable elements in the *Gentianaceae* family containing approximately 1,200 species comprising many valuable pharmaceutical and ornamental plants.

Materials and methods

Plant materials

Ornamental gentian plants, *Gentiana scabra* cv. Momokorin (pink flower), a breeding line 13–98 (pink flower), cv. Alta (blue flower), and *G. triflora* cv. Maciry (blue flower), were provided by the Iwate Agriculture Research Center. The petals and leaves collected from each cultivar were stored at −80°C until used for analysis.

Anthocyanin extraction and HPLC analysis

For anthocyanidin analysis, frozen petal samples (250 mg) were ground in liquid nitrogen and then extracted using 1 ml EAA solution (ethanol:acetic acid:water = 10:1:9) with gentle shaking at 4°C overnight. For hydrolysis of anthocyanin, 4.5 ml 3 N HCl was added and incubated at 100°C for 90 min. After the hydrolysis, the solution was extracted with 1 ml isoamylalchol and filtered through a 0.22 μm filter (Millipore); HPLC analysis was carried out by a high performance liquid chromatography (HPLC) system (D-7000 HPLC system manager, Hitachi). HPLC was carried out with a reverse phase column J' sphere ODS M80 (4.6×150 mm, YMC) using 15% acetonitril and 3% acetic acid as the solvent for 20 min at 40°C at a flow rate of 1.0 ml/min. Anthocyanin was quantified by monitoring the peak area of absorbance at 500 nm, with pelargonidin, cyanidin, and delphinidin as the standard.

Northern analysis of F3′,5′H and F3′H genes

Total RNA was isolated from 100 mg fresh weight of gentian petals at flower development stage 3 as defined by Nakatsuka et al. (2005) using concert Plant RNA reagent (Invitrogen, CA, USA). The 5 μg of total RNA was separated on 1.25% MOPS-agarose gel, and transferred to a Hybond-N+ membrane (Amersham Pharmacia, Sweden). The membrane was hybridized and detected as described by Nakatsuka et al. (2005).

### The characterization of F3′,5′H cDNA in pink-flowered gentian plants

The first strand cDNA was synthesized using the total RNA of the gentian petals by using RNA PCR Kit (AMV) version 2.1 (TaKaRa, Tokyo) according to the manufacturer's protocol. The primers designed for amplification of the *F3′,5′H* open reading frame (ORF), were used as follows: forward primer 5′-<u>ATG</u> CCC ATA

AAA ATG TCA CCC ATT-3′ (start codon indicated by the underline) and reverse primer 5′-<u>TTA</u> AGG AGC ATA AAC ATG AAG AGG -3′ (stop codon indicated by the underline). For 3′RACE analysis, the combination of the forward primer for the ORF amplification and M13-M4 primer was used. The reaction mixture in a total volume of 50 μl contained 1 μl first strand cDNA, 200 μM dNTPs, 2.5 mM MgCl₂, 5 μM each primer, 2.5 U of LA Taq polymerase (TaKaRa, Tokyo), and 1× LA PCR buffer. The reaction conditions were preheating at 94°C/1.5 min, 35 cycles at 95°C/20 s, 55°C/40 s, and 72°C/2 min, and then extension at 72°C/10 min. The amplified fragments were purified by GENE-CLEAN II (BIO101, CA, USA) and then subcloned into a pCR4-TOPO cloning vector for sequencing (Invitrogen). Sequence analyses were carried out using Big-Dye Terminator cycle sequencing Kit version 1.1 and DNA sequencer model ABI PRISM 377 (Applied Biosystems Japan, Tokyo). Sequence analyses were carried out using GENETYX-MAC version 12.0 (GENETYX, Tokyo) and the BLAST network service from the National Center for Biotechnology Information (NCBI) (Altschul et al. 1990).

The $F3′,5′H$ genomic sequence of pink-flowered gentian plants was amplified using the primer sets described earlier. The reaction mixture contained 100 ng genomic DNA as the template and was subjected to the same conditions as in the RT-PCR

Fig. 1 Flower colors and HPLC profiles of anthocyanidin in blue- and pink-flowered gentian plants. Three gentian cultivars used in this study are shown, cv. Alta (**a**), line 13–98 (**b**), and cv. Momokorin (**c**). Anthocyanin was extracted from the petals and then hydrolyzed to convert it to anthocyanidin. Anthocyanidin extractions from each cultivar were subjected to HPLC analysis as described in Materials and methods. *dp* and *cy* indicate delphinidin and cyanidin, respectively
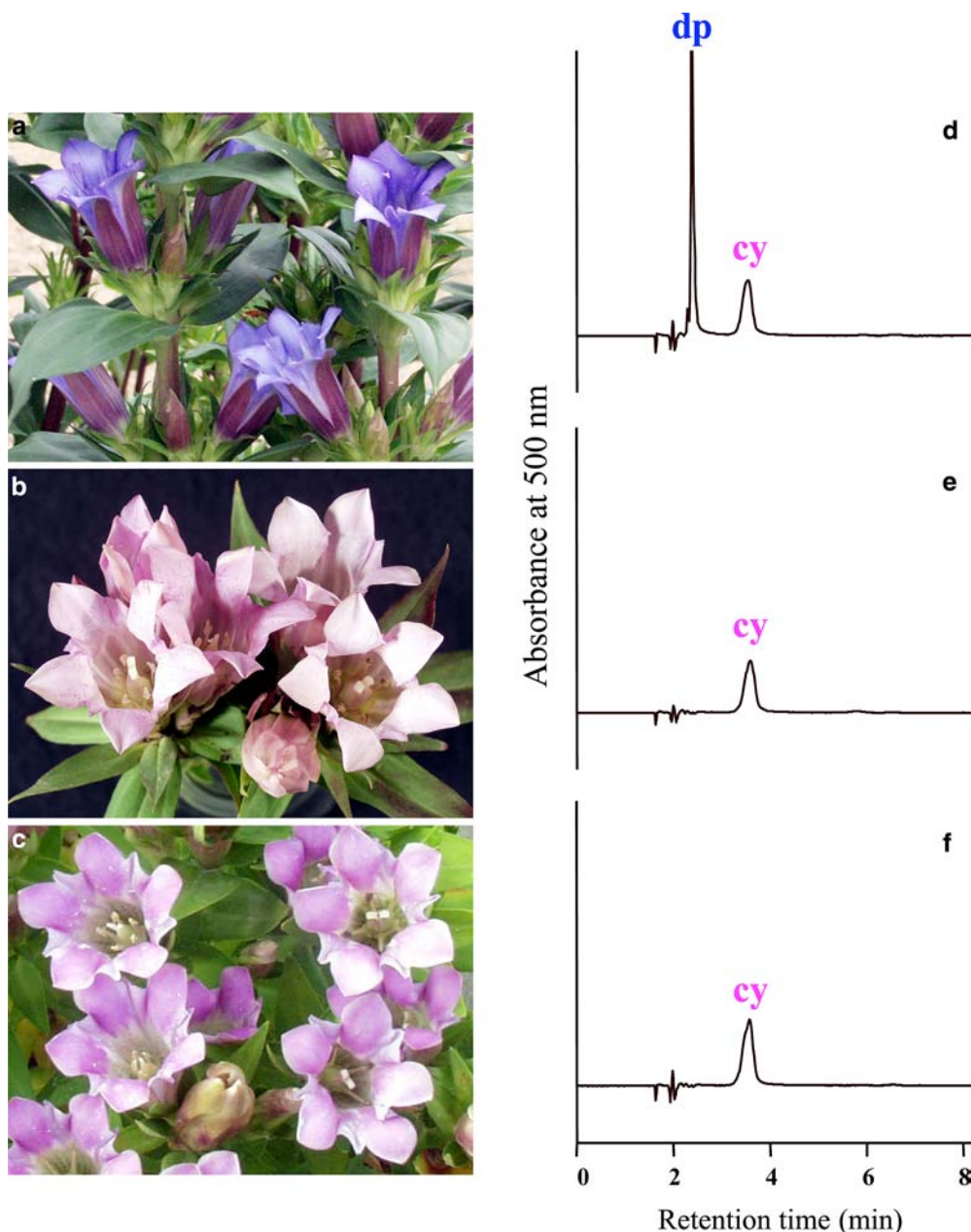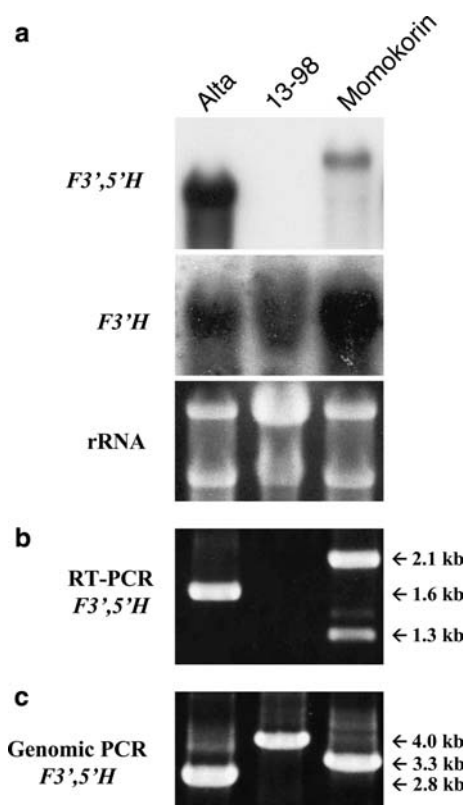
Fig. 2 Expression analyses and genomic-PCR analysis of flavonoid hydroxylase genes in the petals of blue- and pink-flowered gentian plants. **a** Northern blot analysis was performed using total RNA (5 μg) from the petals of blue-flowered gentian cv. Alta, pink-flowered gentian line 13–98 and cv. Momokorin. The blot was hybridized by each probe for $F3',5'H$ or $F3'H$ gene. The ribosomal RNA bands stained with ethidium bromide used as a control are also shown. **b** RT-PCR analysis was performed to amplify $F3',5'H$ ORF using total RNAs isolated from the petals. *Numerals on the right* indicate the sizes of the amplified fragments estimated by DNA marker. **c** PCR analysis was performed to amplify the genomic $F3',5'H$ sequence with the primer set used by RT-PCR analysis. *Numerals on the right* indicate the same as (**b**)

analysis. The obtained fragments were subcloned into a pCR4-TOPO cloning vector for sequencing (Invitrogen). Sequence analysis was performed as described earlier.

Southern blot analysis of $F3',5'H$ and transposable elements

Total genomic DNA was isolated from a 1 g leaf sample using Nucleon PhytoPure (Amersham Biosciences) as described in the manufacturer's instruction. The 10 μg of genomic DNA digested by *Eco*R I and *Hin*d III (TaKaRa, Tokyo) was separated on 0.6% agarose gel and then transferred to a Nytran N membrane (Schleicher & Schuell, Germany). $F3',5'H$ fragments were amplified using the primer sets described earlier, the truncated transposase sequence of the *dTgs1* fragment was amplified using primer sets as follows:

dTgs1-U, 5′-CGA CGT GGA GTT GAT GGT AG-3′; *dTgs1*-L, 5′-CTC CTA GCA TTT TTA GAA GC-3′, and *GsTRIM1* was amplified using M13-M4 and M13-RV attached within the pCR4-TOPO vector. Amplified fragments were labeled with a DIG-High Prime DNA labeling Kit (Roche Diagnostics, Germany) according to the manufacturer's instructions. Each probe was hybridized and detected as described by Nakatsuka et al. (2005).
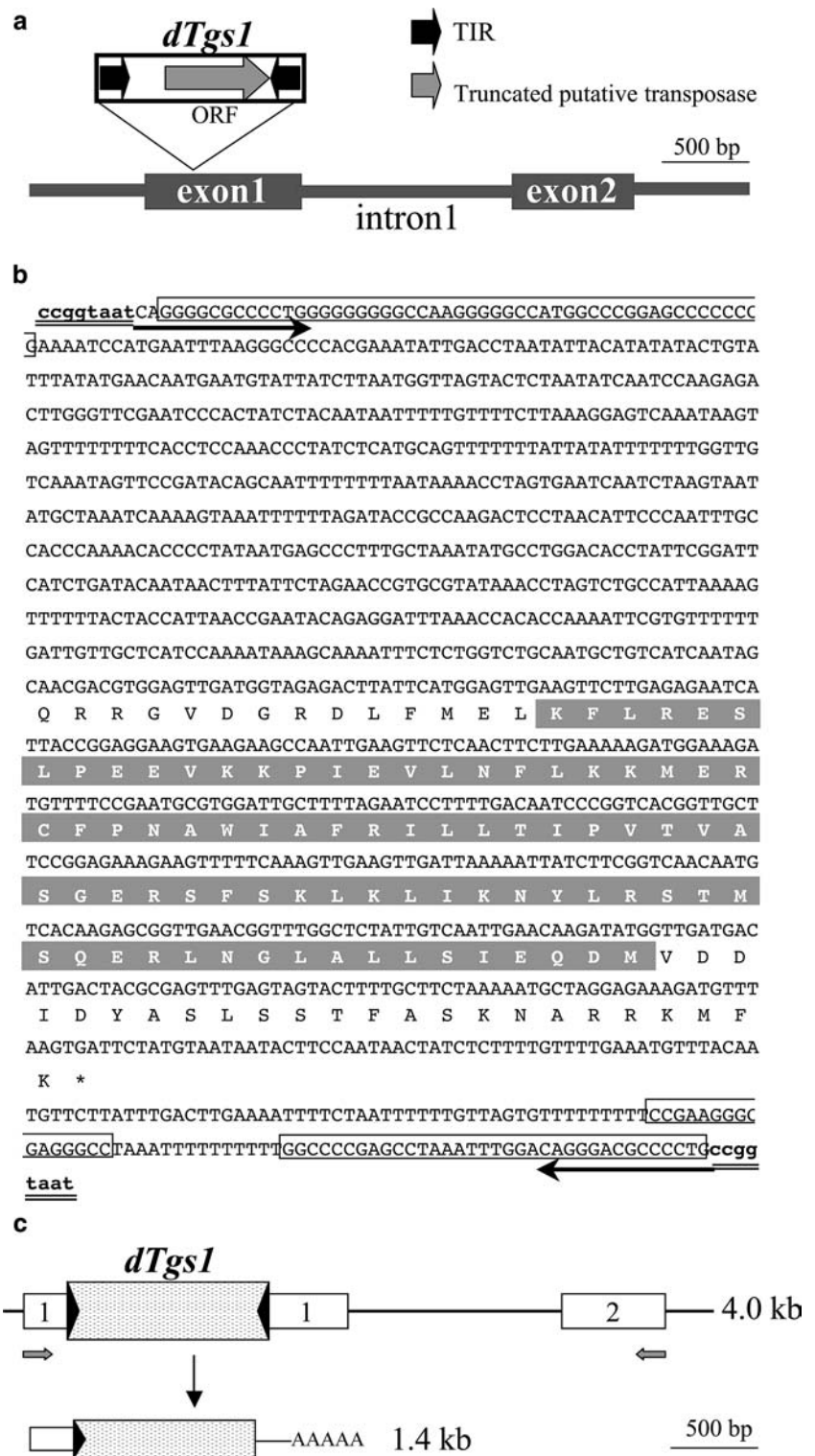
Results

Anthocyanidin accumulation in pink-flowered gentian plants

Pictures of three gentian plants used in this study are shown in Fig. 1. To measure the anthocyanidin compounds, HPLC samples were extracted from the petals of one blue- and two pink-flowered gentian plants. The HPLC profile of the blue-flowered gentian cv. Alta (Fig. 1a) showed two peaks, indicating co-accumulation of major delphinidin and minor cyanidin derivatives in the petals (Fig. 1d). On the other hand, the two pink-flowered gentian plants, line 13–98 (Fig. 1b) and cv. Momokorin (Fig. 1c), showed a single major peak, indicating an accumulation of cyanidin derivatives, but not delphinidin derivatives (Fig. 1e, f, respectively). The amount of cyanidin derivatives in the two pink-flowered gentian plants did not remarkably increase in comparison with that in the blue-flowered gentian.

The expression of flavonoid hydroxylase genes in pink-flowered gentian plants

Because pigment analysis confirmed that pink-flowered gentian plants accumulated only cyanidin derivatives, the difference in hydroxylation activity at the B-ring in the flavonoid skeleton was considered to cause this change. Therefore, one of the flavonoid biosynthetic structural genes, $F3',5'H$, encoding an enzyme catalyzing the hydroxylation at the 3′- and 5′-positions of the flavonoid skeletons, was first analyzed with respect to gene expressions between the blue- and pink-flowered gentian plants. No transcript of the $F3',5'H$ gene in the pink-flowered gentian line 13–98 was observed by either northern blot (Fig. 2a) or RT-PCR analyses (Fig. 2b) using primer pairs for full length of $F3',5'H$ ORF, although minor transcripts of 1.4 kb in lengths were amplified by 3′RACE analysis (see subsequently). In contrast, the $F3',5'H$ transcript of another pink-flowered cv. Momokorin was detected at a higher molecular weight with a remarkably reduced amount compared with blue-flowered cv. Alta (Fig. 2a). By RT-PCR analysis using primers designed to amplify the full-length ORF of the $F3',5'H$ gene, cv. Momokorin showed the amplified fragments of about 1.3 and 2.1 kb, differing from the 1.6 kb single amplified fragment in

Fig. 3 Schematic representation of structure, sequence, and transcripts of $F3',5'H$ gene in line 13–93. **a** Structure of genomic $F3',5'H$ gene. The $F3',5'H$ genomic structure of the pink-flowered gentian line 13–98 is 4.0 kb in length, containing 1.2 kb of $dTgs1$ transposable element (hAT family) within the first exon. **b** Nucleotide and deduced amino acid sequence of $dTgs1$. $dTgs1$ in the $F3',5'H$ gene of line 13–98 has 1,187 bp of hAT superfamily transposon. 8 bp of target sequence duplication (TSD) is shown *double underlined*. 5′ and 3′ inverted repeats (IR) are indicated by *arrows*, and CG-rich sequences of subterminal regions are *boxed by lines*. A truncated deduced amino acid sequence of transposase is shown in a lower sequence, and C-terminal conserved domain of hAT transposon, which is prospected by programs of CD-Search (Marchler-Bauer and Bryant 2004), is indicated by a *closed gray box*. **c** Minor transcripts of $F3',5'H$ gene. The primer positions with RT-PCR and genomic PCR analysis in Fig. 2 are shown by *thick arrows*. The *numeral at the right of the illustration* represents the fragment length obtained by genomic PCR or 3′-RACE analysis. The $F3',5'H$ gene in line 13–98 underwent alternative polyadenylation, resulting in short incomplete $F3',5'H$ transcripts

blue-flowered gentian (Fig. 2b). These results show that neither of the pink-flowered gentian plants accumulated normal $F3',5'H$ transcripts in the petals. On the other hand, mRNA accumulation of $F3'H$, encoding an enzyme catalyzing the hydroxylation at 3′-position of the flavonoid B-ring, was similar among the petals of blue- and pink-flowered gentian plants (Fig. 2a),

suggesting that the $F3'H$ gene did not contribute to the inability to accumulate delphinidin derivatives in the pink flowers. All other flavonoid biosynthesis-related structural genes (*CHS, CHI, F3H, DFR, ANS, 3GT, 5AT,* and *FSII*) also exhibited similar expression among the petals of the pink- and blue-flowered gentian plants (data not shown).

Cloning of genomic F3′,5′H gene in pink-flowered gentian plants

To examine the genome structures of the F3′,5′H gene, the genomic DNAs of the blue- and pink-flowered gentian plants were used to amplify full-length fragments containing F3′,5′H ORF. Different length amplified fragments were obtained from each gentian line (Fig. 2c). Blue-flowered cv. Alta showed about a 2.8 kb amplified fragment, whereas, the amplified fragments of genomic F3′,5′H ORF in the pink-flowered gentian line 13–98 and cv. Momokorin were observed to be about 1.2 and 0.5 kb longer than that of blue-flowered gentian, respectively. Southern blot analysis of F3′,5′H also supported this result (see subsequently). Both the 4.0 and 3.3 kb of the genomic F3′,5′H fragments in the pink-flowered gentian plants were subcloned and their sequences were determined.

Transposable element inserted in the F3′,5′H gene of pink-flowered gentian line 13–98

4.0 kb of the F3′,5′H genomic region in line 13–98 contained an insertion of 1,187 bp in length at the 239 bp position from the start codon within the first exon, designated as dTgs1 (Fig. 3a). dTgs1 showed distinctive features of the hAT superfamily members of class 2 transposable elements, such as an 8 bp of target site duplication (TSD) and a 14 bp of terminal inverted repeats (TIR). The 14 bp of TIR had one substitution between 5′- and 3′-TIRs (Fig. 3b and Table 1). They showed a unique perfect or imperfect palindrome structure, and exhibiting partial similarity to that of dTph1 in petunia (Table 1). A sequence containing truncated ORF near the 3′ flank of this insertion has homology to similar regions in the A. thaliana hAT dimerization domain-containing protein (accession number NP_178689, 582 residues; 71% amino acid identity over 116 residues). The nucleotide sequences of the proximal region of both TIRs, recognized by Ac transposase, contained a CG-rich sequence, but not the conserved AAACGG motifs observed in Ac transposable elements in maize.

### Transposable element inserted in the F3′,5′H gene of pink-flowered gentian cv. Momokorin

A 3.3 kb of genomic F3′,5′H in cv. Momokorin contained an insertion of 513 bp in length at the 87 bp position from the start codon within the first exon, designated as GsTRIM1 (Fig. 4a). GsTRIM1 showed the feature of TRIM belonging to class 1 transposable elements, such as a 5 bp of TSD and a 76 bp of terminal direct repeats (TDR) containing 8 bp inverted repeats (TGTTAAAG) at both ends (Fig. 4b). However, unlike the usual TRIM elements, GsTRIM1 contained three TDRs the second of which had an incomplete structure in comparison with the TDRs of both ends. Further examination of this insertion revealed the presence of primer binding sites (PBS) and polypurine tracts (PPT) between each TDR. This insertion did not contain sequences encoding any proteins, indicating that this is a nonautonomous transposable element like other TRIM elements. Sequence analysis also showed that the GsTRIM1 transposable element was similar to some sequences in the Lotus japonicus genomic library (data not shown). In particular the 3′-end of TDR of GsTRIM1 was similar to that of the L. japonicus GsTRIM1 homologue, designated as LjTRIM. LjTRIM also contained some copies at independent loci within the L. japonicus genome and had the same feature of 138–141 bp of TDR, PBS, and PPT motifs as GsTRIM1 (Fig. 5a). Because TRIM elements

Table 1 Comparison of plant hAT element terminal inverted repeats (TIRs)

Name	Sequence	Length (bp)	Species	Reference
DTgs1	5′-**CAGGGGCG**cCCCTG 3′-**CAGGGGCG**tCCCTG	14	Gentina scabra	In this study
DTph1	5′-**CAGGGGCG**GAGC 3′-**CAGGGGCG**GAGC	12	Petunia hybrida	Gerats et al. (1990)
Ac/Ds	5′-c**AGGG**ATGAAA 3′-t**AGGG**ATGAAA	11	Zea mays	Muller-Neumann et al. (1984)
Bg	5′-**CAGGG** 3′-**CAGGG**	5	Zea mays	Brown et al. (1989)
Tag2	5′-cAGAAccGTCAATTGGGC 3′-tAGAAgtGTCAATTGGGC	18	Arabidopsis thaliana	Henk et al. (1999)
Tag1	5′-CAATTGTTTTCACGCCCGACCCG 3′-CAAT-GTTTTCACGCCCGACCCG	22	A. thaliana	Liu and Crawford (1998)
Tam3	5′-TAAAGATGTGAA 3′-TAAAGATGTGAA	12	Antirrhinum majus	Hehl et al. (1991)
Slide-124	5′-TAATGCTG 3′-TAATGCTG	8	Nicotiana tabacum	Grappin et al. (1996)

The sequence of each transposable element similar to TIR of dTgs1 is shown in bold

Fig. 4 Schematic representation of structure, sequence, and transcripts of *F3',5'H* gene in cv. Momokorin. **a** Structure of genomic *F3',5'H* gene. The *F3',5'H* genomic structure of the pink-flowered gentian cv. Momokorin is 3.3 kb in length, containing 520 bp of *GsTRIM1* transposable element within the first exon. **b** Nucleotide sequence of *GsTRIM1*. *GsTRIM1* has the features of terminal-repeat retrotransposons in miniature (*TRIM*), containing 5 bp of target sequence duplication (TSD, *double underlined*) and triple 76 bp of terminal direct repeats (TDR, *boxed lines*) with 8 bp of inverted repeats (IR, *thick arrows*). Primer binding sites (*PBS*) and polypurine tract (*PPT*) are indicated by *solid* and *dashed lines*, respectively. A 5'-splicing recognition site of alternative splicing as observed in Fig. 2b is *boxed by thick lines*. **c** Transcripts of *F3',5'H* gene. The primer positions with RT-PCR and genomic PCR analysis in Fig. 2 are shown by *thick arrows*. The *numeral at the right of the illustration* represents the fragment length obtained by Genomic PCR or RT-PCR analysis. Two abnormal *F3',5'H* transcripts accumulated in pink-flowered gentian cv. Momokorin by the insertion of *GsTRIM1* or alternative splicing
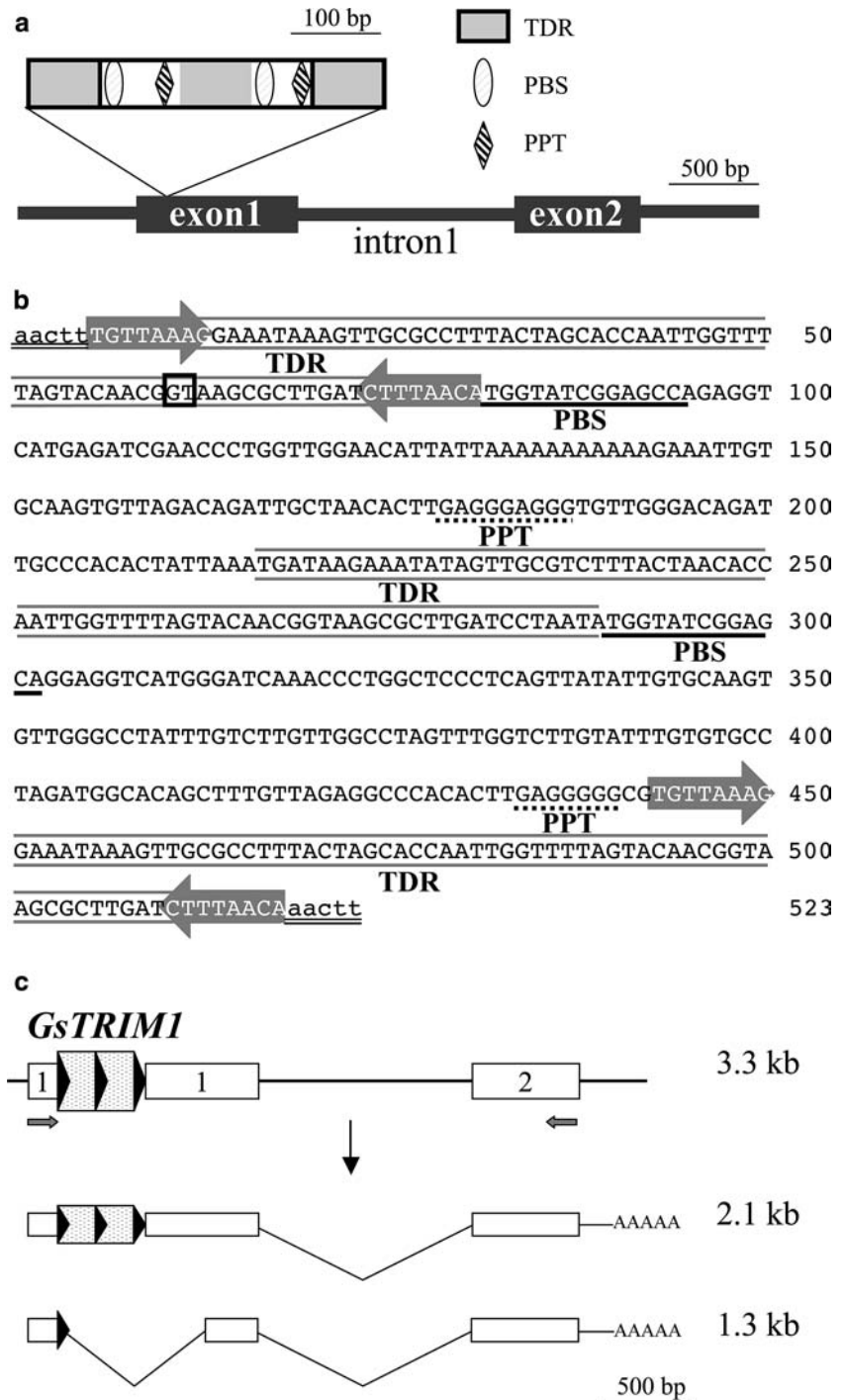

were first identified by database analysis and have the striking peculiarity to be quite conserved among plant species, we performed a phylogenic tree analysis (Fig. 5b) using TDR sequences. The result clearly showed that *GsTRIM1* and *LjTRIM* positioned apart from the previous reported TRIMs, indicating that these two are evolutionarily different from other plant TRIM elements. Notably, the TDR sequences of *GsTRIM1* had an incomplete palindrome structure (Fig. 5c).

Effect of transposable element insertions on the *F3',5'H* transcription in pink-flowered gentian plants

Since two pink-flowered gentian plants had insertions of independent transposable elements in the *F3',5'H* gene, functional *F3',5'H* mRNA could not be transcribed. Namely, *F3',5'H* transcripts of line 13–98 were suppressed by the insertion of *dTgs1*, while two different sizes of transcripts were observed by the insertion of

Fig. 5 Comparison of TRIM elements among higher plants. **a** Comparison of overall structure of TRIM elements between *G. scabra* (*GsTRIM*) and *L. japonicus* (*LjTRIM*). *GsTRIM* and *LjTRIM* were presumed from *GsTRIM1* and 15 independent *GsTRIM1* homologues of *L. japonicus*, respectively. **b** Phylogenic analysis of TDR of *GsTRIM1* with those of other higher plants. The phylogenic tree was generated using CLUSTALW (Thompson et al. 1994) and TREEVIEW (Page 1996) programs. *Numerals* indicate bootstrap values from 1,000 replicates. The *bar* indicates an evolutionary distance of 0.1%. Accession numbers in the GenBank/DDBJ databases are as follows: *A. thaliana At1* (gi 6598490, position 8,088–8,204), *A. thaliana At2* (gi 7649355, position 68,877–68,994), *G. scabra GsTRIM1* (the present study, AB222606), *Lycopersicon esculentum* (gi 4220970, position 38–157), *L. japonicus LjTRIM* (AJ580824, position 45,546–45,170), *Medicago sativa* (gi 19642, position 8,088–8,204), *Nicotiana tabacum* (gi 9392606, position 3,821–3,947), *O. sativa* (gi 10800055, position 68,877–68,994), *Phaseolus vulgaris* (gi 2576326, position 1,878–1,761), *Solanum tuberosum* (gi 14599414, position 1,522–1,651). **c** Predicted secondary structure of TDR of *GsTRIM1*. The potential secondary structure is predicted with GENETYX-Mac software
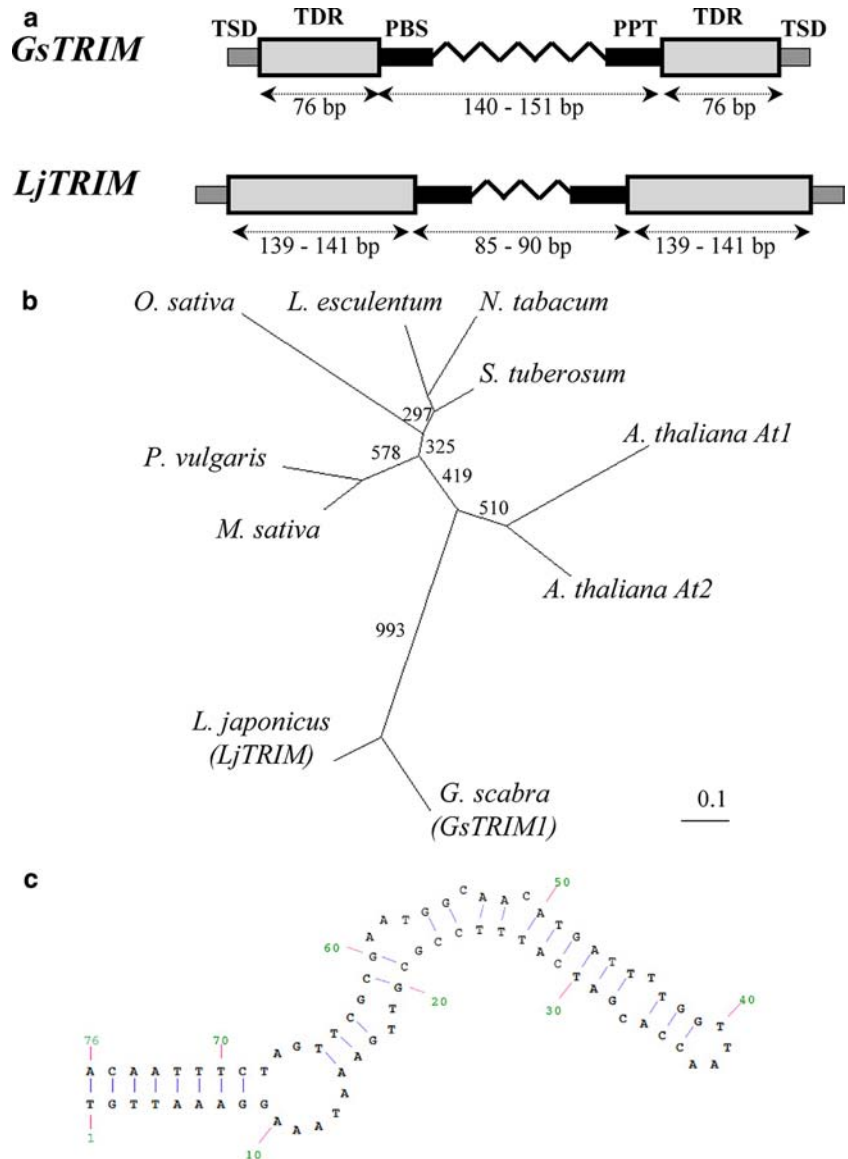
GsTRIM1 in cv. Momokorin (Fig. 2a, b). However, when 3′RACE analysis was performed using forward primer within exon 1, minor transcripts of 1.4 kb in length were also detected in line 13–98. Sequence analysis revealed that the 1.4 kb of the minor *F3′,5′H* transcripts were derived from alternative polyadenylation after the stop codon of putative transposase of *dTgs1*, resulting in short incomplete *F3′,5′H* transcripts that could not translate functional F3′,5′H enzyme (Fig. 3c). To investigate alternative *F3′,5′H* transcripts of cv. Momokorin, the RT-PCR products, 1.3 and 2.1 kb fragments (Fig. 2b), were also subcloned and subjected to sequence analysis. The 2.1 kb of the *F3′,5′H* transcript contained the sequence of 513 bp *GsTRIM1* within the *F3′,5′H* mRNA, whereas the 1.3 kb fragment of *F3′,5′H* was shorter than the transcripts of blue-flowered gentian, which might be due to an alternative splicing within the first TDR of *GsTRIM1* (Fig. 4b, c). Neither *F3',5'H* transcripts of cv. Momokorin could

translate any functional F3′,5′H enzymes by the non-sense codon within the first TDR of *GsTRIM1*. This was also confirmed by 3′RACE analysis of cv. Momokorin transcripts.

Existence of *dTgs1* and *GsTRIM1* transposable elements in the *Gentiana* genome

The copy numbers of *F3′,5′H*, *dTgs1*, and *GsTRIM1* were determined in the three *G. scabra* lines by Southern blot analysis (Fig. 6). Moreover, to investigate whether both transposable elements were present universally in the genus *Gentiana*, *G. triflora* that is a related species to *G. scabra* was also analyzed. Because the restriction enzymes that did not cut the *F3′,5′H* gene were used, the analysis suggested that *G. scabra* cv. Alta, line 13–98 and cv. Momokorin contain at least two copies of *F3′,5′H* in the genome, while *G. triflora* cv. Maciry has four copies.

Both *F3',5'H* loci of cv. Momokorin observed an increase of about 0.5 kb in length by the insertion of the *GsTRIM1* transposable element compared with cv. Alta. About a 1.2 kb shift was also observed in *Eco*R I-digested DNAs in line 13–98 by the insertion of the *dTgs1* transposable element, while three *Hin*d III fragments that differ from cv. Alta were observed. In addition, about a 2.5 kb *Eco*R I fragment was newly observed compared with cv. Alta, suggesting rearrangement or base substitution by the insertion of a transposable element occurring in line 13–98. A number of both *dTgs1* and *GsTRIM1* transposable elements were detected in the genome of all samples, suggesting that they are members of a relatively high copy number family, the existence of which is widespread in the genome of the genus *Gentiana*.

Discussion

In this study, we showed that two pink-flowered gentian plant lines are caused by independent insertions of transposable elements into the first exon of the *F3',5'H* gene. Two transposable elements, *dTgs1* and *GsTRIM1*, were classified as class 2 and class 1 transposable elements, respectively, and are interesting cases in which such different types of transposable elements cause the same phenotypic change of flower color. In petunia, corollas with a small number of purple sectors on a predominantly pink background have been identified as the result of inserting a petunia *Spm-like* (*Psl*) transposable element into the *Hf1* allele, encoding *F3',5'H* (Snowden and Napoli 1998).

dTgs1 showed the feature of the *hAT* superfamily, containing an 8 bp direct duplication of the target insertion site and truncated sequences of putative *hAT* transposase (Fig. 3, Table 1). The *hAT* transposon family, such as those containing *Ac/Ds* of maize (Kunze and Sterlinger 1989) and *Tam3* of snapdragons (Hehl et al. 1991), was one of the largest families in the plant mobile elements. *dTgs1* contains imperfect inverted repeats of 14 bp that share homology with a number of other *hAT* transposable elements, with the most similarity to the *dTph1* of petunia (Gerats et al. 1990). Furthermore, the terminal 5 nucleotides of *dTgs1* were completely consistent with *dTph1*, *Ac/Ds* (Muller-Neumann et al. 1984), and *Bg* of maize (Brown et al. 1989). *dTgs1* contained only a truncated putative transposase gene and is probably a nonautonomous transposable element, depending on other transposable elements having mobility components. Insertion of *dTgs1* in *F3',5'H* gene inhibited normal *F3',5'H* transcriptions (Fig. 2a, b) and caused low levels of incomplete *F3',5'H* transcripts (Fig. 3c). These suggest that the 3' region of putative transposase has a functional structure in relation to polyadenylation. Southern blot analysis showed that *dTgs1* is abundant and widespread in the genomes of *G. scabra* and *G. triflora* (Fig. 6), perhaps a few of them might also contain *Tgs1* autonomous transposable elements. However, the 'revertant' phenotype, exhibiting blue flowers, has not yet been obtained from line 13 to 98, suggesting that *dTgs1* inserted in the *F3',5'H* gene has not been activated under normal cultivated conditions. It should be noted that *Tam3* of snapdragons was known to be strictly controlled by low temperature (Carpenter et al. 1987), and transposition of transposable elements could also be controlled by DNA methylation (Miura et al. 2001; Hashida et al. 2003). The subterminal sequence of *dTgs1* contains a CG-rich sequence (Fig. 3) and it might possibly be silenced by DNA methylation. There is also another possibility to explain why no excision event has been observed in this line. The nucleotide sequences of the terminal ends of *dTgs1* were imperfect inverted repeats, namely one nucleotide difference was present between TIR sequences. Such an imperfect TIR sequence

Fig. 6 Southern blot analysis of *F3',5'H*, *dTgs1* and *GsTRIM1* elements in *Gentiana* sp. Five microgram genomic DNA of *G. scabra* cv. Alta (*line 1*, blue flower), line 13–98 (*line 2*, pink flower), cv. Momokorin (*line 3*, pink flower) and *G. triflora* cv. Maciry (*line 4*, blue flower) digested by *Eco*R I (*E*) or *Hin*d III (*H*) was separated and hybridized by the probe of *F3',5'H*, truncated putative transposase of *dTgs1* and *GsTRIM1*. *Numbers on the left* of the figure indicate molecular size markers
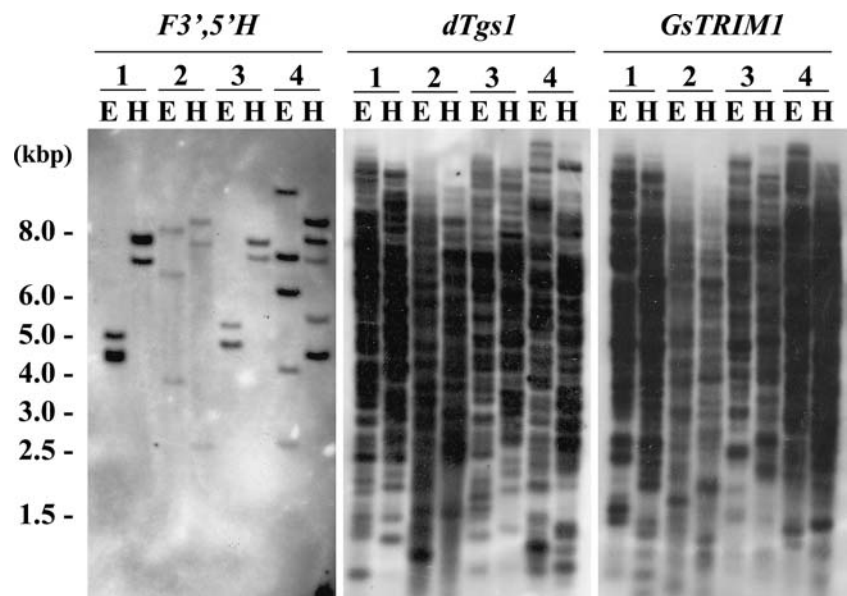
could not be recognized by the transposase of *dTgs1*, so that *dTgs1* could not be excised anymore. Further studies will be needed to elucidate the possible factor involved in *dTgs1* mobility and to identify autonomous *Tgs1* with complete transposase controlling its translocation.

GsTRIM1 showed the feature of TRIM, containing a 5 bp direct duplication of the target insertion site and the triple 76 bp direct repeats of TDR (Fig. 4). Interestingly, the TDR of *GsTRIM1* is the shortest among those of the previously reported plant TRIM and LTR-retrotransposons. For example, the TDR of TRIMs presented in both monocotyledonous and dicotyledonous plants was 100–250 bp in length (Witte et al. 2001, 2005), whereas the LTR of *Tos17* in *O. sativa* was 138 bp (Hirochika et al. 1992), and *Tnt1* in tobacco was 610 bp (Grandbastien et al. 1989). Because the TDR of *GsTRIM1* was an incomplete palindrome sequence (Fig. 5c), its structural features might have not been developed from other LTR-retrotransposons previously reported. Witte et al. (2001) has reported that TRIM was derived from LTR-retrotransposons and also mobilized in trans by other retroelements. Although *GsTRIM1* was possibly mobilized by functional autonomous element counterparts, a full-length LTR-retrotransposon with mobility-related polyproteins has not yet been discovered in gentian. *L. japonicus*, the model plant of the *Fabaceae* family, was found to contain some *GsTRIM1* homologues by a search of genomic data base sequences, but their homologue could not be found in the genomic sequences of other model plants, *Arabidopsis thaliana* and *O. sativa*. *GsTRIM1* homologues (*LjTRIM*) are different from the TRIMs reported previously in *L. japonicus* (Witte et al. 2001). It is probable that the horizontal transmission of retrotransposons between *Gentiana* and *Lotus* occurred in the distant past. This was in part supported by the fact that *GsTRIM* and *LjTRIM* resemble each other apart from the previously reported other plant TRIMs by a phylogenic analysis (Fig. 5b). Southern blot analysis showed that the genome of the *Gentiana* species has high copy numbers of *GsTRIM1*. Because Japanese ornamental gentian species contain a comparatively higher genome size than other closely related gentian species (unpublished data), retrotransposons like *GsTRIM1* might contribute to increasing their genome size. In the case of maize, the large genome size might result from amplification of retrotransposon, comprising about 50–80% of the maize nuclear genome (SanMiguel et al. 1996; SanMiguel and Bennetzen 1998). On the contrary, in the light of the fact that the structure of *GsTRIM1* consists of triple TDRs, homologous recombination between neighboring TDRs might occur as proposed in LTR-retrotransposons (SanMiguel et al. 1996; Vicient et al. 1999). Such unusual TRIM elements are also reported to be present in *Arabidopsis*, though the precise formation mechanism is unknown (Witte et al. 2001). Formation of nested *GsTRIM1* might be involved in genome size reduction and deactivating retrotransposon (Vicient et al. 1999).

Alternative splicing of *F3′,5′H* by the insertion of a *GsTRIM1* transposable element was observed by RT-PCR analysis (Fig. 2b). This 1.3 kb of transcripts was generated de novo by splicing a site within the TDR of *GsTRIM1*, and the first exon of *F3′,5′H*. TRIM (*Katydid*-At1) in *Arabidopsis* has also been hypothesized to create a new intron within a gene (Witte et al. 2001). *Waxy* gene in maize generated multiple alternative splicings but not a newly spliced site by insertion of *Stonor*, classifying into LTR-retrotransposon, because these insertions disrupted long-range splice site recognition leading to novel transcripts (Varragona et al. 1992). In contrast, ca. 5% of all alternative spliced internal exons in the human genome are derived from *Alu* elements (Sorek et al. 2002; Kreahling and Graveley 2004). The insertion of a *GsTRIM1* transposable element in *F3′,5′H* generated a de novo 5′ splice site (Fig. 4); thus, the *GsTRIM1* transposable element might have the capacity of inducing alternative splicing like *Alu* elements. Moreover, the insertion of *GsTRIM1* also caused reduction of *F3′,5′H* transcripts (Fig. 2a). A nonsense codon generated by the point mutations of *F3′H* in *Ipomoea* has been reported to cause nonsense-mediated decay (NMD), affecting the stability of mature mRNA (Hoshino et al. 2003). The stop codon generated in the *GsTRIM1* of cv. Momokorin might also induce the reduction of mature *F3′,5′H* mRNA by NMD. Retrotransposons have been known to be activated by biotic and abiotic stresses, e.g., salicylic acid, methyl jasmonate, wounding, and cell culture (Kumar and Bennetzen 1999). Since it is unknown what stimuli activate TRIMs, activation factor(s) regulating *GsTRIM1* transposition must be further studied.

Taken together, we identified two different transposable elements isolated from pink-flowered gentian plant lines in which their independent insertion events caused the same flower color mutation. There is almost no doubt that transposable elements have influenced genome evolution in gentian plants. Further characterization should lead to a deeper understanding of the roles of these transposable elements during the evolution of gentians.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Brown JJ, Mattes MG, O'Reilly C, Shepherd NS (1989) Molecular characterization of *rDt*, a maize transposon of the "Dotted" controlling element system. Mol Gen Genet 215:239–244

Carpenter R, Martin C, Coen ES (1987) Comparison of genetic behaviour of the transposable element *Tam3* at two unlinked pigment loci in *Antirrhinum majus*. Mol Gen Genet 207:82–89

Feschotte C, Jiang N, Wessler SR (2002) Plant transposable elements: where genetics meets genomics. Nat Rev Genet 3:329–341

Fujiwara H, Tanaka Y, Yonekura-Sakakibara K, Fukuchi-Mizutani M, Nakao M, Fukui Y, Yamaguchi M, Ashikari T, Kusumi T (1998) cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from *Gentiana triflora*. Plant J 16:421–431

Fukuchi-Mizutani M, Okuhara H, Fukui Y, Nakao M, Katsumoto Y, Yonekura-Sakakibara K, Kusumi T, Hase T, Tanaka Y (2003) Biochemical and molecular characterization of a novel UDP-glucose:anthocyanin 3′-O-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from gentian. Plant Physiol 132:1652–1663

Gerats AG, Huits H, Vrijlandt E, Marana C, Souer E, Beld M (1990) Molecular characterization of a nonautonomous transposable element (*dTph1*) of petunia. Plant Cell 2:1121–1128

Goto T, Kondo T, Tamura H, Imagawa H, Iino H, Takeda K (1982) Structure of gentiodelphin, an acylated anthocyanin isolated from *Gentiana makinori*, that is stable in dilute aqueous solution. Tetrahedron Lett 23:3695–3698

Graig NJ (2002) Mobile DNA. In: Craig NJ, Craigle R, Gellert M, Lambowitz AM (eds) Mobile DNA II. American Society for Microbiology Press, Washington DC, pp 3–11

Grandbastien MA, Spielmann A, Caboche M (1989) *Tnt1*, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature 337:376–380

Grappin P, Audeon C, Chupeau MC, Grandbastien MA (1996) Molecular and functional characterization of *Slide*, an *Ac*-like autonomous transposable element from tobacco. Mol Gen Genet 252:386–397

Hashida S, Kitamura K, Mikami T, Kishima Y (2003) Temperature shift coordinately changes the activity and the methylation state of transposon *Tam3* in *Antirrhinum majus*. Plant Physiol 132:1207–1216

Hehl R, Nacken WK, Krause A, Saedler H, Sommer H (1991) Structural analysis of *Tam3*, a transposable element from *Antirrhinum majus*, reveals homologies to the *Ac* element from maize. Plant Mol Biol 16:369–371

Henk AD, Warren RF, Innes RW (1999) A new *Ac*-like transposon of *Arabidopsis* is associated with a deletion of the *RPS5* disease resistance gene. Genetics 151:1581–1589

Hirochika H, Fukuchi A, Kikuchi F (1992) Retrotransposon families in rice. Mol Gen Genet 233:209–216

Holton TA, Brugliera F, Lester DR, Tanaka Y, Hyland CD, Menting JG, Lu CY, Farcy E, Stevenson TW, Cornish EC (1993) Cloning and expression of cytochrome P450 genes controlling flower colour. Nature 366:276–279

Holten TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7:1071–1083

Hoshino A, Morita Y, Choi JD, Saito N, Toki K, Tanaka Y, Iida S (2003) Spontaneous mutations of the flavonoid 3′-hydroxylase gene conferring reddish flowers in the three morning glory species. Plant Cell Physiol 44:990–1001

Hosokawa K, Fukushi E, Kawabata J, Fujii C, Ito T, Yamamura S (1995) Three acylated cyanidin glucosides in pink flowers of *Gentiana*. Phytochemistry 40:941–944

Hosokawa K, Fukushi E, Kawabata J, Fujii C, Ito T, Yamamura S (1997) Seven acylated anthocyanins in blue flowers of *Gentiana*. Phytochemistry 45:167–171

Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, Wessler SR (2003) An active DNA transposon family in rice. Nature 421:163–167

Kikuchi K, Terauchi K, Wada M, Hirano H (2003) The plant MITE *mPing* is mobilized in anther culture. Nature 421:167–170

Kobayashi H, Oikawa Y, Koiwa H, Yamamura S (1998) Flower-specific expression directed by the promoter of a chalcone synthase gene from *Gentiana triflora* in *Petunia hybrida*. Plant Sci 131:173–180

Koes R, Verweij W, Quattrocchio F (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends Plant Sci 10:236–242

Kreahling J, Graveley BR (2004) The origins and implications of alternative splicing. Trends Genet 20:1–4

Kumar A, Bennetzen JL (1999) Plant retrotransposons. Annu Rev Genet 33:479–532

Kunze R, Starlinger P (1989) The putative transposase of transposable element *Ac* from *Zea mays* L. interacts with subterminal sequences of *Ac*. EMBO J 8:3177–3185

Kunze R, Weil CF (2002) The *hAT* and CACTA superfamilies of plant transposons. In: Craig NJ, Craigle R, Gellert M, Lambowitz AM (eds) Mobile DNA II. American Society for Microbiology Press, Washington DC, pp 565–610

Liu D, Crawford NM (1998) Characterization of the germinal and somatic activity of the *Arabidopsis* transposable element *Tag1*. Genetics 148:445–456

Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res 32:W327–331

Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T (2001) Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. Nature 411:212–214

Muller-Neumann M, Yoder JI, Starlinger P (1984) The DNA sequence of the transposable element *Ac* of *Zea mays* L. Mol Gen Genet 198:19–24

Mol J, Grotewold E, Koes R (1998) How genes paint flowers and seeds. Trends Plant Sci 3:212–216

Nakatsuka T, Nishihara M, Mishiba K, Yamamura S (2005) Temporal expression of flavonoid biosynthesis-related genes regulates flower pigmentation in gentian plants. Plant Sci 168:1309–1318

Nakazaki T, Okumoto Y, Horibata A, Yamahira S, Teraishi M, Nishida H, Inoue H, Tanisaka T (2003) Mobilization of a transposon in the rice genome. Nature 421:170–172

Page RDM (1996) TREEVIEW: an application to display phylogenic trees on personal computers. Comp Appl Biosci 12:357–358

Pereira A, Cuypers H, Gierl A, Schwarz-Sommer Z, Saedler H (1986) Molecular analysis of the *En/Spm* transposable element system of *Zea mays*. EMBO J 5:835–841

SanMiguel P, Bennetzen JL (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann Bot 82:37–44

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274:765–768

Snowden KC, Napoli CA (1998) *Psl*: a novel *Spm*-like transposable element from *Petunia hybrida*. Plant J 14:43–54

Sorek R, Ast G, Graur D (2002) *Alu*-containing exons are alternatively spliced. Genome Res 12:1060–1067

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 Cell Physiol 37:711–716

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22:4673–4680

van Houwelingen A, Souer E, Spelt K, Kloos D, Mol J, Koes R (1998) Analysis of flower pigmentation mutants generated by random transposon mutagenesis in *Petunia hybrida*. Plant J 13:39–50

Varagona MJ, Purugganan M, Wessler SR (1992) Alternative splicing induced by insertion of retrotransposons into the maize waxy gene. Plant Cell 4:811–820

Vicient CM, Kalendar R, Anamthawat-Jonsson K, Schulman AH (1999) Structure, functionality, and evolution of the *BARE-1* retrotransposon of barley. Genetica 107:53–63

Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol 126:485–493

Witte CP, Le QH, Bureau T, Kumar A (2001) Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. Proc Natl Acad Sci USA 98:13778–13783

Witte CP, Tiller S, Isidore E, Davies HV, Taylor MA (2005) Analysis of two alleles of the urease gene from potato: polymorphisms, expression, and extensive alternative splicing of the corresponding mRNA. J Exp Bot 56:91–99