The transposable element *Tip100* and genome rearrangements in the common morning glory

Naoko Ishikawa

DOCTOR OF PHILOSOPHY

2001

Depertment of Molecular Biomechanics
School of Life Science
The Graduate University for Advanced Studies

CONTENTS

General introduction

1. Mobile genetic elements in plants

1.1. Transposable elements and MITEs in plants

Plant transposable elements are mainly divided into two groups. One is the retrotransposon transposed by an RNA intermediate (Grandbastien 1992; Wessler et al. 1995; Kunze et al. 1997; Kumar and Bennetzen 1999; Bennetzen 2000), and the other is the DNA transposable element moved by excision and reintegration (Kunze et al. 1997).

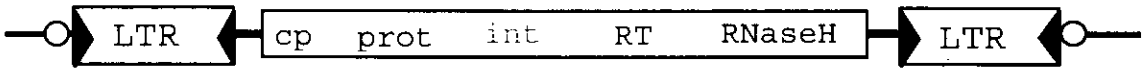
The retrotransposons consist of long terminal repeat (LTR) retrotransposons and non-LTR retrotransposons (Fig.1). The LTR retrotransposons have direct long terminal repeats (LTRs) that can range from a few 100 bp to over 5 kb in size. The LTR retrotransposons are further sub-classified into the two subgroups, *Ty1/copia* and *Ty3/gypsy*, named after their counterparts in *Saccharomyces serevisiae* and *Drosophila melanogaster*, respectively. The characteristic differences between the two subgroups are their degree of sequence similarity and gene order (Fig.1). The non-LTR retrotransposons are classified into long interspersed elements (LINEs) and short interspersed elements (SINEs). The LINEs are structurally simpler than LTR retrotransposons, and both of them carry several genes, which seem to have similar functions (Fig.1).

Fig. 1 General structures of the *Ty1/copia*, *Ty3/gypsy*, LINE, and SINE retrotransposons.

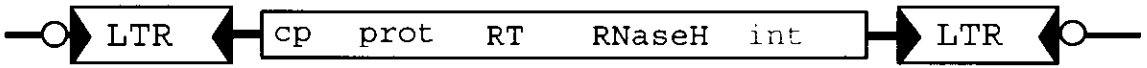
The LTR retrotransposons have long terminal repeats (LTRs) in direct orientation at each end. The LTRs contain signals for initiation and termination of transcription, and are bordered by short inverted repeats (indicated by solid arrows) typically terminating in 5'-TG...CA-3'. The open circles indicate sequences of target site duplications (TSDs). The symbol (A)n within the LINE and SINE boxes indicates poly A sequence. The genes within the retrotransposons encode capsid-like proteins (cp), endonuclease (en), integrase (int), protease (prot), reverse transcriptase (RT), and RNaseH. The two boxes in the SINE box labeled A/B are the regions of homology to the split RNA palymerase III promoters. These figures are not drawn to scale, as the LTR retrotransposons range from a few kb up to 15 kb in size. LINEs usually range in size from less than 1 kb to 8kb, while SINEs are normally 100bp to 300 bp in size.

LTR retrotransposons

***Ty1/copia* group**



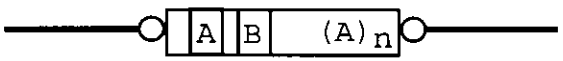
***Ty3/ gypsy* group**



Non-LTR retrotransposons

LINE

SINE



seem to have similar functions (Fig.1). The SINEs are very different from the other class of retrotransposons and they do not encode any *trans*-acting proteins for transposition. All known SINEs are derived from RNA products transcribed by the RNA polymerase III which also produce tRNAs and snRNAs (Kumar and Bennetzen 1999). Unlike genes that are transcribed by RNA polymerase II, the genes transcribed by RNA polymerase III usually contain their promoter sequences within the RNA coding region. Indeed, sequences similar to the promoters for RNA polymerase III can be found within the SINE sequences.

The DNA transposable elements in plants are mainly divided into three families based on their sequences of the terminal inverted repeats (TIRs) and homologies among the transposase proteins: *Ac* family; *En/Spm* family; and *Mu* family (Fig.2). All DNA transposable elements belonging to these families share common structural features that their terminal regions containing sequences required in *cis* for transposition. These *cis*-acting regions are thought to be recognized by *trans*-acting transposases (Gierl et al. 1988; Becker and kunze 1997; Benito et al. 1997).

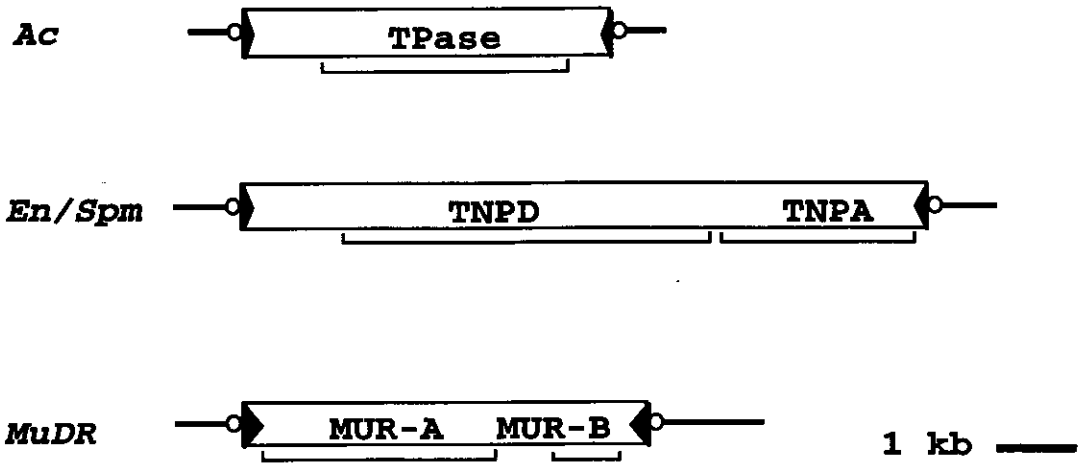
Functionally, transposable elements usually consist of the autonomous and non-autonomous elements (Kunze et al. 1997). The autonomous elements carry both *cis*-elements required for transposition and active genes for *trans*-acting transposases,

Fig. 2 Characteristics of *Ac*, *En/Spm* and *Mu* in maize.

(A) Structures of the *Ac*, *En/Spm* and *MuDR* elements. Filled triangles represent TIRs. The open circles indicate TSDs. Approximate position of the coding regions of the transposase, TNPD, TNPA, MURA and MURD are also shown.

(B) Structural characteristics of the *Ac*, *En/Spm* and *MuDR* elements. TIRs of the *Ac* element are imperfect, therefore both sequences of TIRs are indicated as 5'TIR and 3'TIR. Only 14 bp of 220 bp sequence at the termini is shown for *MuDR*.

A



B

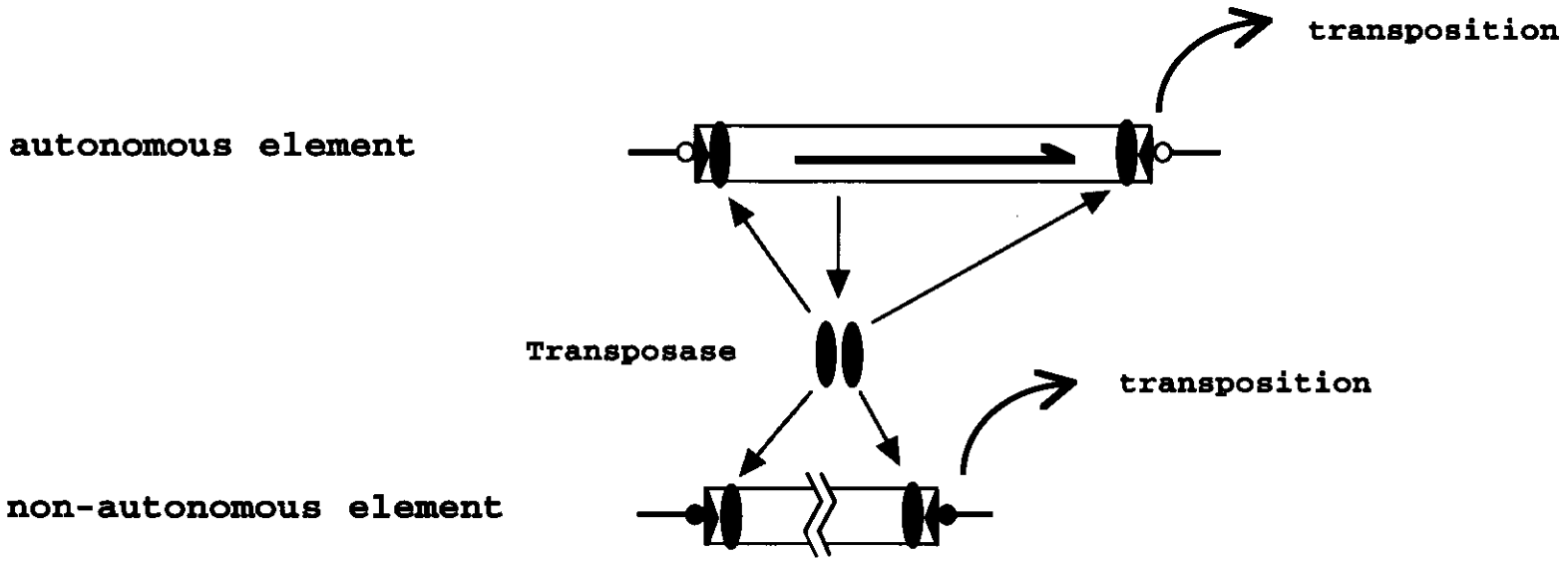
	Ac	*En/Spm*	*MuDR*
size (bp)	4565	8287	4942
length of TSD (bp)	8	3	9
length of TIR (bp)	11	13	~220
TIR sequence	5'TIR 5'-CAGGGATGAAA-3' 3'TIR 5'-TAGGGATGAAA-3'	5'-CACTACAAGAAAA-3'	5'-AGAGATAATTGCCA...-3'

whereas the non-autonomous elements deficient in active transposases can be mobilized only when the active transposases are supplied in *trans* by the autonomous element in the genome (Fig.3). The first transposable element that was discovered by Barbara Mcclintock was the maize element *Dissociation* (*Ds*). She named it after its ability to induce chromosome breakage, which was called dissociation, at the site of its insertion. The chromosome breakage activity requires the presence of another element, which was named *Activator* (*Ac*). Subsequently she noted that *Ac*, as well as *Ds* in the presence of *Ac*, can move from one place to another in the genome. *Ac* is an autonomous element whereas *Ds* elements are non-autonomous.

The DNA transposable elements belonging to the *Ac* family generate 8 bp target site duplications (TSDs) upon integration. Most of the *Ac* family members are less than 4 kb in their sizes (Kempken and Windhofer 2001). *Ac* is the best-characterized element among the *Ac* family characterized so far. The *cis*-required regions of *Ac* consist of imperfect 11 bp TIRs and approximately 240 bp subterminal regions at both ends (Fig.2). The element carries a single gene comprising five exons producing a 112 kDa transposase (TPase). The TPase binds to the subterminal regions, probably in a cooperative manner using a domain located in the C-terminal end of the protein (Becker and Kunze 1997, Essers et al. 2000). It also binds with significantly lower

Fig. 3 The autonomous and non-autonomous elements

The filled triangles at the ends of the boxes represent TIRs. The solid arrow in autonomous element box indicates the transposase gene. The shaded ovals represent the transposase proteins, for transposition, which are supplied by the autonomous element. The open and dotted circles indicate TSDs.


transposition
autonomous element
Transposase
transposition
non-autonomous element

affinity to the TIRs by using a domain located in the N-terminal end (Becker and Kunze 1997). Transposable elements related to the *Ac* family have been identified in most eukaryotic lineges, including plants, fungi and animals (Kempken and Windhofer 2001).

Transposable elements belonging to the *En/Spm* family generate 3 bp TSDs and often called the CACTA element family because their TIRs end with the sequence CACTA (Kunze et al. 1997). The *cis*-required regions of the *En/Spm* element consist of perfect 13 bp TIRs (Fig.2) and 12 bp subterminal motifs, which are scattered within 180 bp at the 5'end and 300 bp at the 3'end. The *En/Spm* element produces two different transposases, TPNA (67kDa) and TPND (131kDa), by relatively complicated differential splicing. Both proteins are shown to be required for transposition (Frey et al. 1990; Masson et al. 1991). TPNA binds to the subterminal motifs (Gierl et al. 1988) and carries both DNA binding and multimerization domains (Trentmann et al. 1993, Raina et al. 1998). The other transposase TNPD has been proposed to bind their TIRs and to catalyze the cleavage reaction for transposition (Frey et al. 1990).

The elements belonging to the *Mu* family generate 9 bp TSDs upon integration. The autonomous element of the *Mu* family, called *MuDR*, was discovered in maize and contains approximately 220-bp TIRs (Chomet et al. 1991). It carries two genes, *mudrA* and *mudrB*, transcribed in convergent manner from promoters located in the

TIRs. The *mudrA* and *mudrB* genes encode proteins of 823 and 207 amino acids, respectively (Fig.2). The *mudrA* gene product, MURA, is required for transposition (Raizada and Walbot 2000) and capable of binding specifically to the TIRs of the *MuDR* element (Benito and Walbot 1997; Raizada and Walbot 2000). It also shows some similarity to the transposase of the *IS256* family of the elements (Eisen et al. 1994). The function of MURB is not well understood. MURB appears not to be required for excisions of element, although it may be involved in its transposition processes (Lisch et al. 1999).

Short interspersed elements of around 100 - 500 bp, named MITEs (Miniature Inverted-Repeat Transposable Elements), are another group of the DNA transposable elements (Wessler et al. 1995; Zhang et al.2001). Majority of characterized MITE families can be divided into two groups, *Stowaway*-like MITEs and *Tourist*-like MITEs, based on similarity of their TIRs and TSDs (Bureau and Wessler 1992; 1994; Wessler et al. 1995; Turcotte 2001).

1.2. Mobile genetic elements in the morning glories

In the common morning glory (*Ipomoea purpurea*) and the Japanese morning glory (*Ipomoea nil* or *Pharbitis nil*), a considerable number of DNA transposable elements were isolated mainly from anthocyanin biosynthesis genes (Table 1; Fig.4).

Fig. 4 A simplified pathway for anthocyanin biosynthesis and mutations for flower pigmentation in the morning glories.

Mutations are indicated in parentheses. Mutants of the Japanese morning glory are indicated by [J] and those of the common morning glory are by [M]. Abbreviations; ANS, anthocyanidin synthase; CHI, chalcone isomerase; CHS, chalcone synthase; DFR, dihydroflavonol 4-reductase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3'-hydroxylase; UF3GT, UDP-glucose; flavonoid 3-O-glycosyltransferase.

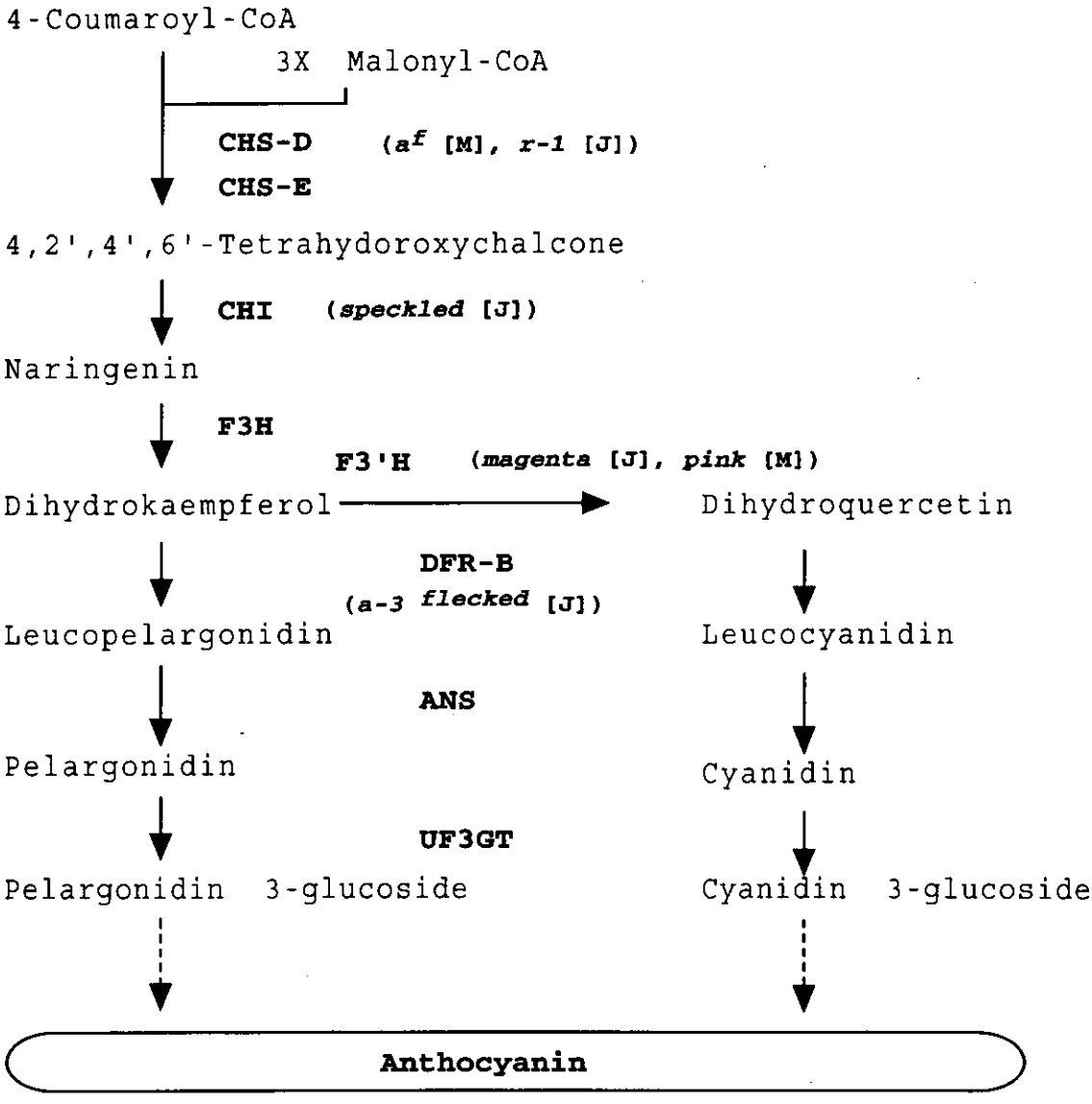

4-Coumaroyl-CoA
3X Malonyl-CoA
CHS-D
CHS-E
(af [M], r-1 [J])
4,2',4',6'-Tetrahydoroxychalcone
CHI
(speckled [J])
Naringenin
F3H
F3'H
(magenta [J], pink [M])
Dihydrokaempferol
Dihydroquercetin
DFR-B
(a-3 flecked [J])
Leucopelargonidin
Leucocyanidin
ANS
Pelargonidin
Cyanidin
UF3GT
Pelargonidin 3-glucoside
Cyanidin 3-glucoside
Anthocyanin

Table 1 DNA transposable elements found within the genes for flower pigmentation in the morning glories

Family	Element	Species	Size(bp)	Gene (mutant allele)	Reference
Ac/Ds					
	Tip100	*I.purpurea*	3873	*CHS-D* (*flaked*)	Habu et al.,1998
	Tip201	*I.purpurea*	512	*F3H* (*blue/pink*)	Morita et al.,1998
En/Spm					
	Tpn1	*I.nil*	6412	*DFR-B* (*flecked*)	Inagaki et al.,1998
	Tpn2	*I.nil*	6500	*CHI* (*speckled*)	Hoshino and Iida, 1997
	Tpn3	*I.nil*	5570	*CHS-D* (*r-1*)	Hoshino and Iida, 1997
	Tpn4	*I.nil*	12kb	*NHX1*(*purpure*)	Fukada-Tanaka et al.,2000
Mu					
	InMu1	*I.nil*	612	*CHI*	Hoshino et al.,2001

For the genes for flower pigmentation, see Fig.4

Two elements, *Tip100* and *Tip201* belonging to the *Ac* family, were obtained from the common morning glory (Habu et al. 1998; Hoshino et al.2001). *Tip100* found within the *CHS-D* gene of the mutable *flaked* line was shown to be transposed actively whereas *Tip201* resides stably in the *F3'H* gene of the *pink* mutant. In the Japanese morning glory, the *Tpn1* element belonging to the *En/Spm* family was isolated from the *DFR-B* gene of the mutable *flecked* line. Subsequently, four other related elements of the *Tpn1* family were obtained (Table 1). The transposable elements related to the *Mu* family have around 220 bp TIRs, and *InMu1* was detected within the *CHI* gene in the Japanese morning glory (Hoshino et al.2001).

At least seven retrotransposons are found in the morning glories (Table 2). The LTR retrotransposon *RTip1* found near 3' of the *ANS* gene of the common morning glory is a defective *Ty3/gypsy*-like element containing two LTR-sequences of about 590 bp (Hisatomi et al.1997a). SINE-like elements have only recently been identified in plants. Most of these SINEs have a region homologous to tRNA sequences (Yoshioka et al.1993, Pozueta-Romero et al.1998). *SINEIp1* (for SINE *I. purpurea*) was found in the *CHS-D* gene in a certain line of the common morning glory (Durbin et al.2001). The split RNA polymerase III promoters were found within the element. Three additional elements related

Table 2 Retrotransposons in the morning glories

Family	Element	Species	Size(bp)	Gene	Reference
(LTR retroposons)					
Ty3/gypsy	*RTip1*	*I.purpurea*	12422	*ANS*	Hisatomi et al.1997
(non-LTR retrotransposons)					
SINE	*SINEIp1*	*I.purpurea*	235	*CHS-D*	Dubin.ML et al.2001
	SINEIp2	*I.purpurea*	246	*CHS-D*	Dubin.ML et al.2001
	SINEIp3	*I.purpurea*	221	*CHS-D*	Dubin.ML et al.2001
	SINEIp4	*I.purpurea* *I.nil*	267	*CHS-D*	Dubin.ML et al.2001
	(*IPSE1*)	*I.purpurea* *I.nil*	-	*CHS-D*	Tanaka M et al.2001

to *SINEIp1* were also detected in the *CHS-D* gene of *I. purpurea* by Durbin et al. (2001), and another SINE sequence were also found in the genome of the morning glories (Tanaka et al. 2001; Table 2).

Mobile element-like sequences (*MELSs*) were characterized in the genomic *DFR* and *CHS* gene regions (Johzuka-Hisatomi et al. 1999, Inagaki et al. 1999, Hoshino et al. 2001, Durbin et al 2001). The structures of *MELSs* are similar to those of MITEs (Wessler et al. 1995). *MELS* identified in the two morning glories are shown in table 3.

1.3. DNA rearrangements mediated by transposable elements

Since the discovery of the transposable elements by Babara McClintock in 1940s, they have been found to be directly and indirectly associated with various genome rearrangements and also known to affect gene expression (Kunze et al. 1997; Fedoroff 2000).

The genome rearrangements such as deletion, inversion, duplication take place as results of the transposition. The rearrangements can also occurred by homologous recombination between the multiple copies of the transposable elements present in the genome (Pichersky 1990; Walker et al. 1995; Gray 2000). In addition, transposable elements can act as a spontaneous

Table 3 *MELSs* in the morning glories

Element	Species	Size(bp)	TIR	Location	Reference
MELS1	*I. nil*	56	11	*DFR-B* 2nd intron	Inagaki et al. 1999
	I. purpurea	57	11	*DFR-B* 2nd intron	Hisatomi et al. 1997b
MELS2	*I. purpurea*	415	23	*DFR-B* and *DFR-C* intergenic	Hisatomi et al.1997b
MELS3-1	*I. nil*	189	51/50	*DFR-A* 5th intron	Inagaki et al. 1999
	I. purpurea	137	51	*DFR-A* 5th intro	Inagaki et al. 1999
MELS3-2	*I. nil*	182	51	*DFR-A* and *DFR-B* intergenic	Inagaki et al. 1999
MELS3-3	*I. nil*	182	51	*CHS-D* 1st intron	Johozuka-Hisatomi et al 1999
MELS3-4	*I. nil*	182	51	*DFR-B* 5th intron	Hoshino et al. 2001
MELSE4	*I. nil*	222	49	*DFR-A* and *DFR-B* intergenic	Inagaki et al. 1999
	I. purpurea	210	49	*DFR-A* and *DFR-B* intergenic	Inagaki et al. 1999
MELS5	*I. nil*	251	13	*DFR-B* 5th intron	Inagaki et al. 1999
MELS6-1	*I. nil*	83	11	*CHS-D* 5' flanking	Johozuka-Hisatomi et al 1999
	I. purpurea	260	11	*CHS-D* 5' flanking	Johozuka-Hisatomi et al 1999
dMELS6-2	*I. purpurea*	220	11	*CHS-D* 1st intron	Johozuka-Hisatomi et al 1999
MELS6-3	*I. nil*	298	11	*CHS-D* 1st intron	Johozuka-Hisatomi et al 1999
	I. purpurea	318	11	*CHS-D* 1st intron	Johozuka-Hisatomi et al 1999
MELS6-4	*I. nil*	274	11	*CHS-E* 3'flanking	Johozuka-Hisatomi et al 1999
	I. purpurea	307	11	*CHS-E* 3'flanking	Johozuka-Hisatomi et al 1999
MELS6-5	*I. purpurea*	286	11	*DFR-A* 1st intron	Johozuka-Hisatomi et al 1999
dMELS6-6	*I. purpurea*	193	11	*DFR-B* and *DFR-C intergenic*	Johozuka-Hisatomi et al 1999
MELS6-7	*I. nil*	320	11	*DFR-A* 5' flanking	Johozuka-Hisatomi et al 1999
MELS7	*I. purpurea*	488	11	*DFR-B 2nd intron*	Hoshino et al. 2001
MELS8	*I. nil*	324	-	*CHS-D* 5' flanking	Durbin et al. 2001
dMELS8	*I. purpurea*	215	-	*CHS-D* 5' flanking	Durbin et al. 2001

mutagen in plant genome by inserting into an active gene, which results in generation of various important floricultural traits in the Japanese and common morning glories (Iida et al. 1999).

1.4. Transposition processes of the DNA transposable elements

The plant DNA elements are thought to be transposed by a "cut-and-paste" mechanism, that is, the element is excised and reintegrates at a new position in the genome (Kunze et al. 1997). The mechanisms of the transposition are extensively studied in the elements belonging to the *Ac* family. During excision of the element, small DNA rearrangements, often called footprints, are generated at the excised site (Fig. 5A-2). For the generation of the footprints, two models have been proposed by Saedler and Nervers (1985) and by Coen et al. (1986), which are described in detail in Chapter I (see Fig. I-6).

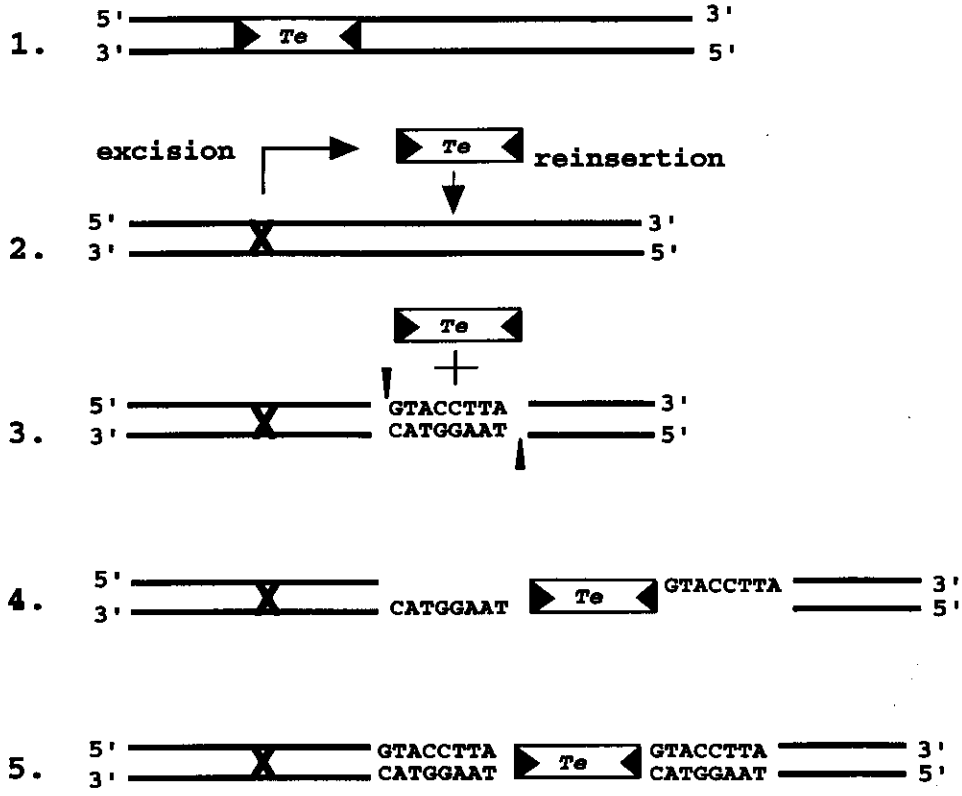
Reinsertion of the transposable element after excision is thought to be a subsequent step of the transposition processes. At the target site staggered cuts are generated (Fig.5A-3), which result in overhanging ends corresponding to the lengths of TSDs characteristic for each family (Fig.5A-4). The excised, presumably flush-ended element is ligated between the single-stranded target site overhangs, followed by fill-in synthesis (Fig.5A-5).

Fig.5 Transposition processes of the DNA transposable elements and the structure of the deletion adjacent to the element.

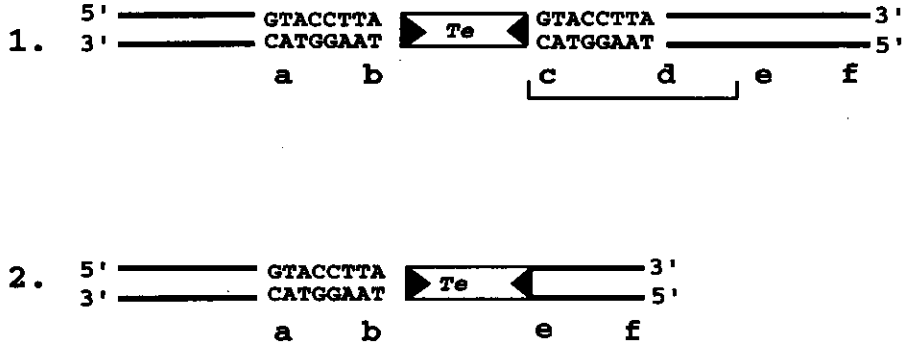
(A) The schematic representation of the process of the transposition. The solid triangles indicate TIRs of the transposable elements (Te) and X indicates the footprints at the excision sites.

(B) The region indicated by bracket in B-1 is deleted in B-2, and note that no sequence alteration occurred at the other side of the element.

A


1.
5'
3'
Te
3'
5'
excision
Te
reinsertion
2.
5'
3'
3'
5'
Te
3.
5'
3'
GTACCTTA
CATGGAAT
3'
5'
4.
5'
3'
CATGGAAT
Te
GTACCTTA
3'
5'
5.
5'
3'
GTACCTTA
CATGGAAT
Te
GTACCTTA
CATGGAAT
3'
5'

B


1.
5'
3'
GTACCTTA
CATGGAAT
Te
GTACCTTA
CATGGAAT
3'
5'
a
b
c
d
e
f
2.
5'
3'
GTACCTTA
CATGGAAT
Te
3'
5'
a
b
e
f

Deletions adjacent to the elements are often observed in various DNA transposable elements including prokaryotic IS elements (Fig.5B). The deletion formation mediated by transposable elements is one of the characteristic features of DNA transposable elements (Iida et al. 1983; Taylor and Walbot 1985; Dooner et al. 1988; Martin et al. 1988; Robbins et al.1989; Lister and Martin 1989; Roberts DE et al.1991; Jilk et al 1993; Gray 2000).

2. Mutable alleles in the common morning glory

2.1. Mutable alleles

The common morning glory was introduced to Europe probably in the 17th century and to Japan in the 18th century. The cultivars with purple, red and white flowers were already recorded in the 18th century (Curtis 1790), and the most extensively characterized mutant displaying white flowers with pigmented spots and sectors was recorded in the early 19th century (Sims 1814). The early genetic studies on the variegated flower line, named *flaked*, were conducted by Barker (1917) in the United States and followed by Imai and Tabuchi (1935) in Japan. More than ten mutable alleles have been described in the common morning (Kasahara 1956), and some of them such as the mutable *flaked* and $yg1^{m}$ alleles for flower and leaf variegation, respectively, can

still be available (Fig.6). New mutable alleles can also be detected (Fig.6). Among these mutable alleles, the *flaked* mutation is the best-studied allele in the common morning glory (Barker 1917; Imai and Tabuchi 1935; Kasahara 1956; Hisatomi et al. 1997a; Habu et al. 1998; Iida et al. 1999; Hoshino et al. 2001).

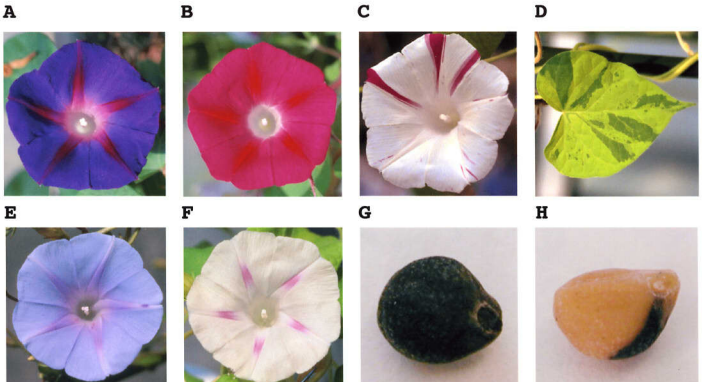
The mutable *flaked* lines bear white flowers with colored flakes and sectors. The variegation is regarded to be caused by recurrent somatic mutation from the white to the pigmented revertant allele. The *flaked* allele at the *A* (*Anthocyanin*) locus, termed *anthocyanin*^{flaked} (a^{flaked}) or a^f, shows incomplete dominance, and the plant in the heterozygous state A/a^f bears lightly colored flowers with intensely colored flakes as well as with occasional white sectors. The a^f mutation is caused by the insertion of a transposable element *Tip100* into the *CHS-D* gene for anthocyanin biosynthesis (Fig.4; Habu et al. 1998).

2.2. Mutable a^* allele

Besides the a^f mutant, the varieties of the common morning glories that show very similar phenotype with a^f mutant were found in the southeastern United States. The common morning glory is thought to be native to Central America and bears dark purple flowers (Ennos et al. 1983; Yoneda 1990; Clegg and Durbin 2000),

Fig.6 Various mutants in the common morning glory.

(A) Flower phenotype of the wild type plant. **(B)** Flower phenotype of a *pink* mutant. The *pink* mutation is caused by insertion of the transposable element *Tip201* into the *F3H* gene (Fig.4). **(C)** Flower phenotype of the a^{f1} allele mutant. The mutant also carries the *pink* allele. **(D)** Leaf phenotype of the ygl^{m} mutant. **(E, F)** Flower phenotypes of plants having new mutable alleles conferring a few fine spots. **(G)** A seed of the wild type plant. **(H)** A variegated seed with unidentified mutable allele.


A
B
C
D
E
F
G
H

and the plant also grows as a common weed in the southeastern United States (Ennos et al. 1983; Simms and Rausher 1987; Clegg and Durbin 2000). Epperson and Clegg (1987) called the mutant, which has similar phenotype to the a^f mutant in the southeastern United States, as a^* mutant plant. It is not clear whether the a^f and a^* alleles have something to do with each other or not.

2.3. Mutable alleles and *Modulator*

Another interesting aspect on the mutable a^f allele is the frequency and timing of the flower variegation vary in different a^f lines. Kasahara (1956) reported a line of the common morning glory showing a variegated leaf pigmentation, *i.e.*, appearance of dark green spots and sectors on a yellow green background (Fig. 6 and see Fig. I-1A). This mutable allele was named *yellow-green leaf*^{mutable}, or ygl^m (previously described as y'). It was further noticed that the timing and frequency of the variegated phenotype in leaves are also generally heritable and that a plant carrying the two alleles, a^f and ygl^m, tends to show a peculiar similarity in both timing and frequency of the variegations in flowers and leaves (see Fig. I-1). Based on these observations, Kasahara (1956) postulated that there must be another genetic element termed *Modulator* (originally termed *Mutator*) acting on both a^f and ygl^m alleles, and that the timing and frequency of the variegations in flowers and leaves are determined by the

heritable state of the *Modulator*. By crossing with the *Modulator* active line, it is possible to distinguish *a* mutant line bearing white flowers due to a stable mutation at the *A* locus from a white flower mutant carrying the a^f allele without active *Modulator*.

In chapter I, I will describe the analysis to elucidate whether *Tip100* in the a^f allele is autonomous element or not. The nature of putative *Modulator* affecting flower and leaf variegations in various mutable lines are also discussed. The structural characteristics of *Tip100* and estimation of *Tip100* copy numbers will be shown. I will also describe if *Tip100* footprints can be explained by the two models proposed by Saedler and Nervers (1985), and Coen et al. (1986). In chapter II, I will show the analysis to reveal the origin of *A* locus mutants and discuss transposable elements associated mutations observed in the *CHS-D* gene of the common morning glory.

Chapter I The transposable element *Tip100* found at the mutable *flaked* allele for flower variegation of the common morning glory is an autonomous element.

1. Introduction

It has been shown that the mutable a^f allele is caused by insertion of the 3.9 kb transposable element *Tip100* into the intron of the *CHS-D* gene encoding chalcone synthase for flower pigmentation (Habu et al. 1998; Shiokawa et al. 2000; Hoshino et al. 2001). *Tip100* in the *CHS-D* intron carries an 11 bp terminal inverted repeats (TIRs) and is flanked by target site duplications (TSDs) of 8 bp. It contains an open reading frame (ORF) encoding a polypeptide of 808 amino acids which exhibits partial similarity to the one of three conserved regions of the transposase of the *Ac* family (Fig. I-1). These structural features are compatible with the notion that *Tip100* may be an autonomous element belonging to the *Ac* family.

Subsequently, it was also found that lightly and heavily variegated flower lines, as well as the white flower line with inactive *Modulator*, all carry a single *Tip100* insertion at the same site in the *CHS-D* intron (Fig. I-2; Habu et al. 1998). In contrast, the stable white flower line with active *Modulator* has two insertions of *Tip100* within the *CHS-D* intron. Excision

Fig. I-1 Comparison of structural features between *Tip100* and other elements of the *Ac* family.

(A) Alignment of TIRs from transposable elements of the *Ac* family. *Tip100* is indicated in red.

(B) Alignment of the best conserved amino acid sequences from members of the *Ac* family. Transposable elements of the *Ac* family have been identified in most eukaryotic lineges, including plants fungi and animals (Calvi BR et al.1991; Kempken and Windhofer 2001). *Tag*I, *Bg*, and *hobo* are isolated from *Arabidopsis thaliana*, maize and *Drosophila melanogaster*, respectively. Dashes indicate gaps introduced to maximum homology.

A

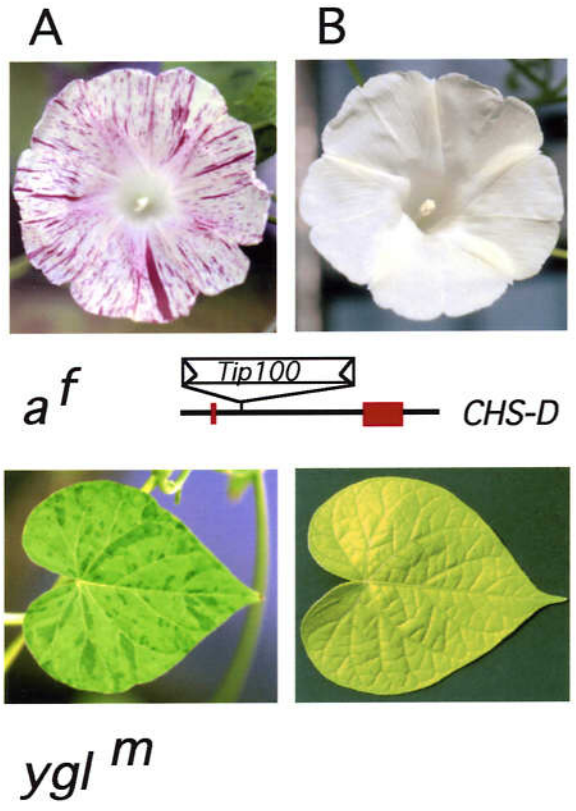
Element	TIR	Species
Ac	CAGGGATGAAA	*Zea mays*
	TAGGGATGAAA	
lpsr	TAGGGGTGGCAA	*Pisum sativum*
Tpc1	TAGGGTGTAAA	*petroselinum crispum*
Tam3	TAAAGATGTGAA	*Antirrhinum majus*
Slid	TAATGCTG	*Nicotiana tabacum*
dTph1	CAGGGGCGGAGC	*Petunia hybrida*
Tip100	CAGGGGCGGAG	*Ipomoea purpurea*
Tst1	CAGGGGCGTAT	*Solanum tuberosum*
	CAGAGGCGTAT	
dTnp1	CAGTGCCGGCTCAA	*Nicotiana plumbaginifolia*

B

```
Tip100   SDLCRWLVKTRKSNIYPLVFRVVTLVLTLPVSTATTERSFSAMNIVKTTLRNK
Tam3     FDILKWW-RQNESL-TPVLARIARDLLSSQMSTVASERAFSAGHRVLTDARNR
Tag1     SQPDEWW-RYFGHD-APNLQKMAIRILSQTASSSGCERNWCVFERIHTKKRNR
Ac       FDILSWW-RGRVAE-YPILTQIARDVLAIQVSTVASESAFSQARLQIGDHRAS
Bg       FFAAEWW-SAYGGE-YKELQMLARRIVSQCLSSSGCERNWSTFALVHTKLRNR
hobo     FEVIEWW-KNNANL-YPQLAKIALKILSIPASSAAAERVFSLAGNIITEKRNR
```

Fig. I-2 Phenotypes of the common morning glories carrying the mutable a^f and ygl^m alleles for flower and leaf variegations, respectively, and the structure of the a^f allele in the *CHS-D* gene.

(A) KK/VR-40 LVS1. **(B)** KK/WR321. The red boxes indicate the exon sequences of the *CHS-D* gene. The mutable ygl^m allele remains to be identified. Note that KK/VR-40 LVS1 is a lightly variegated subline derived from KK/VR-40a displaying heavily variegated phenotypes (Habu et al. 1989).



a f

ygl m

of one of the two copies of *Tip100* from the *CHS-D* gene in the latter stable white flower line appears to be insufficient to restore the *CHS-D* function and probably both elements are rarely excised in the same tissue (Habu et al. 1998).

The simplest explanation for these observations is that *Tip100* and *Modulator* may be a non-autonomous and an autonomous element, respectively. Alternatively, *Tip100* may be an autonomous element and *Modulator* an element which can control the transposition activity of *Tip100*. To examine whether *Tip100* is an autonomous element, *Tip100* and its internal deletion derivative *dTip100* were introduced into tobacco (*Nicotiana tabacum*) to examine whether they are able to transpose in the transgenic tobacco plants. The data presented here strongly indicate that *Tip100* carries the functions necessary for its transposition, thereby represents the first autonomous element characterized in *Ipomoea* species. In addition, I have presented the data showing that a considerable numbers of *Tip100* related elements are present in genome of the common morning glory and majority of them appears to have intact *Tip100* sequence.

2. Materials and methods

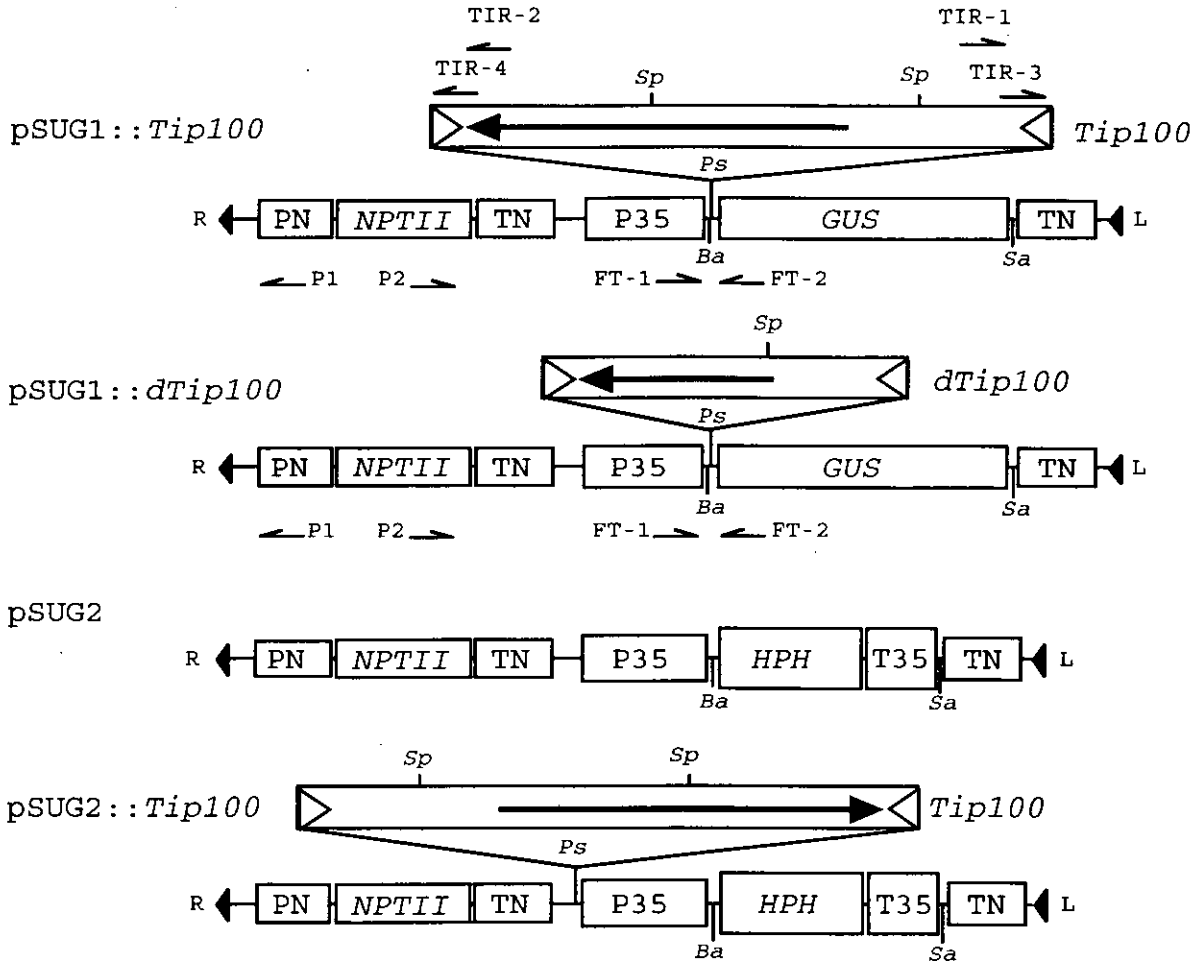
2.1. Plasmid vectors

The *Tip100* element flanked by 8 bp TSDs in the λV-Tip100-4S7 clone (Habu et al. 1998) was amplified by polymerase chain reaction (PCR) and cloned into the *Pst*I site of pHSG398 (TaKaRa Biomedicals) to yield pHSG398::*Tip100*. The primers used for PCR amplification are: Tip100-L (5′-gg**ctgcag***catacgtg* cagggcggaggcac-3′) and Tip100-R (5′-gg**ctgcag***cacgtatg* cagggcggagccaggatta-3′), which contain the *Pst*I site (boldface), the 8 bp TSD sequence (italic) originated from the *CHS-D* intron and the *Tip100* terminal sequences (underlined).

The internal 1.7 kb *Spe*I segment of *Tip100* in pHSG398::*Tip100* was deleted to yield pHSG398::*dTip100*. The plasmid pSUG1 is a derivative of pBI121 (CLONTECH) and carries a unique *Sse*I8387I site at the *Sma*I site of pBI121. To construct pSUG1::*Tip100* and pSUG1::*dTip100* (Fig. I-3), the *Pst*I fragment containing either *Tip100* or *dTip100* was cloned into the *Sse*I8387 site of pSUG1. The *Bam*HI-*Sac*I fragment containing the *GUS* gene for β-glucuronidase in pBI121 was substituted by the *Bam*HI-*Sac*I fragment containing the *HPH* gene for hygromycin B phosphotransferase and the CaMV 35S terminator from pCKR138 (Izawa et al. 1991) to yield pSUG2. For construction of pSUG2::*Tip100*, the *Pst*I fragment containing *Tip100* was cloned

Fig. I-3 Structures of the T-DNA regions in the plasmid vectors.

The filled arrowheads with L or R represent the left and the right border of T-DNA, respectively. The open triangles and the thick horizontal arrows within the *Tip100* boxes indicate TIRs and ORFs, respectively. The boxes with PN and TN indicate the promoter and the terminator of the nopaline synthase gene, respectively, and the boxes with P35 and T35 represent the CaMV 35S promoter and the CaMV 35S terminator, respectively. The structural genes for Km^r, GUS and Hm^r are indicated by *NPTII*, *GUS* and *HPH*, respectively. The small horizontal arrows indicate the positions of primers used for PCR or IPCR amplification. Note that both *Tip100* and *dTip100* on the pSUG1 derivatives are flanked by the 8 bp TSD originated from the *flaked* allele of the *CHS-D* gene (Habu et al. 1998). Restriction sites are: *Ba*, *Bam*HI; *Ps*, *Pst*I; *Sa*, *Sac*I; *Sp*, *Spe*I.


pSUG1::Tip100
TIR-2
TIR-1
TIR-4
TIR-3
Sp
Sp
Tip100
Ps
R
PN
NPTII
TN
P35
GUS
TN
L
Ba
Sa
P1
P2
FT-1
FT-2
pSUG1::dTip100
Sp
dTip100
Ps
R
PN
NPTII
TN
P35
GUS
TN
L
Ba
Sa
P1
P2
FT-1
FT-2
pSUG2
R
PN
NPTII
TN
P35
HPH
T35
TN
L
Ba
Sa
pSUG2::Tip100
Sp
Sp
Tip100
Ps
R
PN
NPTII
TN
P35
HPH
T35
TN
L
Ba
Sa

into the *Sse*8387I site of pSUG2.

2.2. Transgenic plants

Agrobacterium mediated transformation of the plasmid vectors pSUG1::*Tip100* and pSUG1::*dTip100* into *Nicotiana tabacum* cv SR1 plants was performed as previously described (Sugita et al. 1999). After co-cultivation of leaf segments with *A. tumefaciens* LBA4404 (Hoekema et al. 1983) containing an appropriate plasmid on solid hormone-free MS medium containing 50 mg/l of acetosyringone for a 3-day, the explants were transferred to solid SIM medium (MS medium with 1 mg/l of BA, 0.1 mg/l of NAA and 100 mg/l of kanamycin) containing 500 mg/l of ticarcillin to control further bacterial growth. One month after of *Agrobacterium* infection, regenerated adventitious buds conferring kanamycin resistance (Km^r) were removed and transferred to the hormone-free MS medium containing 100 mg/l of kanamycin and 500 mg/l of ticarcillin.

To estimate the integrated T-DNA copy number in transgenic tobacco plants, I employed the inverse PCR (IPCR) procedure using the primers P1 (5'-CGTTGCGGTTCTGTCAGTTCC-3') and P2 (5'-TTGTCAAGACCGACCTGTCC-3') for determining the copy number of the integrated *NPTII* gene for neomycin phosphotransferase (Fig. I-3; Does et al. 1991). To examine the excision of *Tip100* from the introduced *GUS* gene, I have carried out a phenotypic assay for the *GUS* expression in leaves of the transgenic tobacco plants by the GUS staining as previously described (Jefferson et al.

1987). To determine the footprint sequences generated by *Tip100* excision (Fig. I-3), the fragments produced by PCR amplification with the primers FT-1 (5'-ACAATCCCACTATCCTTCGC-3') and FT-2 (5'-GGATAGTCTGCCAGTTCAGT-3') were cloned into the pGEM T-easy plasmid (Promega) and subsequently sequenced with the primer GUS-2 (5'-TCGCGATCCAGACTGAATGCCC-3').

2.3. Excision of *dTip100* promoted by *Tip100*

Agrobacterium mediated transformation of pSUG2::*Tip100* or pSUG2 into the transgenic dTip1 plant carrying one copy of the *NPTII* gene (see Table 1B) was performed in the same way as the transformation of pSUG1::*Tip100* described above, except that hygromycin (50 mg/l) was used instead of kanamycin (100 mg/l) for selection and that the regeneration step with hormone-free MS medium was omitted. The hygromycin resistance (Hm^r) calli were used for further analyses.

2.4. Reintegration of *Tip100*

To determine the sequences at the reintegration sites of *Tip100*, I employed first IPCR amplification using the primers TIR-1 (5'-GGCCAAGCCGCCAAGGCCCCTATTGCCTAATAGGC) and TIR-2 (5'-CGTGACATGCAAGAGAGACGGTCTAAATTTTAGT-3') and subsequently nested PCR amplification with the primers TIR-3 (5'-GGGCCAAGGCCCCCTGCATATAATATGTGC-3') and TIR-4 (5'-TTTCAGTATTCTGTCTTAAATTGTGATT-3'). The positions of these

primers within *Tip100* are shown in Fig. I-3. Since *Tip100* contains no *Kpn*I site, the genomic DNAs extracted from leaves of the transgenic plants to be examined were digested with *Kpn*I. After circularizing the *Kpn*I fragments by self-ligation, DNA samples were treated with *Pst*I to destroy the amplified DNA fragments originated from the integrated pSUG1::*Tip100* sequences where the non-transposed *Tip100* element is flanked by the *Pst*I sequences used for construction of pSUG1::*Tip100*. After the first PCR reaction with the primers TIR-1 and TIR-2, the resulting reaction mixture was diluted 100-fold and subjected to the subsequent nested PCR amplification with the primers TIR-3 and TIR-4. The amplified fragments were cloned into pGEM T-easy and the TSD sequences were determined.

2.5. Southern hybridization

Plant DNA was extracted with Qiagen genomic Tip (QIAGEN). Standard hybridization conditions are; 0.5 M Church-buffer (Church and Gilbert 1984) at 65°C for approximately 12 h. After hybridization, the Hybond-N membranes (Amaersham Life Science) were successively washed with 40 mM Church-buffer (Church and Gilbert 1984) at room temperature for 5 minuets for three times and at 65°C for 25 minuets.

3. Results and discussion

3.1. A phenotypic assay for the excision of *Tip100* from the introduced *GUS* gene in transgenic tobacco plants

To examine whether *Tip100* is an autonomous element, I introduced *Tip100* into tobacco plants and examined whether *Tip100* can be excised from the introduced *GUS* gene in the transgenic tobacco plants. I also employed the *dTip100* element having the ORF of *Tip100* disrupted by deleting the 1.7 kb internal *Spe*I segment, as a control defective element (Fig. I-3). Since the inserted *Tip100* and *dTip100* elements efficiently block the *GUS* expression in pSUG1::*Tip100* and pSUG1::*dTip100*, respectively, the *GUS* gene is activated when the elements have excised. *Tip100* in pSUG1::*Tip100* and *dTip100* in pSUG1::*dTip100* were introduced into tobacco by *Agrobacterium* mediated transformation, and Km^r transgenic plants were obtained. Most of these primary transformants (T0) were found to contain only one copy of the integrated *NPTII* gene in their genomes while a few of them carried two or three copies of the *NPTII* gene (Table I-1).

Excision of *Tip100* or *dTip100* from the original *GUS* gene was examined in leaves from these T0 plants by assaying for GUS activity. As Fig. I-4A shows, the leaves of all twelve independent T0 plants with pSUG1::*Tip100* examined displayed rare but clear

Table I-1 Footprint sequences generated by *Tip100* excision.

(A) Footprint sequences found in the 12 transgenic plants transformed with pSUG1::*Tip100*. The copy numbers of the *NPTII* gene in the transgenic plants are indicated (N.D., not determined). The original sequence shown at the top of the sequences represents the flanking sequences of *Tip100* in pSUG1::*Tip100*. TSDs and their flanking sequences are shown in bold upper case and lower case letters, respectively. The dashes represent the nucleotides deleted, and newly generated sequences within the footprints are also shown in smaller bold upper case letters. The structure of Footprint 16 is drawn in Fig. I-5C. From each transgenic plant, at least 10 clones containing footprints were sequenced. The numerals under the transgenic plants represent the numbers of footprints found in each of the plants indicated. N and T denote the total numbers of clones and transgenic plants carrying the given footprint, respectively.

(B) Footprint sequences found in the 3 Hm^r calli derived from the transgenic plant dTip1 transformed with pSUG2::*Tip100*. The original sequence shown at the top of the sequences represents the flanking sequences of *dTip100* in pSUG1::*dTip100*. N and T denote the total numbers of clones and Hm^r transformants carrying the given footprint, respectively. The copy numbers of the *NPTII* gene in two out of three transgenic plants transformed with pSUG1::*dTip100* are also shown. The other symbols are as in **(A)**.

Table I-1 Footprint sequences generated by *Tip100* excision.

A

Sequences		Number of footprints found.													
	Transgenic Plants	Tip 1	Tip 2	Tip 3	Tip 4	Tip 5	Tip 6	Tip 7	Tip 8	Tip 9	Tip 10	Tip 11	Tip 12		
	Copy number of T-DNA (*NPTII*)	3	1	1	1	1	1	3	2	1	N.D.	1	1	N	T
Original	ccccctgcag**CACGTATG** : ***Tip100*** :	**CACGTATG**ctgcag													
Footprint 1	ccccctgcag**CACGTAT-**	**CACGTATG**ctgcag	3	4	5	20	8	7	5	3	8	5	1	7	76	12
2	ccccctgcag**CACGTATG**	**-ACGTATG**ctgcag	3	6	1			1	2			4	16	3	36	8
3	ccccctgcag--------	**CACGTATG**ctgcag	3	2	1		2			2	1				11	6
4	ccccctgcag**CACGTA--**	**----TATG**ctgcag							2						2	1
5	ccccctgcag**CACGTATG**	**----------**cag						2		1					3	2
6	c----------------	**----------**----g						2					1		3	2
7	ccccctgcag**CACGTA--**	**-ACGTATG**ctgcag										1			1	1
8	ccccctgcag**CACGTA--**	**---------**tgcag									1				1	1
9	ccccctgcag**CACGT---**	**-ACGTATG**ctgcag	1												1	1
10	ccccctgcag**CACGTATG**	**CACGTATG**ctgcag											1		1	1
11	ccccctgcag**CACGTA--** **GTG**	**-ACGTATG**ctgcag					1		1						2	2
12	ccccctgcag**CACG----** **GTG**	**-ACGTATG**ctgcag			2										2	1
13	ccccctgcag**CACGTAT-** **CAT**	**--CGTATG**ctgcag						1							1	1
14	ccccctgcag**CACGTAT-** **CA**	**-----ATG**ctgcag			1			1							2	2
15	ccccctgcag**CACGTATG** **ATACGTGCTGCAGGTG**	**-ACGTATG**ctgcag								8		1			9	2
16	ccccctgcag**CACGTA--** **CG(*ΔTip100*)**	**CACGTATG**ctgcag			2										2	1
Number of clones sequenced.			10	12	12	20	11	14	10	14	10	11	19	10	153	

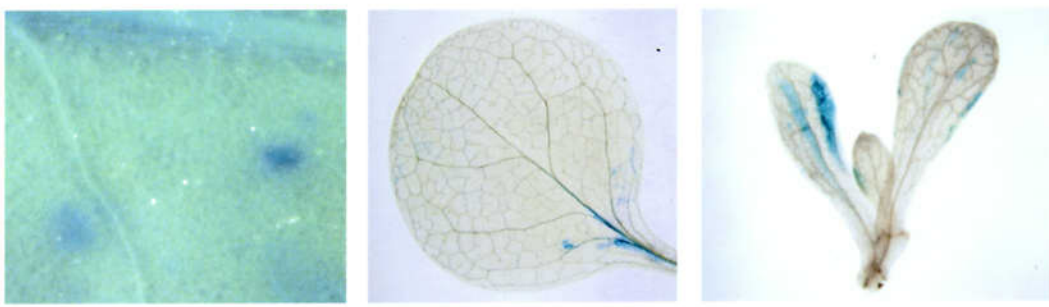
B

Sequences		Transgenic Plants	dTip 1				dTip 2	dTip 3	
		Copy number of T-DNA (*NPTII*)	1				2	N.D.	
			Number of footprints found.						
		Hm^r calli	dTip 1-1	dTip 1-2	dTip 1-3		N	T	
Original	ccccctgcag**CACGTATG** : ***dTip100*** :	**CACGTATG**ctgcag							
Footprint d1	ccccctgcag**CACGTAT-**	**CACGTATG**ctgcag	2	4	2		8	3	
d2	ccccctgcag**CACGTATG**	**-ACGTATG**ctgcag	1	1			2	2	
d3	ccccctgcag--------	**CACGTATG**ctgcag		1			1	1	
Number of clones sequenced.			3	6	2		11		

Fig. I-4 GUS-positive patches observed in leaves of transgenic tobacco plants introduced with pSUG1::*Tip100*.

(A) T0 plant; **(B, C)** T1 plants

A B C



and fine GUS-positive blue patches when stained with X-Gluc (5-bromo-chloro-3-indolyl-beta-D-glucuronide), indicating that *Tip100* excision from the *GUS* gene occurred in the transgenic tobacco plants. I noticed that *Tip100* excision tends to occur much more frequently in the plants at a younger stage, and no apparent GUS-positive blue spots can be seen in a later stage of the same plants. A comparative observation that stabilization of the introduced autonomous *Tam3* element belonging to the *Ac/Ds* family had occurred in transgenic tobacco plants was reported (Martin et al. 1989). However, GUS-positive spots or sectors could be detected in the selfed T1 progeny derived from the T0 plants (Fig. I-4B, C). None of the three transformants with pSUG1::*dTip100* showed any GUS-positive spot at all (data not shown). The results indicate that *Tip100* can be excised from the introduced *GUS* gene in transgenic tobacco plants whereas *dTip100* can not and that *Tip100* is likely to be an autonomous element coding an active transposase of 808 amino acids.

3.2. A phenotypic assay for the excision of *dTip100* from the introduced *GUS* gene promoted by *Tip100*

I asked whether *dTip100* deficient in the active tranposase gene due to the internal deletion of 1.7 kb could be excised from the *GUS* gene when the intact *Tip100* element was introduced into the cells containing the defective *dTip100* element at the *GUS* gene in the genome. To examine this possibility, I chose the T0

tobacco plant dTip1 that appears to carry one copy of the *dTip100* element at the *GUS* gene in the genome (Table I-1B) and *Tip100* in pSUG2::*Tip100* was introduced into the dTip1 plant by *Agrobacterium* mediated transformation. All of the six independently obtained Hm^r calli stained with X-Gluc displayed a few but significant number of GUS-positive blue cells. No such blue cells could be detected in the Hm^r calli obtained by introduction with the control vector pSUG2. The results further support the notion that *Tip100* is an autonomous element and is able to act on the non-autonomous *dTip100* element in the heterologous tobacco plants.

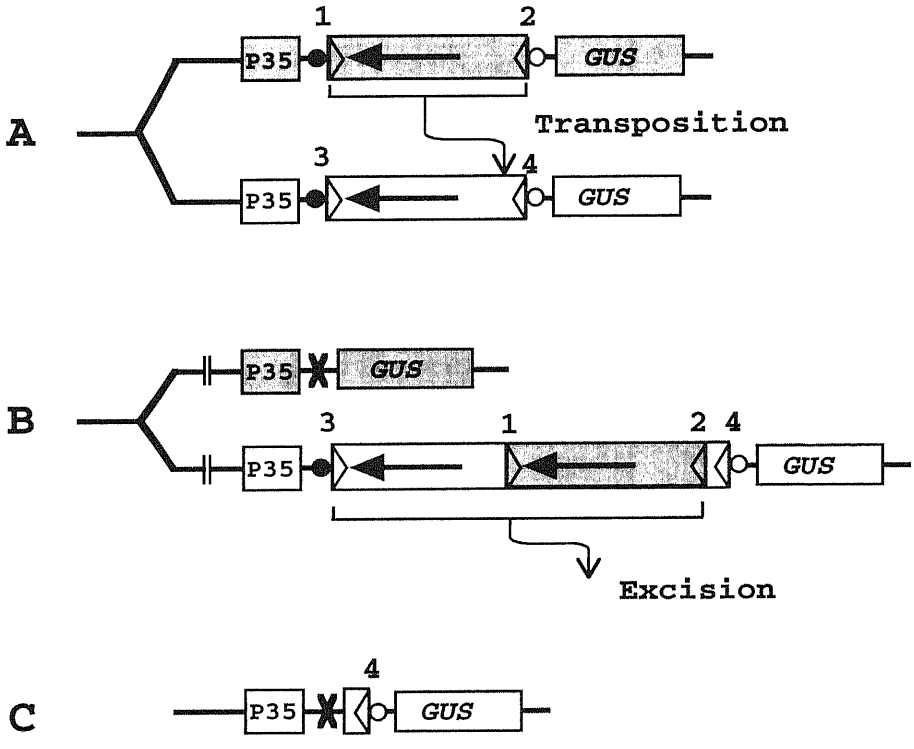
3.3. Footprints generated by excision of *Tip100* and *dTip100*

One of the characteristics of plant DNA transposable elements including the *Ac* family is their ability to generate small sequence alterations at excision sites that are often called footprints (Kunze et al. 1997). To obtain further evidence for transposition of the introduced *Tip100* element in tobacco, I determined the sequences of the *Tip100* excision sites to examine whether footprints had been generated. The 450 bp fragments containing the *Tip100* excision site were PCR-amplified and cloned into plasmids. All 12 T0 plants with pSUG1::*Tip100* displaying the GUS-positive patches gave the PCR-amplified fragments while no amplification was detected in the transgenic plants with pSUG1::*dTip100*. Since these amplified fragments are

likely to contain different footprints resulting from independent *Tip100* excisions, I determined a total of 153 excision sequences consisting of at least 10 excision sequences from each of the 12 different transgenic plants obtained (Table I-1A). Sixteen different sequences were detected and fifteen of them showed the sequences characteristic of footprints. Of these, two most frequently occurred sequences, Footprint 1 and Footprint 2, have one bp deletion at either end of the original 8 bp TSD. Since *Tip100* in pSUG1::*Tip100* is flanked by the 8 bp TSD originated from the *CHS-D* intron in the common morning glory (see Materials and methods), these two footprints were previously observed in *Tip100* excision site at the *flaked* allele of the *CHS-D* gene (Habu et al. 1998). Footprint 16 contains one intact TSD together with the 5' end of the 264 bp *Tip100* segment linked to the other end of the TSD with 2 bp deletion, and the remaining 3.7 kb *Tip100* sequence including its 3' end has also been deleted (Fig. I-5C). A possible model to generate such a structure is shown in Fig. I-5.

To examine whether the footprints were also generated at the excision site of the *dTip100* element promoted by *Tip100* in the Hm^r calli derived from the dTip1 plant, I chose three independent GUS-positive calli introduced with pSUG2::*Tip100*. All of them

Fig. I-5 A possible model for the formation of Footprint 16. The open and filled circles indicate the 8 bp TSD at the 5′ and 3′ ends of *Tip100*, respectively (Table I-1). To distinguish the ends of *Tip100*, the open triangles for TIRs are numbered. Footprints generated by excisions of *Tip100* and its derivative are indicated by X. The other symbols are as in Fig. I-3. **(A)** After the replication, *Tip100* is transposed into *Tip100* on the other daughter strand. **(B)** Subsequently, excision of the composite *Tip100* element with TIRs 2 and 3 results in the complicated Footprint 16 shown in **(C)**.



Footprint 16

produced the 450 bp fragments containing *dTip100* excision sites. I determined 11 footprints from these GUS-positive calli (Table I-1B) and found no significant difference in the footprints generated by excision of *Tip100* from those of *dTip100*. Therefore, I can conclude that excision of *Tip100* generates footprints characteristic of plant DNA transposable elements belonging to the *Ac/Ds* family. No amplification of the 450 bp fragments was detected in the Hm^r calli transformed with pSUG2.

3.4. Assessments of the two alternative models explaining the generation of footprints based on the footprint sequences produced by excision of *Tip100* and *dTip100* from the introduced the *GUS* gene

As Table I-1 shows, I have determined 16 different footprints generated by the excision of *Tip100* and *dTip100* from the introduced the *GUS* gene. Except for Footprint 16, the remaining 15 footprints were considered to be generated by simple excision of *Tip100* and *dTip100*. I thus tried to assess the two alternative models proposed by Saedler and Nervers (1985) and by Coen et al. (1986) by employing these 15 footprint sequences generated by the simple excision.

In the former model proposed by Saedler and Nervers (1985), the initial step of the excision reaction is the creation of 8 bp staggered nicks precisely at the ends of TSDs (Fig. I-

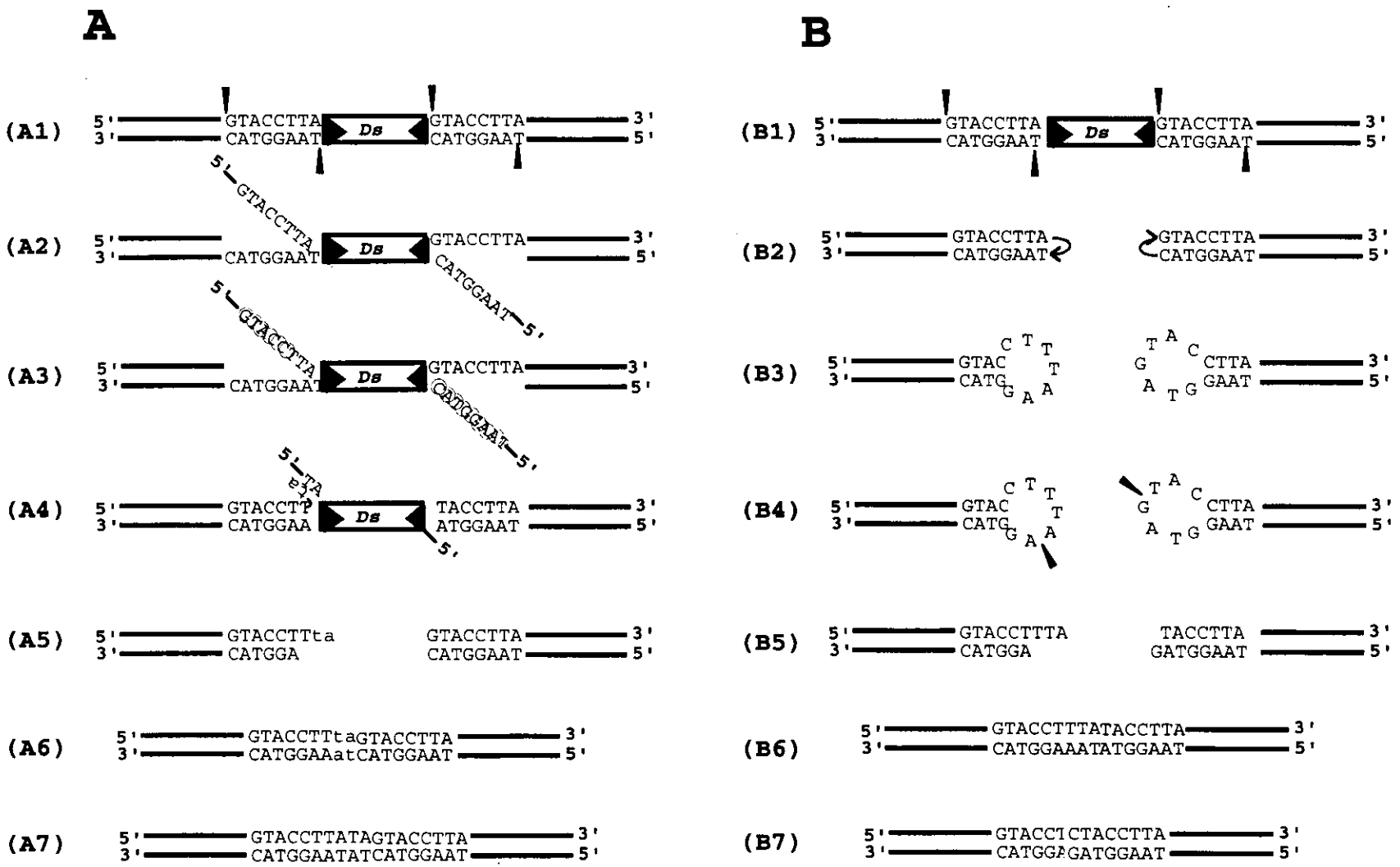
6A). In the second step, the single strand overhangs in a protruding from the flanking DNA are simultaneously subject to repair fill-in synthesis and exonucleolytic attack. When the filling-in processes reach into the end of the template, the DNA polymerase occasionally switches templates and continues synthesis using the overhang attached to the element end as template. Finally, the transposable element is released from the complex and the flanking DNA ends are joined and repaired using one or the other strand as template. Depending on exactly when and where DNA polymerase and exonuclease operate during excision, multiple different transposon footprints will be the outcome. If the repair fill-in synthesis and exonucleolytic attack had occurred only at the single strand overhangs in a protruding from, deletions of the former flanking sequences should take place only within the former 8 bp TSD sequences and small insertions between the former TSD sequences should also be less than 8 bp in length. In these criteria, Footprints 5, 6, 8 and 15 are not well fitted in with the predictions.

In the latter "hairpin model" proposed by Coen et al. (1986), single-strand cuts staggered by only one residue are introduced at the ends of transposable element, producing flanking DNA each having a one base overhang (Fig.I-6B). Subsequently, the free termini at each flanking DNA end are ligated to form a hairpin

Fig.I-6 Models for the transposition mechanism of plant transposable elements.

(A) Excision mechanism according to Saedler and Nevers (1985). At the target site duplications staggered incisions are generated **(A1)**. The overhanging free 5'-ends are simultaneously subjected to exonucleolitic degradation (indicated by encircled bases) and to fill-in repair synthesis **(A2)-(A3)**. Occasionally, the DNA polymerase may switch templates from the excision site overhang to the transposon overhang **(A4)**. After removal of the excised transposon **(A5)**, the chromosome ends are religated and repaired, yielding different possible outcomes **(A6)-(A7)**.

(B) "Hairpin" excision model according to Coen et al. (1986). In the initial step, 1 bp-staggered cuts are generated at the transposon ends **(B1)**. The 5'- and 3'-ends of each chromosome end are ligated to form hairpins **(B2)-(B3)**. Opening of the hairpins by endonucleolitic cleavage can occur at different positions **(B4)**, followed by religation and repair **(B5)**, resulting different possible outcomes **(B6)-(B7)**.


A
B
(A1)
(A2)
(A3)
(A4)
(A5)
(A6)
(A7)
(B1)
(B2)
(B3)
(B4)
(B5)
(B6)
(B7)
Ds
GTACCTTA
CATGGAAT
GTACCTTaGTACCTTA
CATGGAAatCATGGAAT
GTACCTTATAGTACCTTA
CATGGAATATCATGGAAT
GTACCTTTATACCTTA
CATGGAAATATGGAAT

structure (Peacock et al. 1984; Coen et al. 1986). The hairpins are resolved by endonucleolytic nicking, which can occur at different positions within the hairpin sequences, followed by religation of the flanking DNA. Replication of the resulting heteroduplex DNA would give two reciprocal products, a deletion and an inverted duplication. Alternatively, the heteroduplex might be corrected by DNA repair. According to this model, there would be no problem to explain the generation of Footprints 5, 6, 8 and 15. However, it would not be easy to explain the formation of Footprints 10, although this footprint sequence can be easily explained by the model proposed by Saedler and Nervers (1985). Apart from these sequences, Footprints 5, 6, 8, 10 and 15, the remaining 10 footprint sequences can be equally explained by both models. I would like to emphasize here that these 5 footprint sequences could have also been explained by both models, if appropriate modifications were allowed to be introduced into the original models.

3.5. Reintegration of *Tip100* into the tobacco genome

To examine whether the excised *Tip100* elements were reintegrated into new sites in the genomes of the transgenic tobacco plants, I characterized the flanking sequences of *Tip100* insertion sites different from the original site at the *GUS* gene. Since the *Tip100* excision occurred at rather low frequencies,

the reintegration of *Tip100* was also expected to be relatively rare events. I thus applied the IPCR procedure followed by the nested PCR amplification technique with appropriate primers (Fig. I-3) and were able to identify eight different reintegration sites from four transgenic plants among the selfed T1 progeny derived from the GUS-positive T0 plants (Fig. I-7). As expected, *Tip100* generates 8 bp TSDs upon integration. Based on the results obtained in transgenic tobacco plants, I have concluded that *Tip100* is an autonomous element coding the active transposase of 808 amino acids. Apart from the autonomous *Ac* element (Baker et al. 1986), other plant transposable elements of the *Ac* family including *Tam3* of snapdragon and *Tag1* of Arabidopsis were regarded to be autonomous since they were capable of somatic transposition in transgenic tobacco plants (Haring et al. 1989; 1991; Martin et al. 1989; Frank et al. 1997). To our knowledge, *Tip100* is the first autonomous element characterized among various transposable elements in *Ipomoea* species (Hoshino et al. 2001).

3.6. Copy numbers of *Tip100* related elements in the genome

Since the *Tip100* element found within the *CHS-D* gene of the mutable a^f line is an automonous element, I have interested in

Fig. I-7 Sequence analysis of *Tip100* reinsertion sites in transgenic tobacco plants.

Newly generated TSDs and their flanking sequences are shown in bold upper case and lower case letters, respectively.

tcttatta**GGCAAGAC** : ***Tip100*** : **GGCAAGAC**accatggg

aggtcatg**GGTTTGGC** : ***Tip100*** : **GGTTTGGC**ttttgggt

tgaacatc**GTGACGGG** : ***Tip100*** : **GTGACGGG**ttgagcag

atttaatt**CTCACGGC** : ***Tip100*** : **CTCACGGC**ttaaaccc

aaatgtgc**CATGTGCT** : ***Tip100*** : **CATGTGCT**gaaaggaa

ctacagcc**CTTAAGAC** : ***Tip100*** : **CTTAAGAC**gaacatgc

tcaaagtc**TTCGTCAT** : ***Tip100*** : **TTCGTCAT**ccactgct

tcttgagg**CTGCTGTG** : ***Tip100*** : **CTGCTGTG**agagaatg

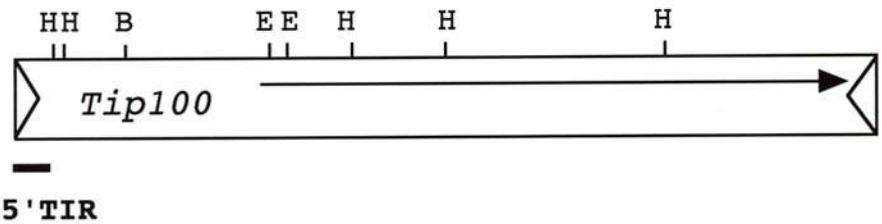
characterizing *Tip100* and its relatives in the genome of the common morning glory. To estimate the copy number of *Tip100* and its related sequences, I have performed Southern hybridization analysis with the 118 bp at the 5′ end of the *Tip100* sequence as a probe. The results indicated that a number of *Tip100*-like sequences are present in the genomes of a line of the common morning glory (Fig. I-8). Based on the hybridization patterns, I have estimated that the *Tip100*-related sequences are present approximately 100 copies per haploid genome. To estimate whether the autonomous elements occupy a large portion of the *Tip100*-related sequences detected in Fig. I-8, the genomic DNAs were cleaved with *Eco*RV that dose not cut the *Tip100* sequence in the a^f allele and hybridized with the five different probes indicated in Fig.I-9A. As Fig.I-9A shows, a significant portion of the *Tip100*-related sequences in the genome of the plant appears to be the intact *Tip100* element. Indeed, I have found that another *Tip100* element in a variety of the common morning glory. Sequencing analysis of the newly found *Tip100* and its flanking sequences revealed that it is identical in sequence to *Tip100* previously characterized in the mutable a^f allele and that it is flanked by 8 bp TSDs.

Fig. I-8 Copy numbers of *Tip100*-related elements in the genome of the common morning glory.

(A) Restriction map of *Tip100*. Resriction sites are: H, *Hin*dIII; B, *Bam*HI; E, *Eco*RI; S, *Sma*I. The thick bar under the *Tip100* box indicates the position and designation of a probe used for the analysis.

(B) Southern analysis with the probes indicated in **(A)**. Gnomic DNA were digested with indicated restriction enzymes and separated on 1.5% agarose gels. Size of hybridising fragments is in kilobases.

A



B

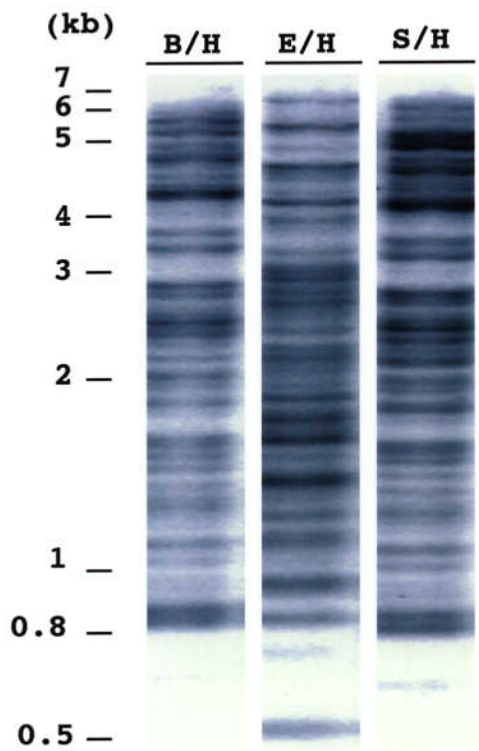
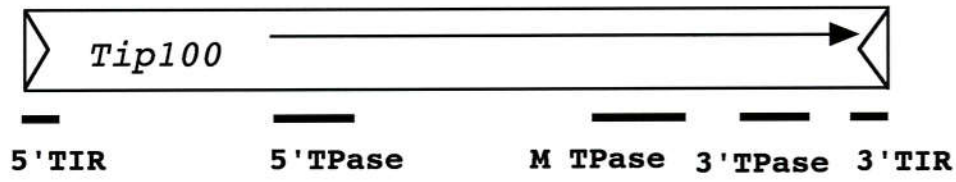


Fig. I-9 Structures of *Tip100*-related elements in the genome of the common morning glory.

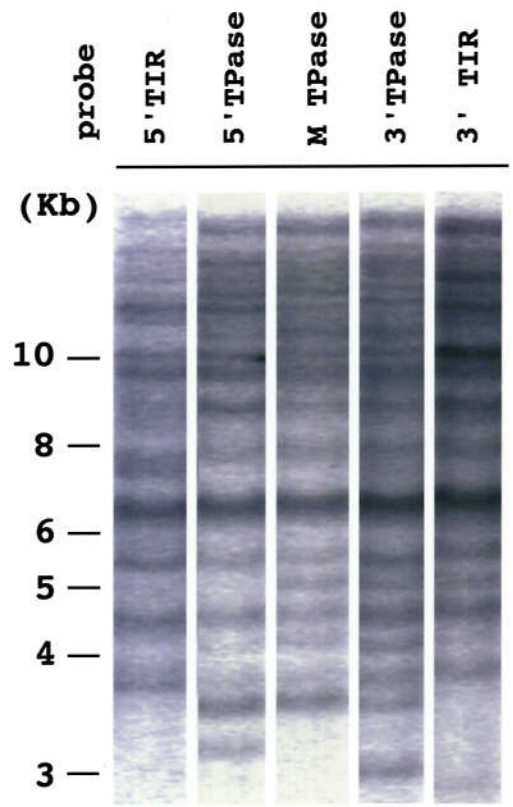
(A) The probes used for Southern analysis. The thick bars under the *Tip100* box indicate the positions of probes and their designations.

(B) Southern analysis with the indicated probes. Gnomic DNA were digested with *Eco*RV, which do not cleave *Tip100* in the *CHS-D* gene, and separated on 0.8% agalose gels. Size of hybridizing fragments is in kilobases.

A


Tip100
5'TIR
5'TPase
M TPase
3'TPase
3'TIR

B


probe
5'TIR
5'TPase
M TPase
3'TPase
3' TIR
(Kb)
10
8
6
5
4
3

3.7. *Tip100* and *Modulator* in the common morning glory

Based on the observations that lines of the common morning glory carrying the two alleles, a^f and ygl^m, tend to show a peculiar similarity in both timing and frequency of the variegations in flowers and leaves and that these variegated phenotypes are also generally heritable by their progeny, Kasahara (1956) postulated that there must be *Modulator* acting on both a^f and ygl^m alleles and that the timing and the frequency of the variegated phenotypes in flowers and leaves are determined by the state of the *Modulator* activity which is generally heritable. Indeed, the *CHS-D* gene of the stable white flower mutant carrying the mutable a^f allele without active *Modulator* is identical to that of mutable a^f lines displaying variegated flowers (Fig. I-2), whereas the white flower line carrying the stable *a* allele was found to have two copies of *Tip100* integrated into the *CHS-D* gene (Habu et al. 1998). Although the simplest assumption for these observations is that *Tip100* and *Modulator* are a non-autonomous and an autonomous element, respectively, our present finding that *Tip100* is an autonomous element are clearly incompatible with this assumption.

The genome of the common morning glory contains about 100 copies of *Tip100* related elements and more than half of them appear to be structurally very similar to *Tip100*, suggesting that a significant number of autonomous *Tip100* copies are present in the genome of *I. purpurea*. A comparable observation that most

copies of *Tam3* in the genome of snapdragon have highly conserved structures of nearly the same size has been reported (Kishima et al. 1999). Interestingly, both *Tip100* and *Tam3* carry single ORFs for active transposase genes without intron sequence (Hehl et al. 1991; Habu et al. 1998).

A possible hypothesis is that *Modulator* may be an element affecting the transposition activity of the autonomous *Tip100* element. In the *Ac/Ds* family of elements, *Stabilizer* of snapdragon and the *IAE* loci of Arabidopsis are reported to control the transposition activity of the autonomous elements, *Tam3* and *Ac*, respectively (Carpenter *et al.* 1987; Jarvis et al. 1997). Recently, DNA hypomethylation mutations in the *DDM1* gene encoding a SWI2/SNF2 chromatin-remodeling factor have also been reported to result in activation of DNA transposable elements and developmental abnormalities in Arabidopsis (Kakutani et al. 1996; Miura et al. 2001; Singer et al. 2001). Since no apparent developmental abnormalities were observed in the lines of the common morning glory displaying highly variegated flowers, it is unlikely that *Modulator* is directly related to the *DDM1* function.

Chapter II. The genomic structures of the *CHS-D* gene regions in natural isolates of the common morning glory

1. Introduction

The common morning glory is thought to be native to Central America and bears dark purple flowers (Ennos et al. 1983; Yoneda 1990; Austin and Huaman 1996; Glover et al. 1996; Clegg and Durbin 2000). The plant also grows as a common weed in the southeastern United States and its varieties displaying white flowers and variegated flowers, similar in phenotype to the mutable a^f line, were described earlier (Epperson and Clegg 1987; 1988; 1992; Tiffin et al. 1998; Rausher et al.1999; Durbin et al.2000). Clegg and his colleagues characterized these natural varieties isolated in Georgia (Ennos and Clegg 1983; Epperson and Clegg 1987), and they found that all the mutants producing either stable white flowers or white flowers with pigmented spots and sectors carry mutations in the *A* locus for *Albino* (Epperson and Clegg 1987; 1988). Moreover, they designated the stable and variegated white flower mutations as *a* and a^*, respectively (Epperson and Clegg 1987; 1992). Later, Rausher and his colleagues also studied natural isolates of the common morning glory obtained in North Carolina, and they followed the nomenclature *a* and a^* in the *A* locus mutations (Tiffin et al.

1998; Rausher et al 1999). In the mean time, Habu et al (1998) demonstrated that the mutable *flaked* of a^f lines have the transposable element *Tip100* inserted into the intron of the *CHS-D* gene. However, it has not been completely clear whether the *A* locus, characterized by Clegg and his colleagues as well as Rausher and his collaborators, corresponds to the *CHS-D* gene or not, although the data suggesting that the *A* locus is likely to be the *CHS-D* gene were reported (Durbin et al. 2000). Thus, I have been interested in examining whether their *a* and a^* mutations in the *A* locus are indeed mutations in the *CHS-D* gene. In addition, I have also examined whether any polymorphisms can be found in the *CHS-D* gene regions of the natural isolates displaying normally pigmented flowers. Indeed, I was able to find polymorphisms apparently caused by mobile genetic elements in the *CHS-D* gene regions.

2. Materials and methods

2.1. Plants used for the analysis

Ten varieties displaying pigmented flowers, 12 *a** variegated flower mutants bearing white flowers with colored spots and sectors and 5 apparent white flower *a* mutants used in this study (Table II-1) were kindly provided by L.C. Coberly (Duke University). All of those plants were collected from four different areas (PR, Bt, Fe, and Lb) in North Carolina (Fig. II-1), and they are named after the areas where the seeds were collected. The assignment of the mutable *a** varieties were based on their characteristic flower variegation phenotype (Epperson and Clegg 1987; 1988). To confirm that all of the stable white flower mutants isolated indeed carry mutations in the *A* locus, they were crossed with the typical *a** mutant PR-N to examine whether complementation occur to generate plants with pigmented flowers or not since the *a** allele is known to be recessive. To compare the structures of the *CHS-D* genes in the wild type and authentic mutable *a*^{f1} variegated lines, I employed the lines KK/FR-35 (*A/A*) and KK/VR-40a (*a*^{f1}/*a*^{f1}) described by Habu et al. (1998). The *a*^{f1} allele is the *CHS-D* gene having *Tip100* inserted within its intron (Fig. II-2). In addition, a stable white flower line KK/VR-40a-GL1, derived from KK/VR-40a, was also used for my characterization (Table II-1).

TableII-1 Varieties carrying different alleles in the *A* locus

	(varieties)	genotype	Primer sets used 5p2/r3p	12LA/I3	5P3/7LAR	comment
Plant with colored flower						
group1	(KK/FR-35;PR-D;FeD)	*A1/A1*	5.3	1.0	1.6	KK/FR-35[#1]
group2	(Bt-7;Fe-A;Fe-B)	*A2/A2*	n.a.	1.2	2.5	
group3	(Fe-C)	*A1/A2*	5.3	1.0+1.2[#2]	1.6+2.5[#2]	
group4	(Bt-2;Fe-1;Fe-2)	*A1/a^{f}1*	5.3+9.1[#2]	1.0	n.d.	heterozygotes
group5	(Fe-3)	*A2/a^{f}1*	9.1	1.0+1.2[#2]	1.6+2.5[#2]	heterozygote
Plant with variegated flowers						
group6	(KK/VR-40a;PR-M;PR-K;PR-S;PR-KK; Bt-5;Bt-3;PR-JJ;Bt-8;Fe-2;Lb-5; Lb-9)	*a^{f}1/a^{f}1*	9.1	1.0	1.6	KK/VR-40[#1]
group7	(PR-W)	*a^{f}1-d1/a^{f}1-d1*	7.1	n.a.	1.6	
group8	(US1)	*a^{f}3/a^{f}3*	n.d.	n.d.	n.d.	(Durbin et al. 2001)
Plant with white flowers						
group9	(PR-N; Bt1)	*a^{f}1/a^{f}1*	9.1	1.0	1.6	inactive *Modulator*[#3]
group10	(KK/WP-3)	*a^{f}2/a^{f}2*	n.d.	n.d.	n.d.	(Habu et al. 1998)
group11	(KK/VR40a-GL1)	*a^{f}1-d2/a^{f}1-d2*	8.1	1.0	n.a.	
group12	(Bt5)	*a^{f}1-d3/a^{f}1-d3*	8.2	1.0	n.a.	
group13	(Lb4; Lb11)	*a^{f}1-d4/a^{f}1-d4*	n.a.	1.0	n.a.	

[#1]KK/FR-35 and KK/VR-40a were used as authentic plants carring the wild type *A1* and mutable *a^{f}1* alleles, respectively.

[#2]Both length of products were amplified.

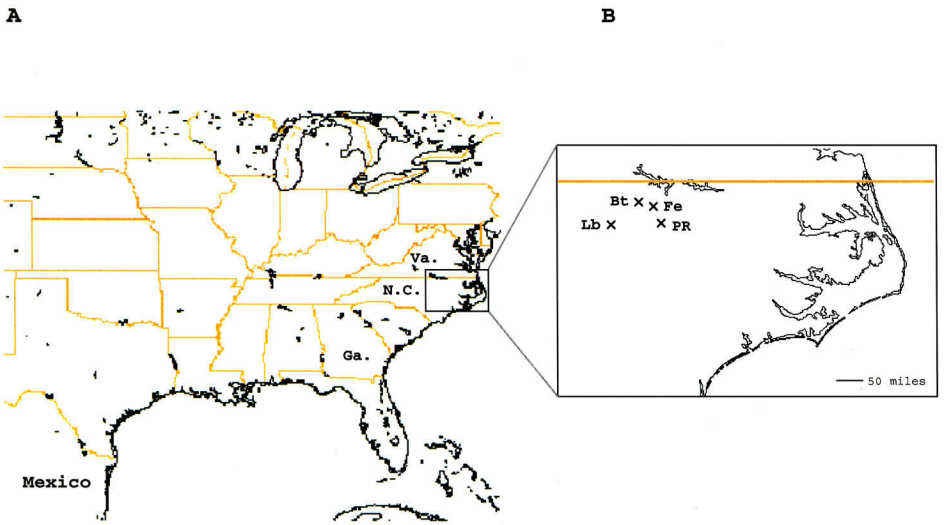
[#3]The group of plants display white flowers due to carrying inactive *Modulator*.

n.a.=not amplified, n.d.=not determined,

Fig. II-1 The areas where the natural isolates of the common morning glory were found.

(A) The map of Georgia (Ga.), and North Carolina (N.C.), and Virginia (Va.) in the United States.

(B) The location of the PR, Bt, Fe and Lb areas where the seeds of the common morning glory were collected.


A
B
Va.
N.C.
Ga.
Mexico
Bt
Fe
Lb
PR
50 miles

2.2. Molecular analysis

The structures of the *CHS-D* gene regions of the varieties were first characterized by PCR analysis using appropriate primers (Fig II-2). The primers used for PCR amplification analysis are:

5p2 (5'-CTCTCGTATAAATAGTAGTGCAAACTTACT-3')

r3p (5'-ATCACATTGGATGTTACTCTCACACTGCAACTGC-3')

5p3 (5'-GAATTCTTATTTAGGTACTTCAATTCATAC-3')

7LAR (5'-CTAACTGATCTATTAAGCATATCTTGTGAAATGTTG-3')

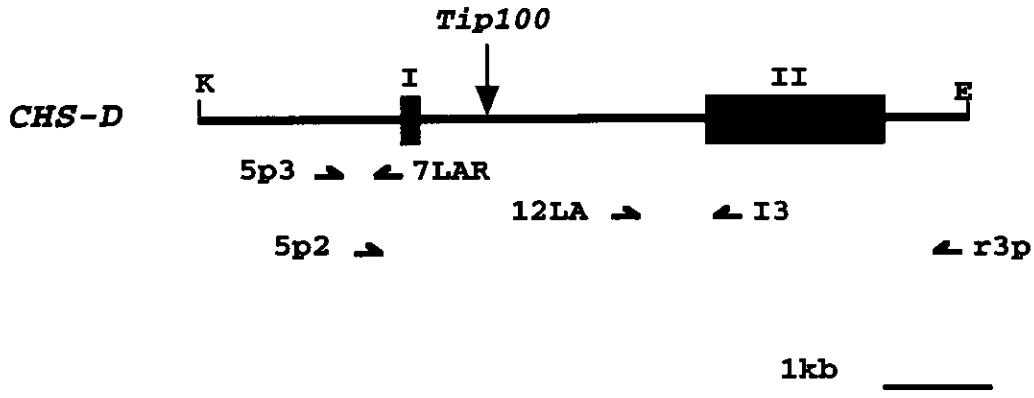
12LA (5'-GTTTGAGACCAAACGTGAAACATGTTATGTTATCTTC-3')

I3 (5'-CATCCAAAGAAGGCGCCATGTATTCACAAAAGCTT-3').

The PCR amplified fragments different in length from those obtained from the control lines KK/FR-35 (*A/A*) and KK/VR-40a for (a^{f1}/a^{f1}) were characterized further by determining their DNA sequences.

Fig. II-2 The primers used for the PCR analysis.

The shaded boxes indicate exon I and exon II of the *CHS-D* gene. The small horizontal arrows indicate the positions and directions of the primers used. The vertical arrow with *Tip100* above the *CHS-D* intron indicates the insertion site of *Tip100* in the a^f allele. Restriction sites are: E, *Eco*RI; K, *Kpn*I.


Tip100
CHS-D
K
I
II
E
5p3
7LAR
12LA
I3
5p2
r3p
1kb

3. Results and discussion

3.1. Genomic structures of the *CHS-D* gene regions in plants displaying colored flowers

Out of 10 varieties bearing colored flowers, four plants, Bt-2, Fe-1, Fe-2 and Fe-3, were found to be heterozygotes containing the wild type *A* and mutable a^f alleles (Table II-1). Segrigants producing variegated flowers appeared in the selfed progeny of these 4 heterozygotes, confirming that all of them carry the mutable a^f allele. Two plants (PR-D and Fe-D) out of the remaining 6 varieties were shown to be identical to the wild type *A* allele carried by KK/FR35 (Habu et al. 1998). Since another *A* allele was also found among the varieties producing colored flowers (see below), I have decided to designate this wild type allele as *A1* and the new allele as *A2* (Table II-1).

The three plants, Bt-7, Fe-A and Fe-B, carrying the *A2* allele failed to produce the 5 kb fragment by PCR amplification using the primers 5p2 and r3p (Table II-1; FigII-2). Subsequent PCR analysis indicated that they appeared to carry two discrete insertions within the *CHS-D* gene region: one is about 0.9 kb insertion at the *CHS-D* promoter region as detected by PCR amplification with the primers 5p3 and 7LAR, and another is around 0.2 kb insertion within the *CHS-D* intron near exon II as detected by PCR amplification with the primers 12LA and I3 (Table II-1; Fig II-2). To characterize further these PCR

amplified fragments were cloned, sequenced and compared with the previously determined *CHS-D* sequence of the wild type *A1* plant KK/FR35 (Habu et al. 1998). As Fig. II-3 shows, the apparent 0.2 kb insertion adjacent to exon II was due to the increase of an additional copy of approximately 200bp tandem repeats (Fig.II-3). These three *A2* plants carry four tandem copies of the 0.2 kb sequence instead of the three tandem copies in the wild type *A1* allele. The other apparent 0.9 kb insertion at the *CHS-D* promoter region was found to be a complicated substitution. Comparison of the *A1* and *A2* alleles revealed that about 0.6 kb part at the *CHS-D* promoter is substituted by around 1.5 kb segment (Fig. II-3; Fig. II-4A). However, the 1.5 kb substituted segment in the *A2* allele contains about 350 bp segment homologous to the corresponding promoter region in the Japanese morning glory (Fig. II-3; Fig. II-4B). Close examination of the region in question revealed that at least two elements, the 235 bp element called SINEIp1 and the 932 bp element named *MELS9* appeared to be integrated into the 350 bp segment homologous to the promoter sequence of the Japanese morning glory (Fig. II-3; Fig. II-4B). In addition, a small 31 bp insertion seemed to have occurred to the right of the inserted *MELS9* element. The 932 bp *MELS9* element carries 14 bp TIRs and is flanked by 3 bp TSDs, as is shown in Fig. II-5. The 3 bp TSD sequences TAA and the 14 bp

Fig. II-3 Comparison of the genomic structures of the *CHS-D* gene.

Structures of the *A1* and the *A2* alleles and corresponding region of the Japanese morning glory are shown. The open boxes with roman numerals indicate exon I and exon II of the *CHS-D* gene. Previously identified transposable elements and mobile element-like sequences (*MELSs*) in the region are also indicated (see Table 1 and Table 2 in general introduction). Open triangles represent TIRs of each element.

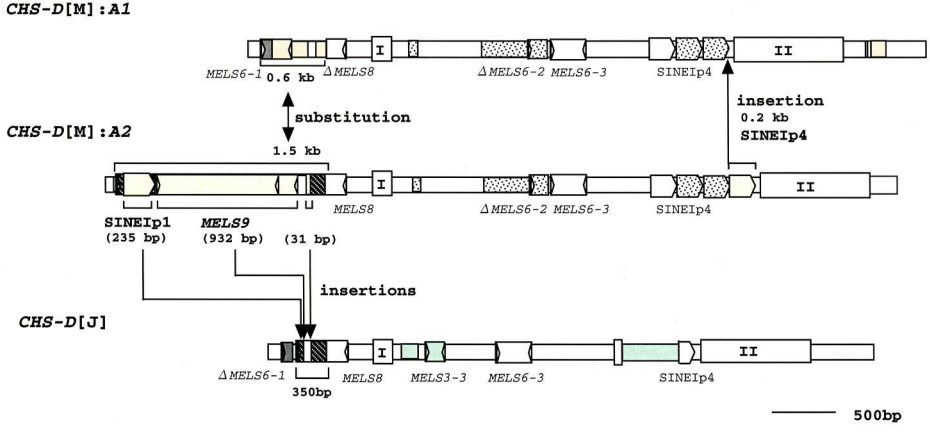
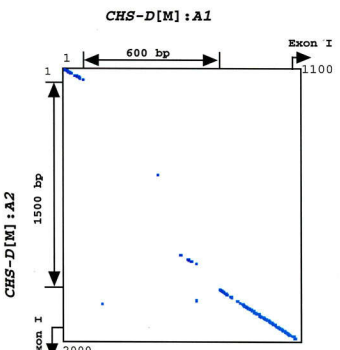

CHS-D[M]:A1
MELS6-1
0.6 kb
ΔMELS8
I
ΔMELS6-2
MELS6-3
SINEIp4
II
insertion
0.2 kb
SINEIp4
substitution
CHS-D[M]:A2
1.5 kb
MELS8
SINEIp1
(235 bp)
MELS9
(932 bp)
(31 bp)
insertions
CHS-D[J]
ΔMELS6-1
350bp
MELS3-3
500bp
Specific region in
CHS-D[M]:A1
CHS-D[M]:A2
CHS-D[J]
Homologous region in
All three
CHS-D[M]:A1 and CHS-D[M]:A2
CHS-D[M]:A1 and CHS-D[J]
CHS-D[M]:A2 and CHS-D[J]

Fig. II-4 Sequence comparison of the *CHS-D* promoter regions.

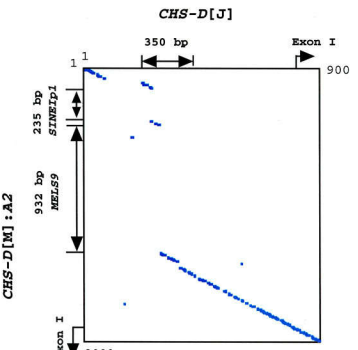
(A) Homology plot of the *A1* allele (*CHS-D*[M]:*A1*) and the *A2* (*CHS-D*[M]:*A2*) allele in the common morning glory.

(B) Homology plot of the *A2* allele (*CHS-D*[M]:*A2*) and *CHS-D* of the Japanese morning glory (*CHS-D*[J]).

A



B



Check Size: 10 Matching Size: 10

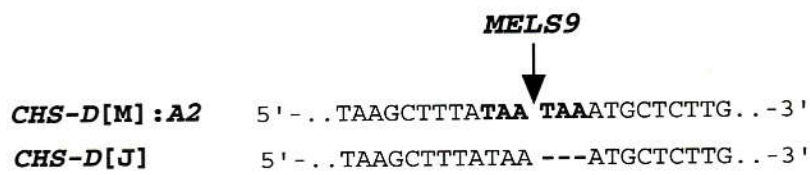
Fig. II-5 Structural characteristics of the *MELS9* element in the *A2* allele

(A) Sequence comparison between the flanking regions of *MELS9* in the *A2* allele of the common morning glory (*CHS-D*[M]:*A2*) and the homologous region of the Japanese common morning glory (*CHS-D*[J]). Three base pair of TSDs (TAA) generated upon insertion are shown in boldface.

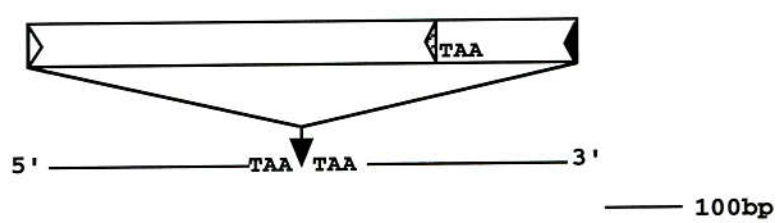
(B) Schematic representation of structure of *MELS9*. The open and solid triangles represent 14 bp imperfect TIRs at the left and right termini of the element, respectively. Another 14 bp sequence, which is the same sequence with TIR at the right end of the element, is also found within the element as is indicated by shaded triangle. Like the right TIR sequence, its adjacent sequence is also TAA.

(C) Alignment of TIRs from members of the *Tourist* elements. TIRs of *MELS9* at the right (*MELS9*-R) and left (*MELS9*-L) termini of the element are compared to TIRs of four different subfamiles of the *Tourist* elements identified from maize, barley, rice and sorghum (Bureau and Wessler 1994)

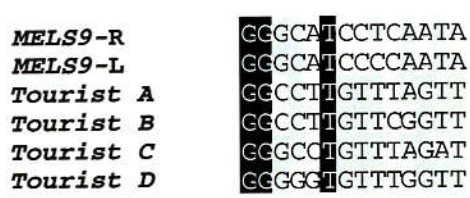
A


MELS9
CHS-D[M]:A2 5'-..TAAGCTTTATAA TAAATGCTCTTG..-3'
CHS-D[J] 5'-..TAAGCTTTATAA ---ATGCTCTTG..-3'

B


TAA
5'
TAA TAA
3'
100bp

C


MELS9-R GGGCATCCTCAATA
MELS9-L GGGCATCCCCAATA
Tourist A GGCCTTGTTTAGTT
Tourist B GGCCTTGTTCGGTT
Tourist C GGGCCTGTTTAGAT
Tourist D GGGGGTGTTTGGTT

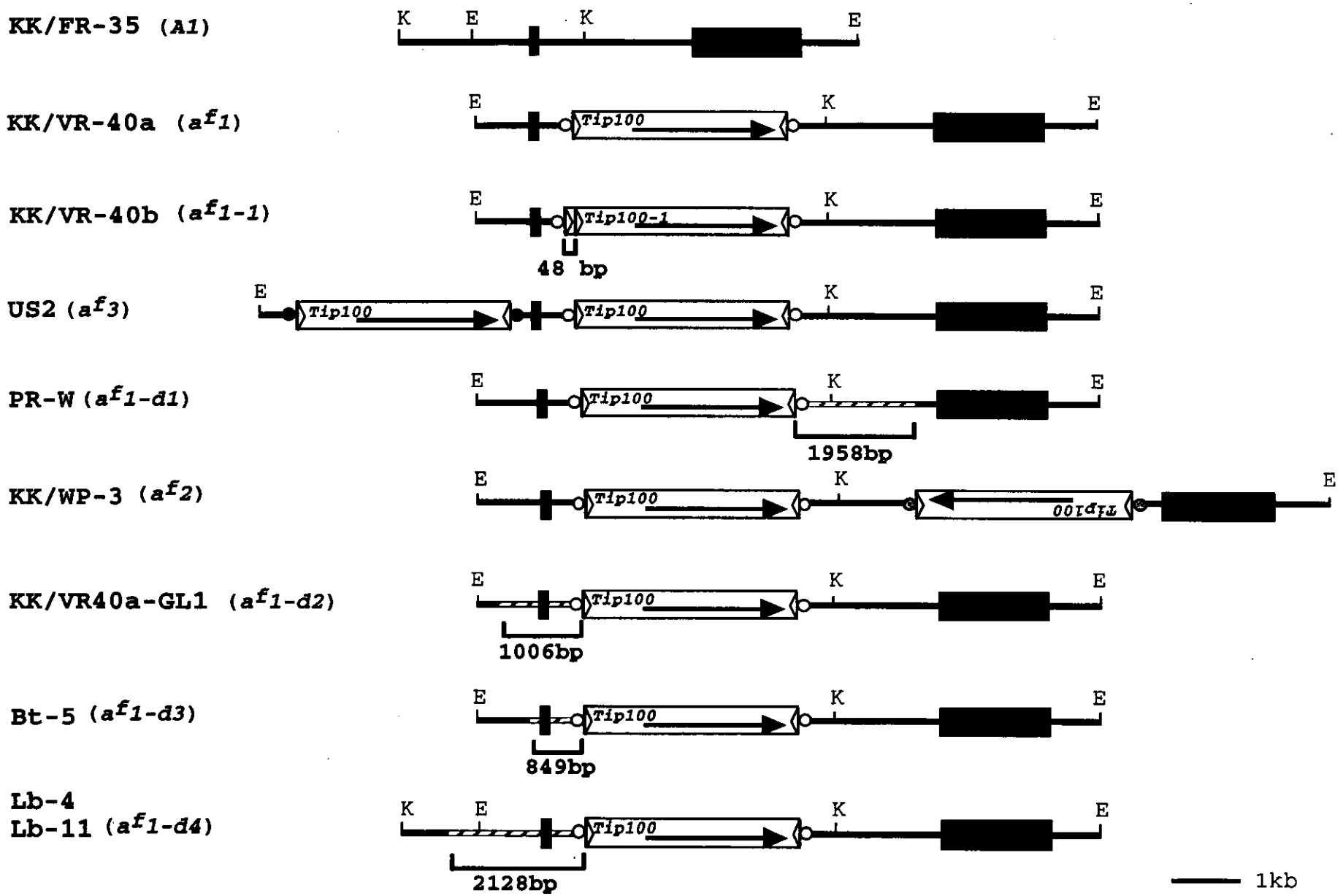
TIR sequence indicate that it is likely to be a member of *Tourist* previously described by Bureau and Wessler (1992; 1994). The 235 bp SINEIp1 element contains a characteristic split RNA polymerase III promoter sequences. It is clear that complicated DNA rearrangements must have taken place in the *CHS-D* promoter region of the *A2* plants. Interestingly, no plants examined carry a recombinant structure between the *A1* and *A2* alleles, even these alleles coexist in the population of the common morning glory. Indeed, the variety Fe-C was found to be a heterozygote bearing both *A1* and *A2* alleles.

3.2. Characterization of the a^* alleles conferring variegation flowers

It has been reported that two mutable a^f alleles, a^{f1} and a^{f1-1}, conferring white flowers with pigmented spots and sectors (Habu et al. 1998). The a^{f1} allele is the *CHS-D* gene having *Tip100* inserted within its intron (Fig. II-6), which is the most frequently detected allele found in the mutable lines KK/VR-37, KK/VR-40a, KK/VP-347 and KK/VP-SU2001 (Habu et al. 1998). The line KK/VR-40a (a^{f1}/a^{f1}) is a subline of KK/VR-40, and another subline KK/VR-40b (a^{f1-1}/a^{f1-1}) derived from KK/VR-40 carries *Tip100-1*, which contains an additional 48 bp terminal sequence as tandem repeats, although its insertion site within the *CHS-D*

Fig. II-6 Structures of various alleles derived from the a^{f1} allele.

The shaded boxes indicate the exon sequences of the *CHS-D* gene. The arrows within the *Tip100* boxes represent the transposase gene. The circles flanking *Tip100* indicate 8 bp of TSDs. The hatched thin boxes indicate the deleted regions and the length of individual deletion is shown under the each map. Restriction sites are: *E*, *Eco*RI; *K*, *Kpn*I.


KK/FR-35 (A1)
K
E
K
E
KK/VR-40a (af1)
E
Tip100
K
E
KK/VR-40b (af1-1)
E
Tip100-1
K
E
48 bp
US2 (af3)
E
Tip100
Tip100
K
E
PR-W (af1-d1)
E
Tip100
K
E
1958bp
KK/WP-3 (af2)
E
Tip100
K
Tip100
E
KK/VR40a-GL1 (af1-d2)
E
Tip100
K
E
1006bp
Bt-5 (af1-d3)
E
Tip100
K
E
849bp
Lb-4
Lb-11 (af1-d4)
K
E
Tip100
K
E
2128bp
1kb

gene is identical to that in the mutable a^{f1} allele (Fig. II-6).

Out of 12 a^* mutants bearing white flowers with colored spots and sectors examined, 11 varieties produced the 9.1 kb fragment which coincides in length with the *CHS-D* segment containing *Tip100* found in the plants carrying the a^f allele, when PCR amplification was carried out with the primers 5p2 and r3p. To examine whether they also contain *Tip100* at the *CHS-D* intron, I sequenced the PCR amplified fragments and found that all of them carry the a^{f1} allele bearing *Tip100* with the identical 8 bp TSDs. It is thus clear that the a^* allele, originated from North Carolina in the southeastern United States and characterized by Rausher and his colleagues (Tiffin et al. 1998), is identical to the a^{f1} allele described by Habu et al. (1998). Very recently, another a^* allele originated from Georgia in the southeastern United States and characterized by Clegg and his colleagues (2001) was also reported to carry *Tip100* at the same insertion site as the *Tip100* integration site in the a^{f1} allele, although their a^* mutant also carries an additional *Tip100* insertion at 333 bp in front of the ATG initiation codon in the *CHS-D* exon I sequence (Durbin et al. 2001; Fig. II-6 US2). It is likely that their a^* mutant was also derived from the common a^{f1} ancestor. In addition, the insertion of *Tip100* in the *CHS-D* promoter region does not affect the expression of the *CHS-D* gene because their a^* mutant displays white flowers with pigmented

spots and sectors. Probably, the essential *cis*-acting elements for the *CHS-D* promoter activity reside the 333 bp region between the second *Tip100* insertion site and the ATG initiation codon in the *CHS-D* gene.

Sequence analysis of the remaining one variety PR-W among the 12 a^* mutants revealed that it carries an 1958 bp deletion adjacent to the *Tip100* insertion within the *CHS-D* intron, and I have designated this allele as a^f*1-d2* (Fig. II-6). It is clear that the a^f*1-d2* allele is also derived from the a^f*1* allele and that the 1958 bp deletion must have been promoted by the *Tip100* element. Therefore, I can conclude that both a^f characterized before (Barker 1917; Imai and Tabuchi 1935; Hisatomi et al. 1997 b; Habu et al, 1998) and a^* originated from the southeastern United States (Epperson and Clegg 1987; Tiffin et al. 1998) are originated from the common ancestor carrying the a^f*1* allele (Fig. II-6).

3.3. Characterization of the *a* alleles conferring stable white flowers

It has been documented that apparent stable white flower mutations in the *A* locus can be classified into two subclasses (Habu et al. 1998). One subclass of the plants was shown to carry the mutable a^f*1* allele without active *Modulator*, and the other was found to contain two copies of *Tip100* integrated into the

CHS-D gene. In the latter stable allele designated as $a^{f}2$, excision of one of two copies of *Tip100* from the *CHS-D* gene is thought to be insufficient to restore the *CHS-D* function and both elements are unlikely to be rarely excised in the same tissue (Habu et al. 1998).

I have characterized 5 apparent white flower mutants carrying mutations in the *A* locus and originated from the southeastern United States and one mutant KK/VR-40a-GL1, derived from KK/VR-40a (Table II-1). Of these total 6 stable white flower mutants, two varieties PR-N and Bt-1 originated from the southeastern United States were found to carry the mutable $a^{f}1$ allele, indicating that they are likely to bear inactive *Modulator* plants. The results indicate that the $a^{f}1$ allele plants due to the absence of active *Modulator* displaying white flowers are present among the natural isolates.

All of the remaining four varieties, Bt-5, Lb-4, Lb-11 and KK/VR-40a-GL1, have *Tip100* mediate deletions removing the exon I sequence. Two of them, Lb-4 and Lb-11, carrying the identical deletion are probably derived from the common ancestor because they were isolated from the same Lb field (Fig. II-1). The deletion in these Lb-4 and Lb-11 plants appears to be removed both exon I and the *CHS-D* promoter since insertion of *Tip100* at 333 bp in front of the ATG initiation codon in the *CHS-D* exon I sequence was not seemed to affect the expression of the *CHS-D* gene as discussed above (Durbin et al. 2001; Fig. II-6). Deletion

formation mediated by transposable elements including prokaryotic IS elements is one of the characteristic features commonly found in DNA transposable elements (Iida et al. 1983; Taylor and Walbot 1985; Dooner et al. 1988; Martin et al. 1988; Robbins et al.1989; Lister and Martin 1989; Roberts DE et al.1991; Gray 2000; Jilk et al 1993).

3.4. A possible evolutionary process to generate the variants and mutants carrying sequence alterations in the *CHS-D* gene region of the common morning glory

As discussed above, both the a^f and a^* alleles are highly likely to be derived from the single common ancestor carrying the a^{f1} allele (Fig. II-6). All of the stable white flower mutants so far characterized are also derivatives of the common ancestor with the a^{f1} allele carrying *Tip100* inserted into the *CHS-D* intron. Two subclasses of the stable white flower mutants can be found: the plants with the mutable a^{f1} allele without active *Modulator* and those with deficient *CHS-D* gene caused by *Tip100* mediated DNA rearrangements. The DNA rearrangements mediated by *Tip100* so far observed are either an insertion of the second copy of *Tip100* into the *CHS-D* gene or deletions removing the critical sequences for the *CHS-D* gene activity.

As described in the section 3.1, there are two distinct alleles, *A1* and *A2*, in the *CHS-D* gene among the plants bearing

pigmented flowers. The allele *A1* appears to occupy in larger populations than the allele *A2*. When compared to the *A1* allele, the *A2* allele has complicated DNA rearrangements including insertions of 235 bp SINEIp1 and 932 bp *MELS9*. Interestingly, the mutable $a^{f}1$ allele is clearly derived from the *A1* allele since no DNA rearrangements associated with the *A2* allele was observed in the mutable $a^{f}1$ allele and its derivatives, indicating that the *A1* allele must be the progenitor of the mutable $a^{f}1$ alleles. The mutable $a^{f}1$ allele must have formed by a simple insertion of 3.9 kb *Tip100* into the *A1* allele. Subsequent *Tip100* mediated DNA rearrangements including transposition of *Tip100* from a replicated sister chromatid into another replicated sister chromatid and *Tip100* mediated deletion formation must have resulted in various *a* alleles described in Fig. II-6, -7. These observations strongly support the notion that mobile genetic elements act as major spontaneous mutagens and play important roles in generating floricultural traits in the common morning glory (Fig. II-8). A comparable observation with respect to this notion has been discussed in the spontaneous mutants in the Japanese morning glory (Iida et al. 1999; Hoshino et al. 2001).

Fig. II-7 Possible pathways to generate various alleles from the *a^{f1}* allele.

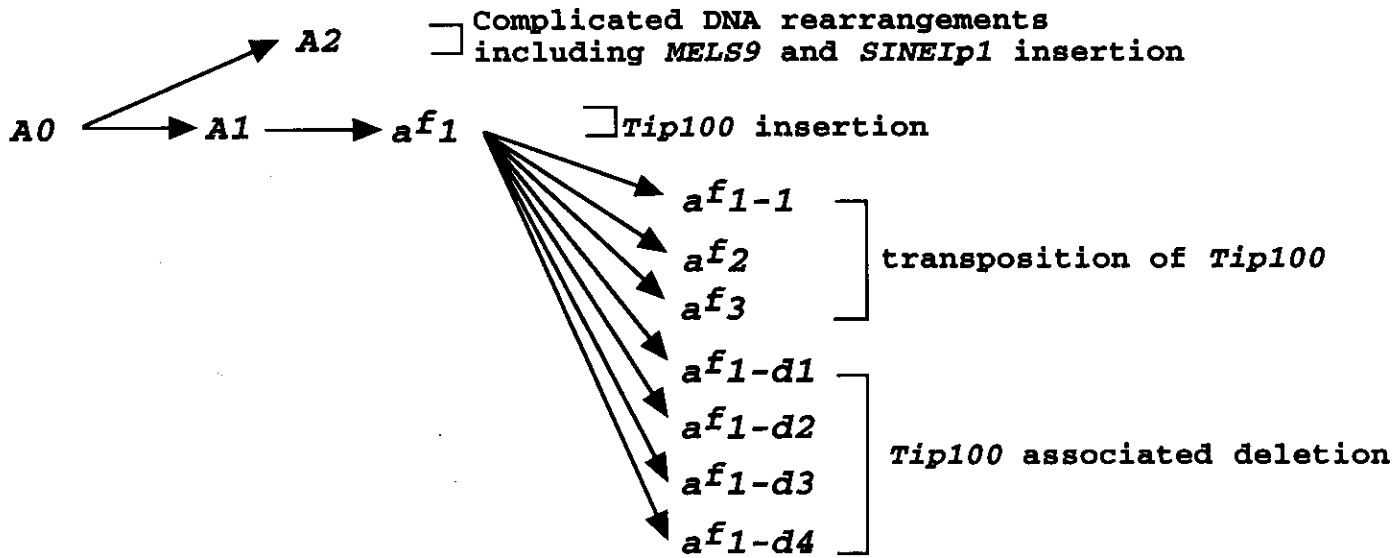
(A) After the replication, *Tip100* in one of the daughter strands transpose into the other strand. The boxes represent *Tip100* and the bold short arrows indicate its transposase gene. The triangles indicate TIRs. The arrows with the lower case characters indicate one of the three possible transposition events.

(B) Structures after transposition. Footprints generated by *Tip100* transposition are indicated by X. **(a)** *Tip100* transposed into the other daughter strand in front of *Tip100*. **(b)** *Tip100* transposed into the *Tip100* sequence. Subsequent homologous recombination between the directly repeated *Tip100* sequences, as indicated by the bracket under the map, would generate the *a^{f1-1}* allele. **(c)** *Tip100* transposed into a region to the right of the *Tip100* insertion site.

(C) The alleles derived from the *a^{f1}* allele.

Fig. II-8 Generation of the various alleles studied in the *CHS-D* locus.

The structure of the individual allele is drawn in Fig.II-3 and Fig.II-6.


A2
Complicated DNA rearrangements
including MELS9 and SINEIp1 insertion
A0
A1
af1
Tip100 insertion
af1-1
af2
af3
transposition of Tip100
af1-d1
af1-d2
af1-d3
af1-d4
Tip100 associated deletion

Summary and conclusion

The common morning glory (*Ipomoea purpurea*), a native to Central America, is a common weed in the southeastern United States. The plant was introduced to Europe probably in the 17th century and to Japan in the 18th century. The cultivars with purple, red and white flowers were already recorded in the 18th century, and the most extensively characterized mutant displaying white flowers with pigmented spots and sectors was recorded in the early 19th century. The early genetic studies on the variegated flower line, named *flaked*, were conducted by Barker (1917) in the United States and followed by Imai and Tabuchi (1935) in Japan. The *flaked* mutation, also called *anthocyanin*^{flaked} (a^{flaked} or a^{f}), was recently shown to be caused by the insertion of the transposable element *Tip100* into the *CHS-D* gene for anthocyanin biosynthesis. The 3.9 kb *Tip100* element belonging to the *Ac/Ds* family contains an open reading frame encoding a polypeptide of 808 amino acids. The patterns of the flower variegations in the mutable *flaked* (a^{f}) lines are determined by the frequency and timing of the excision of *Tip100* from the *CHS-D* gene. In addition to the mutable a^{f} allele, there is a leaf variegation mutation *yellow-green leaf*^{mutable}, or ygl^{m}, conferring dark green spots and sectors on a yellow green background. A plant carrying the two alleles, a^{f} and ygl^{m}, tends to show a peculiar similarity in both timing and frequency of

the variegations in flowers and leaves, and the patterns of the flower and leaf variegations are also generally heritable. Based on the observation, it has been proposed that there must be another genetic element termed *Modulator* acting on both a^f and ygl^m alleles and that the timing and frequency of the variegations in flowers and leaves are determined by the heritable state of the *Modulator*.

As an initial step to understand the complicated flower and leaf variegation system in the common morning glory, I examined whether *Tip100* is an autonomous element carrying an active transposase gene within the element. *Tip100* and its internal deletion derivative *dTip100* were introduced into tobacco (*Nicotiana tabacum*), and their capability to transpose in the transgenic tobacco plants was examined. The introduced *Tip100* element was able to transpose from the integrated vector into the tobacco genome whereas *dTip100* was not. The defective *dTip100* could be transposed by introduction of *Tip100*, indicating that *Tip100* is the first autonomous element characterized in *Ipomoea* species. Based on these results, I have speculated the nature of the putative *Modulator* element affecting flower and leaf variegations in various mutable lines.

To characterize whether many different mutant alleles in the *CHS-D* gene are present in the common morning glory producing either white flowers or white flowers with pigmented spots and sectors, the structures of the *CHS-D* gene regions in the mutants

grown as a weed in the southeastern United States. All of the mutants examined were found to be derivatives of the a^f mutant previously characterized, and some of them carry additional DNA rearrangements associated with *Tip100*. Based on the results, I have concluded that there had been a "founder" mutation in the *CHS-D* gene and that subsequent DNA rearrangements mediated by *Tip100* and other mobile genetic elements had occurred in some of the geographically spread mutants originated from the founder. I have also discussed a possible evolutionary process to generate variants and mutants carrying sequence alterations in the *CHS-D* gene region of the common morning glory.

References

Austin DF, Huaman Z (1996) A synopsis of *Ipomoea* (*Convolvulaceae*) in the Americas. Taxon 45: 3-38

Baker B, Schell J, Loerz H, Fedoroff N (1986) Transposition of the maize controlling element "*Activator*" in tobacco. Proc Natl Acad Sci USA 83: 4844-4848

Barker EE (1917) Heredity studies in the morning glory (*Ipomoea purpurea*). Bull Cornell Univ Agric Exp Sta 392: 1-38

Becker HA, Kunze R (1997) Maize *Activator* transposase has a bipartite DNA binding domain that recognizes subterminal sequences and the inverted repeats. Mol Gen Genet 254: 219-230

Benito MI, Walbot V (1997) Characterization of the maize *Mutator* transposable element MURA transposase as a DNA-binding protein. Mol Cell Biol 17: 5165-5175

Bennetzen JL (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol 42: 251-269

Bureau TE, Wessler SR (1992) *Tourist*: a large family of small inverted repeat elements frequently associated with maize genes. Plant Cell 4: 1283-1294

Bureau TE, Wessler SR (1994) Mobile inverted-repeat elements of the *Tourist* family are associated with the genes of many cereal grasses. Proc Natl Acad Sci USA 91: 1411-1415

Calvi BR, Hong TJ, Findley SD, Gelbart WM (1991) Evidence for a common evolutionary origin of inverted repeat transposons

in *Drosophila* and plants: *hobo*, *Activator*, and *Tam3*. Cell 66: 465-471

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

Chomet P, Lisch D, Hardeman KJ, Chandler VL, Freeling M (1991) Identification of a regulatory transposon that controls the *Mutator* transposable element system in maize. Genetics 129: 261-270

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81: 1991-1995

Clegg MT, Durbin ML (2000) Flower color variation: A model for the experimental study of evolution. Proc Natl Acad Sci USA 97: 7016-7023

Coen ES, Carpenter R, Marctin C (1986) Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. Cell 47: 285-296

Curtis W (1790) Convolvulus purpureas. Purple bindweed, or Convolvulus major. Bot. Mag. 3: 113

Does MP, Dekker BMM, de Groot MJA, Offringa R (1991) A quick method to estimate the T-DNA copy number in transgenic plants at an early stage after transformation, using inverse PCR. Plant Mol Biol 17: 151-153

Dooner HK, English J, Ralston EJ (1988) The frequency of transposition of the maize element *Activator* is not affected

by an adjacent deletion. Mol Gen Genet 211: 485-491

Durbin ML, McCaig B, Clegg MT (2000) Molecular evolution of the chalcone synthase multigene family in the morning glory genome. Plant Mol Biol 42: 79-92

Durbin ML, Denton AL, Clegg MT (2001) Dynamics of mobile element activity in chalcone synthase loci in the common morning glory (*Ipomoea purpurea*). Proc Natl Acad Sci USA 98: 5084-5089

Eisen JA, Benito MI, Walbot V (1994) Sequence similarity of putative transposases links the maize *Mutator* autonomous element and a group of bacterial insertion sequences. Nucleic acids Res 22: 2634-2636

Ennos RA, Clegg MT (1983) Flower color variation in the morning glory, *Ipomoea purpurea*. J Hered 74: 247-250

Epperson BK, Clegg MT (1987) Instability at a flower color locus in the morning glory. J Hered 78: 346-352

Epperson BK, Clegg MT (1988) Genetics of flower color polymorphism in the common morning glory. J Hered 79: 64-68

Epperson BK, Clegg MT (1992) Unstable white flower color genes and their derivatives in the morning glory. J Hered 83: 405-409

Essers L, Adolphs RH, Kunze R (2000) A highly conserved domain of the maize Activator transposase is involved in dimerization. Plant Cell 12: 211-223

Fedoroff N (2000) Transposons and genome evolution in plants.

Proc Natl Acad Sci USA 97: 461-468

Frank MJ, Liu D, Tsay YF, Ustach C, Crawford NM (1997) *Tag1* is an autonomous transposable element that shows somatic excision in both Arabidopsis and tobacco. Plant Cell 9: 1745-1756

Frey M, Reinecke J, Grant S, Saedler H, Gierl A (1990) Excision of the *En/Spm* transposable element of Zea mays requires two element-encoded proteins. EMBO J 9: 4037-4044

Gierl A, Lutticke S, Seadler H (1988) TnpA product encoded by the transposable element *En-1* of *Zea mays* is a DNA binding protein. EMBO J 7: 4045-4053

Glover D, Durbin ML, Huttley G, Clegg MT (1996) Genetic diversity in the common morning glory. Plant Species Biol. 11: 41-50

Grandbastien MA (1992) Retroelements in higher plants. Trands in Genet 8: 103-108

Gray YHM (2000) It takes two transposons to tango. Trands in Genet 16: 461-468

Habu Y, Hisatomi Y, Iida S (1998) Molecular characterization of the mutable *flaked* allele for flower variegation in the common morning glory. Plant J 16: 371-376

Haring MA, Gao J, Volbeda T, Rommens CMT, Nijkamp HJJ, Hille J (1989) A comparative study of *Tam3* and *Ac* transposition in transgenic tobacco and petunia plants. Plant Mol Biol 13: 189-201

Haring MA, Teeuwen-de Vroomen Mj, Nijkamp HJJ, Hille J (1991) *Trans*-activation of an artificial *dTam3* transposable element in transgenic tobacco plants. Plant Mol Biol 16: 39-47

Hehl R, Nacken WKF, 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

Hisatomi Y, Hanada K, Iida S (1997a) The retrotransposon *RTip1* is integrated into a novel type of minisarellite, *MiniSip1*, in the genome of the common morning glory and carries another new type of minisatellite, *MiniShip2*. Theor Appl Gene 95: 509-515

Hisatomi Y, Yoneda Y, Kasahara K, Inagaki Y, Iida S (1997b) DNA rearrangements at the region of the dihydroflavonol 4-reductase gene for flower pigmentation and incomplete dominance in the morning glory carrying the mutable *flaked* mutation. Theor Appl Gene 95: 509-515

Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA (1983) A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303: 179-180

Hoshino A, Johzuka-Hisatomi Y, Iida S (2001) Gene duplication and mobile genetic elements in the morning glories. Gene 265: 1-10

Iida S, Meyer J, Arber W (1983) Prokaryotic IS elements. In: Shapiro JA (eds) Mobile Genetic Elements: ACADEMIC PRESS, INC.,pp173-178

Iida S, Hoshino A, Johzuka-Hisatomi Y, Habu Y, Inagaki Y (1999) Floricultural traits and transposable elements in the Japanese and common morning glories. Annal New York Acad Sci 870: 265-274

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

Imai Y, Tabuchi K (1935) Recurrent mutation in the flaked alleles of *Pharbitis purpurea*. J Genet 35: 433-446

Izawa T, Miyazaki C, Yamamoto M, Terada R, Iida S, Shimamoto K (1991) Introduction and transposition of the maize transposable element *Ac* in rice (*Oryza sativa L.*). Mol Gen Genet 227: 391-396

Jarvis P, Belzile F, Page T, Dean C (1997) Increased *Ac* excision (*iae*): *Arabidopsis thaliana* mutations affecting *Ac* transposition. Plant J 11: 907-919

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: β-glucuronidase as a sensitive and vertile gene fusion marker in higher plants. EMBO J 6: 3901-3907

Jilk RJ, Markris JC, Borchardt L, Reznikoff WS (1993) Implications of *Tn5*-Associated adjacent deletions. J Bacteriol 175: 1264-1271

Johzuka-Hisatomi Y, Hoshino A, Mori Y, Habu Y, Iida S (1999) Characterization of the chalcone synthase genes expressed in flowers of the common and Japanese morning glories. Genes Genet Syst 74: 141-147

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ (1996) Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci USA 93: 12406-12411

Kasahara K (1956) Genetic studies on mutable genes in *Ipompea purpurea*. Res Lib Arts Hosei Univ 1: 1-13 (in Japanese)

Kempken F, Windhofer (2001) The hAT family; a versatile transposon group common to plants, fungi, animals, and man. Chromosama 110: 1-9

Kishima Y, Yamashita S, Martin C, Mikami T (1999) Structural conservation of the transposon *Tam3* family in *Antirrhinum majus* and estimation of the number of copies able to transpose. Plant Mol Biol 39: 299-308

Kumar A, Bennetzen L (1999) Plant retroposons. Annu Rev Genet 33: 479-532

Kunze R, Seadler H, Loennig WE (1997) Plant transposable elements. Adv Bot Res 27: 331-470

Lisch D, Girard L, Donlin M, Freeling M (1999) Functional analysis of deletion derivatives of the maize transposon *MuDR* delineates roles for the MURA and MURB proteins. Genetics 151: 331-341

Lister C, Martin C (1989) Molecular analysis of a transposon-induced deletion of the *nivea* locus in *Antirrhinum majus*. Genetics 123: 417-425

Martin C, Mackay S, Carpenter R (1988) Large-scale chromosomal restructuring is induced by the transposable element *Tam3* at the *nivea* Locus of *Antirrhinum majus*. Genetics 119: 171-184

Martin C, Prescott A, Lister C, MacKay S (1989) Activity of the transposon *Tam3* in *Antirrhinum* and tobacco: possible role of DNA methylation. EMBO J 8: 997-1004

Masson P, Strem M, Fedoroff N (1991) The *tnpA* and *tnpD* gene products of the *Spm* element are required for transposition in tobacco. Plant Cell 3: 73-85

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

Peacock WJ, Dennis ES, Gerlach WL, Sachs MM, Schwartz D (1984) Insertion and excision of *Ds* controlling elements in maize. Cold Spling Harbor Symposia on Quantitive Biology. 49: 347-354

Pichersky E (1990) Nomad DNA- A model for movement and duplication of DNA sequences in plant genomes. Plant Mol Biol 15: 437-448

Pozueta-Romero J, Houlne G, Schantz (1998) Identification of a short interspersed repetitive element in partially spliced transcripts of the ball pepper (Capsicum annum) PAP gene: new evolutionary and regulatory aspects on plant tRNA-related *SINEs*. Gene 214: 51-58

Raina R, Schlappi M, Karunanandaa B, Elhofy A, Fedoroff N (1998) Concerted formation of macromolecular *Suppressor-mutator* transposition complex. Proc. Natl. Acad. Sci. USA 95: 8526-8531

Raizada MN, Walbot V (2000) The late developmental pattern of *Mu* transposon excision is conferred by a Cauliflower Mosaic Virus 35S-Driven MURA cDNA in transgenic maize. Plant Cell 12: 5-21

Rausher MD, Miller RE, Tiffin P (1999) Patterns of evolutionary rate variation among genes of the anthocyanin biosynthetic pathway. Mol Biol Evol 16: 266-274

Robbins PR, Carpenter R, Coen ES (1989) A chromosome rearrangement suggests that donor and recipient sites are associated during *Tam3* transposition in of *Antirrhinum majus*. EMBO J 8: 5-13

Roberts DE, Ascherman D, Kleckner N (1991) *IS10* promotes adjacent deletion at low frequency. Genetics 128: 37-43

Saedler H, Nevers (1985) Transposition in plants: a molecular model. EMBO J 4: 585-590

Shiokawa K, Inagaki Y, Morita H, Hsu TJ, Iida S, Noguchi H (2000) The functional expression of the *CHS-D* and *CHS-E* genes of the common morning glory (*Ipomoea purpurea*) in *Escherichia*

coli and characterization of their gene products. Plant Biotech 17: 203-210

Simms E, Rausher M (1987) Costs and benefits of plant resistance to herbivory. American Naturalist 130: 570-581

Sims J (1814) Convolvulus purpureas (δ.) varius. Striped-flowered bindweed. Bot. Mag. 41: 1682

Singer T, Yordan C, Martienssen RA (2001) Robertson's *Mutator* transposons in *A. thaliana* are regulated by the chromatin-remodeling gene *Decrease in DNA Methylation* (*DDM1*). Genes Dev 15: 591-602

Sugita K, Matsunaga E, Ebinuma H (1999) Effective selection system for generating marker-free transgenic plants independent of sexual crossing. Plant Cell Reports 18: 941-947

Tanaka M, Yoshinaga M, Nakazawa Y, Nakatani M (2001) Identification of a novel *SINE*-like sequence in Retrotransposon fragments of sweet potato. Plant Bio. Tec. 4: 869-876

Taylor LP, Walbot V (1985) A deletion adjacent to the maize transposable element *Mu-1* accompanies loss of *Adh1* expression. EMBO J 4: 869-876

Tiffin P, Miller RE, Rausher MD (1998) Control of expression patterns of anthocyanin structural genes by two loci in the common morning glory. Genes Genet Syst 73: 105-110

Trentmann SM, Saedler H, Gierl A (1993) The transposable element *En/Spm*-encoded TNPA protein contains a DNA binding and a dimerization domain. Mol Gen Genet 238: 201-208

Turcotte K, Srinivasan S, Bureau T (2001) Survey of transposable elements from rice genomic sequences. Plant J 25: 169-179

Walker EL, Robbins TP, Bureau TE, Kermicle J, Dellapoets SL (1995) Transposon-mediated chromosomal rearrangements and gene duplications in the formation of the maize *R-r* complex. EMBO J 14: 2350-2363

Wessler SR, Bureau TE, White SE (1995) LTR-retrotransposons and MITEs: important players in the evolution of plant genes. Curr Opin Genet Dev 5: 814-821

Yoneda Y (1990) Japanese morning glory. In: Ammirato PV, Evans DA, Sharp WR, Bajaj YPS (eds) Handobook of Plant Cell Culture. Vol. 5: McGrawHill Publishing Co. New York.,pp509-533.

Yoshioka Y, Matsumoto S, Kojima S, Ohshima K, Okada N, Machida Y (1993) Molecular characterization of a short interspersed repetitive element from tobacco that exhibits sequence homology to specific tRNAs. Proc Natl Acad Sci USA 90: 6562-6566

Zhang X, Feschotte C, Zhang Q, Jiang N, Eggleston WB, Wessler SR (2001) P instability factor; An active maize transposon system associated with the amplification of Tourist-like MITEs and a new superfamily of transposases. Proc Natl Acad Sci USA 98: 12572-12577

Mol Genet Genomics (2002) 266: 732–739
DOI 10.1007/s00438-001-0603-z

ORIGINAL PAPER

N. Ishikawa · Y. Johzuka-Hisatomi · K. Sugita
H. Ebinuma · S. Iida

The transposon *Tip100* from the common morning glory is an autonomous element that can transpose in tobacco plants

Received: 18 June 2001 / Accepted: 4 October 2001 / Published online: 8 November 2001

Abstract The mutable *flaked* or a^{flaked} (a^f) line of the common morning glory (*Ipomoea purpurea*) displays white flowers with colored flakes, and the a^f mutation is caused by the insertion of a transposable element named *Tip100* into the *CHS-D* gene for anthocyanin biosynthesis. The 3.9-kb *Tip100* element belongs to the *Ac/Ds* family and contains an ORF encoding a polypeptide of 808 amino acids. The frequency and timing of flower variegation vary in different a^f lines, and a genetic element termed *Modulator* has been postulated to affect the variegation pattern. Since the pattern of flower variegation is determined by the frequency and timing of excision of *Tip100* from the *CHS-D* gene, we wished to determine whether *Tip100* is an autonomous element that is itself capable of transposition in a heterologous host. To do this, we introduced the element into the genome of tobacco plants by *Agrobacterium*-mediated transformation. The intact *Tip100* element was able to excise from its original position in the chromosome and reinsert into new sites in the tobacco genome, whereas an internal deletion derivative was not. Based on these results, we conclude that *Tip100* is an autonomous element. We also discuss the nature of the putative *Modulator* element affecting flower and leaf variegation in various mutable lines of the morning glory.

Communicated by J. Schell

N. Ishikawa · Y. Johzuka-Hisatomi · S. Iida (✉)
National Institute for Basic Biology, Myodaiji,
Okazaki-shi, Aichi 444-8585, Japan
E-mail: shigiida@nibb.ac.jp
Fax: +81-564-557685

N. Ishikawa · S. Iida
The Graduate University for Advanced Studies,
Okazaki-shi, Aichi 444-8585, Japan

K. Sugita · H. Ebinuma
Pulp and Paper Research Laboratory,
Nippon Paper Industries Co Ltd., Kita-ku,
Tokyo 114-0002, Japan

Keywords *Ac/Ds* family · Autonomous element · Flower and leaf variegation · *Ipomoea purpurea* · Transposable element *Tip100*

Introduction

Transposable DNA elements that move by excision and reintegration can be classified as autonomous or non-autonomous (Kunze et al. 1997). Autonomous elements carry both *cis*-acting elements required for transposition and the complete coding regions for *trans*-acting transposases, whereas non-autonomous elements, in which the transposase gene is defective or lacking, can be mobilized only when active transposases are supplied by an autonomous element elsewhere in the genome. Somatic excisions of these elements from genes involved in pigmentation often cause a variegated phenotype.

The best studied mutation in the common morning glory (*Ipomoea purpurea* or *Pharbitis purpurea*) is a mutable allele of the *A* locus termed *flaked*, also called *anthocyanin*^{flaked} (a^{flaked} or a^f) (Barker 1917; Imai and Tabuchi 1935; Kasahara 1956; Epperson and Clegg 1987; Hisatomi et al. 1997; Habu et al. 1998; Iida et al. 1999; Durbin et al. 2001; Hoshino et al. 2001). In the homozygous state a^f/a^f, the mutable *flaked* lines display white flowers with colored flakes (Fig. 1A). The flower variegation is thought to be caused by recurrent somatic reversion of the non-functional allele (white) to functional (pigmented). One of the characteristics of the mutable *flaked* allele is that the timing and frequency of flower variegation vary in different lines, and the variegation pattern is generally heritable (Kasahara 1956; Habu et al. 1998). We have shown that the mutable a^f allele has an insertion of the 3.9-kb transposable element *Tip100* in the intron of the *CHS-D* gene encoding chalcone synthase for flower pigmentation (Habu et al. 1998; Shiokawa et al. 2000; Hoshino et al. 2001). *Tip100* in the *CHS-D* intron carries 11-bp terminal inverted repeats (TIRs) and is flanked by

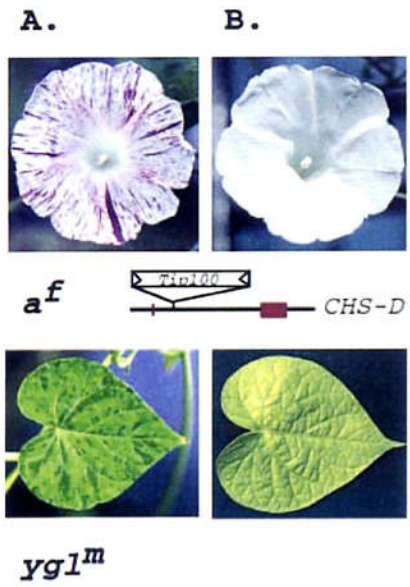


Fig. 1A, B Phenotypes of *Ipomoea purpurea* carrying the mutable a^f and ygl^m alleles for flower and leaf variegation, respectively. The structure of the a^f allele with *Tip100* inserted in the *CHS-D* gene is also shown. The *filled boxes* indicate the exon sequences of the *CHS-D* gene. The mutable ygl^m allele remains to be identified. **A** KK/VR-40a (*Modulator* active). **B** KK/WR321 (*Modulator* inactive)

target-site duplications (TSDs) of 8 bp. It contains an ORF encoding a polypeptide of 808 amino acids which exhibits partial homology to the conserved regions of the transposase of the *Ac/Ds* family. These structural features are compatible with the notion that *Tip100* may be an autonomous element belonging to the *Ac/Ds* family.

Kasahara (1956) described a line of the common morning glory showing variegated leaf pigmentation, characterized by the appearance of dark green spots and sectors on a yellow-green background (Fig. 1A). This mutable allele was named *yellow-green leaf*^{mutable}, or ygl^m (previously described as *y'*). It was further noted that the timing and frequency of the variegated phenotype in leaves are also generally heritable and that a plant carrying both a^f and ygl^m tends to show a striking similarity in both timing and frequency of the variegations in flowers and leaves (Fig. 1). Based on these observations, Kasahara (1956) postulated that there must be another genetic element, termed *Modulator* (originally named *Mutator*), that acts on both the a^f and ygl^m alleles, and that the timing and frequency of variegation in flowers and leaves are determined by the heritable state of the *Modulator*. By crossing them with an active *Modulator* line, it is possible to distinguish *a* mutant lines bearing white flowers due to a stable mutation at the *A* locus from white-flower mutants carrying the a^f allele without an active *Modulator*. Subsequently, we found that lightly and heavily variegated flower lines, as well as the white-flower line with inactive *Modulator*, all carry a single *Tip100* insertion at the same site in the *CHS-D* intron (Fig. 1; Habu et al. 1998). In contrast, the stable white-flower line with active *Modulator* has two insertions of *Tip100* within the *CHS-D* intron. Excision of one of the two copies of *Tip100* from the *CHS-D* gene in the latter stable white-flower line appears to be insufficient to restore the *CHS-D* function, and probably both elements are rarely excised in the same tissue (Habu et al. 1998).

The simplest explanation for these observations is that *Tip100* and *Modulator* correspond to a non-autonomous and an autonomous element, respectively. Alternatively, *Tip100* may itself be an autonomous element and *Modulator* an element which can control the transposition activity of *Tip100*. To determine whether *Tip100* is an autonomous element, we introduced *Tip100* and its internal deletion derivative *dTip100* into tobacco (*Nicotiana tabacum*), and examined whether they are able to transpose in the transgenic tobacco plants. Our data strongly indicate that *Tip100* carries all the functions necessary for its own transposition, and therefore represents the first autonomous element to be characterized in the genus *Ipomoea*.

Materials and methods

Plasmid vectors

The *Tip100* element flanked by 8-bp TSDs in the clone λV-Tip100-4S7 (Habu et al. 1998) was amplified by PCR and cloned into the *Pst*I site of pHSG398 (TaKaRa Biomedicals) to yield pHSG 398::*Tip100*. The primers used for PCR amplification were: Tip100-L (5'-GG**CTGCAG***CATACGTG*CAGGGGCGGAGGCA C-3') and Tip100-R (5'-GG**CTGCAG***CACGTATG*CAGGGGCGGAG CCAGGATTA-3'), which contain a *Pst*I site (boldface), the 8-bp TSD sequence (italics) originating from the *CHS-D* intron, and the *Tip100* terminal sequences (underlined). The internal 1.7-kb *Spe*I segment of *Tip100* in pHSG398::*Tip100* was deleted to yield pHSG398::*dTip100*.

The plasmid pSUG1 is a derivative of pBI121 (Clontech) and carries a unique *Sse*8387I site at the *Sma*I site of pBI121. To construct pSUG1::*Tip100* and pSUG1::*dTip100* (Fig. 2), the *Pst*I fragment containing either *Tip100* or *dTip100* was cloned into the *Sse*8387I site of pSUG1, thus separating the *GUS* gene from the 35S promoter.

The *Bam*HI-*Sac*I fragment containing the *GUS* gene for β-glucuronidase in pBI121 was replaced by the *Bam*HI-*Sac*I fragment containing the *hph* gene for hygromycin B phosphotransferase and the CaMV 35S terminator from pCKR138 (Izawa et al. 1991) to yield pSUG2. To obtain pSUG2::*Tip100*, the *Pst*I fragment containing *Tip100* was cloned into the *Sse*8387I site of pSUG2.

Transgenic plants

Agrobacterium-mediated transformation of the plasmid vectors pSUG1::*Tip100* and pSUG1::*dTip100* into *N. tabacum* cv. SR1 plants was performed as previously described (Sugita et al. 1999). After co-cultivation of leaf segments with *A. tumefaciens* LBA4404 (Hoekema et al. 1983) containing the appropriate plasmid on solid hormone-free MS medium containing 50 mg/l acetosyringone for 3 days, the explants were transferred to solid SIM medium (MS

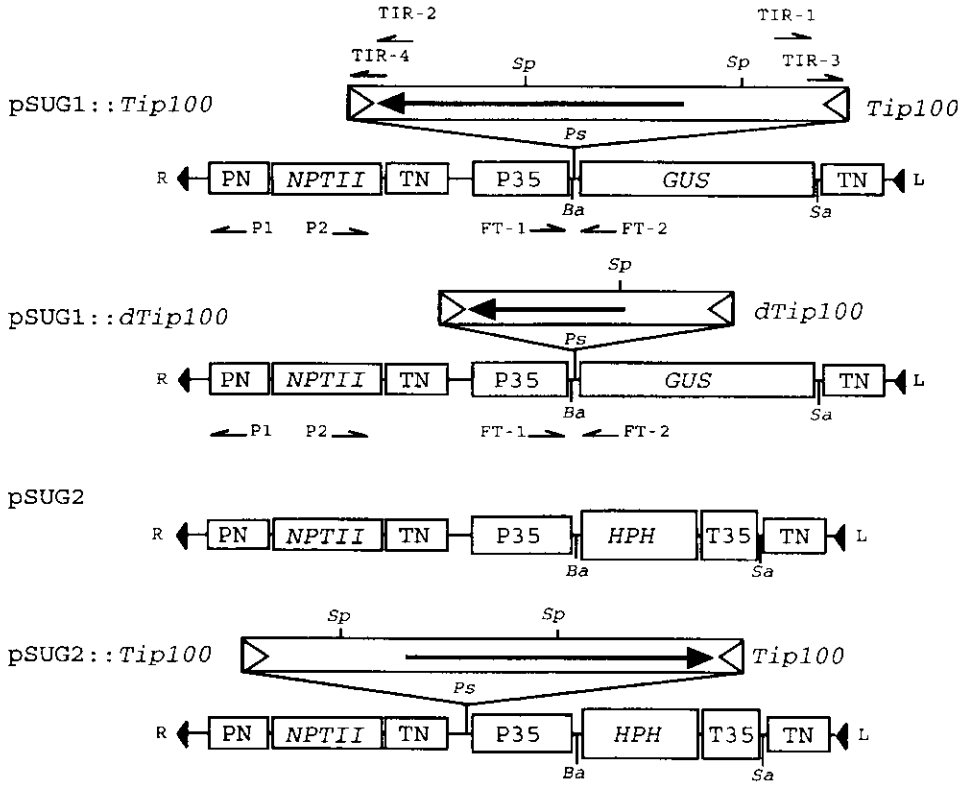


Fig. 2 Structures of the T-DNA regions in the plasmid vectors used. The *filled arrowheads* labeled L and R represent the left and the right borders of T-DNA, respectively. The *open triangles* and the *thick horizontal arrows* within the *Tip100* boxes indicate TIRs and ORFs, respectively. The *boxes* marked PN and TN indicate the promoter and the terminator of the nopaline synthase gene, respectively, and the *boxes* labeled P35 and T35 represent the CaMV 35S promoter and the CaMV 35S terminator, respectively. The structural genes for Km^r, GUS and Hm^r are indicated by *NPTII*, *GUS* and *HPH*, respectively. The *thin arrows* below the maps indicate the positions of primers used for PCR or IPCR amplification. Note that both *Tip100* and *dTip100* on the pSUG1 derivatives are flanked by the 8-bp TSDs originating from the *flaked* allele of the *CHS-D* gene (Habu et al. 1998). Restriction sites are: *Ba*, *Bam*HI; *Ps*, *Pst*I; *Sa*, *Sac*I; *Sp*, *Spe*I

medium containing 1 mg/l benzyladenine, 0.1 mg/l naphthalene acetic acid and 100 mg/l kanamycin) containing 500 mg/l ticarcillin to prevent further bacterial growth. One month after infection with *Agrobacterium*, regenerated adventitious buds that were resistant to kanamycin (Km^r) were removed and transferred to hormone-free MS medium containing 100 mg/l kanamycin and 500 mg/l ticarcillin.

To estimate the number of integrated T-DNA copies present in transgenic tobacco plants, we employed the inverse PCR (IPCR) procedure using the primers P1 (5′-CGTTGCGGTTCTGTCAG TTCC-3′) and P2 (5′-TTGTCAAGACCGACCTGTCC-3′) to determine the copy number of the *NPTII* gene for neomycin phosphotransferase (Fig. 2; Does et al. 1991).

To examine the excision of *Tip100* from the introduced *GUS* gene, we carried out a phenotypic assay for *GUS* expression in the leaves of transgenic tobacco plants using the GUS staining procedure previously described by Jefferson et al. (1987). To determine the footprint sequences generated by *Tip100* excision (Fig. 2), the fragments produced by PCR amplification with the primers FT-1 (5′-ACAATCCCACTATCCTTCGC-3′) and FT-2 (5′-GGATAGT CTGCCAGTTCAGT-3′) were cloned into the pGEM T-easy plasmid (Promega) and subsequently sequenced with the primer GUS-2 (5′-TCGCGATCCAGACTGAATGCCC-3′).

Excision of *dTip100* promoted by *Tip100*

Agrobacterium-mediated transformation of pSUG2::*Tip100* or pSUG2 into a transgenic plant (dTip1) carrying one copy of the *NPTII* gene was performed in the same way as the transformation of pSUG1::*Tip100* described above, except that hygromycin (50 mg/l) was used instead of kanamycin (100 mg/l) for selection and the regeneration step on hormone-free MS medium was omitted. Hygromycin-resistant (Hm^r) calli were used for further analyses.

Reintegration of *Tip100*

To determine the sequences at the sites of reintegration of *Tip100*, we used IPCR amplification using the primers TIR-1 (5′-GGC CAA GCCGCCAAGGCCCCTATTGCCTAATAGGC-3′) and TIR-2(5′-CGTGACATGCAAGAGAGACGGTCTAAATT TT-AGT-3′), followed by nested PCR amplification with the primers TIR-3 (5′-GGGCCAAGGCCCCCTGCATATAATATGTGC-3′) and TIR-4 (5′-TTTCAGTATTCTGTCTTAAAT TGTGATT-3′). The positions of these primers within *Tip100* are shown in Fig. 2. Since *Tip100* contains no *Kpn*I site, the genomic DNAs extracted from leaves of the transgenic plants to be examined were digested with *Kpn*I. After circularizing the *Kpn*I fragments by self-ligation, DNA samples were treated with *Pst*I to destroy the amplified DNA fragments originating from the integrated pSUG1::*Tip100* sequences (the initially integrated, non-transposed, *Tip100* element is flanked by the *Pst*I sequences used for construction of pSUG1::*Tip100*). After the first PCR with the primers TIR-1 and TIR-2, the reaction mixture was diluted 100-fold and subjected to nested PCR amplification with the primers TIR-3 and TIR-4. The amplified fragments were cloned into pGEM T-easy and the TSD sequences were determined.

Results and discussion

A phenotypic assay for the excision of *Tip100* from a *GUS* gene introduced into transgenic tobacco plants

To examine whether *Tip100* is an autonomous element, we introduced *Tip100*, inserted in the *GUS* gene, into tobacco plants, and examined whether it can excise from the reporter gene in the transgenic plants. For this test, we also employed the *dTip100* element, in which the ORF of *Tip100* is disrupted by deletion of the internal 1.7-kb *Spe*I fragment, as a control defective element (Fig. 2). Since the inserted *Tip100* and *dTip100* elements efficiently prevent *GUS* expression from pSUG1::*Tip100* and pSUG1::*dTip100*, respectively, the *GUS* gene will be activated if the elements are excised. *Tip100* in pSUG1::*Tip100* and *dTip100* in pSUG1::*dTip100* were introduced into tobacco by *Agrobacterium*-mediated transformation, and Km^r transgenic plants were obtained. Most of these primary transformants (T0) were found to contain only one copy of the integrated *NPTII* gene in their genomes; the rest carried two or three copies of the *NPTII* gene (see Table 1).

Excision of *Tip100* or *dTip100* from the *GUS* gene was examined in leaves from these T0 plants by assaying for GUS activity. As Fig. 3A shows, the leaves of all twelve independent T0 plants with pSUG1::*Tip100* examined displayed rare but clear and small blue, GUS-positive, patches when stained with X-Gluc (5-bromo-chloro-3-indolyl-beta-D-glucuronide), indicating that *Tip100* excision from the *GUS* gene occurred in the transgenic tobacco plants. We noticed that *Tip100* excision tends to occur much more frequently in the plants at a younger stage, and no obvious GUS-positive blue spots are observed to arise in the same plants at later stages of development. Preliminary results suggested that the introduced *Tip100* elements appear to be heavily methylated at later stages in the development of the transgenic plants. A comparable observation was reported for an autonomous *Tam3* element belonging to the *Ac/Ds* family that was introduced into transgenic tobacco plants; stabilization appeared to be due to rapid methylation of the element (Martin et al. 1989). However, GUS-positive spots or sectors could be detected in the selfed T1 progeny derived from the T0 plants (Fig. 3B, C). None of the three transformants with pSUG1::*dTip100* showed any GUS-positive spots at all (data not shown). These results indicate that *Tip100* can be excised from the *GUS* gene in transgenic tobacco plants, whereas *dTip100* cannot. This suggests that *Tip100* is likely to be an autonomous element coding for an active transposase of 808 amino acids.

A phenotypic assay for the *Tip100*-mediated excision of *dTip100* from the *GUS* gene

We asked whether *dTip100*, in which the transposase gene is non-functional due to an internal 1.7-kb deletion, could be excised from the *GUS* gene when the intact *Tip100* element was introduced into cells containing the defective element in the *GUS* gene integrated in the genome. To examine this possibility, we chose the T0 tobacco plant dTip1, which appears to carry one copy of the *dTip100* element inserted in the *GUS* gene in the genome (see Table 2). *Tip100* in pSUG2::*Tip100* was introduced into the dTip1 plant by *Agrobacterium*-mediated transformation. All of the six independently obtained Hm^r calli displayed a small but significant number of GUS-positive blue cells when stained with X-Gluc (data not shown). No such blue cells could be detected in the Hm^r calli obtained by introduction of the control vector pSUG2. The results further support the notion that *Tip100* is an autonomous element and is able to act on the non-autonomous *dTip100* element in the heterologous tobacco system.

Footprints generated by excision of *Tip100* and *dTip100*

One of the characteristics of plant transposable DNA elements, including those of the *Ac/Ds* family, is their ability to generate small sequence alterations at excision sites, which are referred to as footprints (Kunze et al. 1997). To obtain further evidence for transposition of the *Tip100* element in tobacco, we determined the sequences of the *Tip100* excision sites to examine whether footprints had been generated. The 450-bp fragments containing the *Tip100* excision site were amplified by PCR and cloned into plasmids. All 12 T0 plants with pSUG1::*Tip100* that displayed the GUS-positive patches yielded amplified fragments, while no amplification products were detected in the transgenic plants with pSUG1::*dTip100*. Since these amplified fragments are likely to contain different footprints resulting from independent *Tip100* excisions, we determined a total of 153 excision sequences consisting of at least 10 excision sequences from each of the 12 different transgenic plants obtained (Table 1). Sixteen different sequences were detected and fifteen of them showed the characteristics of footprints. Of these, the two most frequently occurring sequences, Footprints 1 and 2, have a 1-bp deletion at either end of the original 8-bp TSD. *Tip100* in pSUG1::*Tip100* is flanked by the 8-bp TSDs originating from the *CHS-D* intron in the common morning glory (see Materials and methods), and these two footprints were previously observed following *Tip100* excision from the *flaked* allele of the *CHS-D* gene (Habu et al. 1998). Footprint 16 contains one intact TSD followed by a 264-bp segment of the 5′ end of *Tip100* linked to the the other TSD which has a 2-bp deletion; the remaining 3.7 kb of *Tip100* sequence including its 3′ end has been deleted. A similar aberrant excision structure was found at the $A^{rev}3$ allele in a revertant of the mutable a^f line of the common morning glory (Habu et al. 1998).

To examine whether footprints were also generated at sites of *Tip100*-mediated excision of the *dTip100* element

Table 1 Footprint sequences generated by *Tip100* excision

Footprint[a]	Sequence[b]			Number of footprints found (copy number of T-DNA)[c]												N[d]	T[d]
	5′ Flank	Insert	3′ Flank	Tip1	Tip 2	Tip 3	Tip 4	Tip 5	Tip 6	Tip 7	Tip 8	Tip 9	Tip 10	Tip 11	Tip 12		
Original	ccccctgcag **CACGTATG**	*Tip100*	**CACGTATG**ctgcag	(3)	(1)	(1)	(1)	(1)	(1)	(3)	(2)	(1)	(N. D.)	(1)	(1)		
1	ccccctgcag **CACGTAT**		**CACGTATG**ctgcag	3	4	5	20	8	7	5	3	8	5	1	7	76	12
2	ccccctgcag **CACGTATG**		**ACGTATG**ctgcag	3	6	1	–	–	1	2	–	–	4	16	3	36	8
3	ccccctgcag		**CACGTATG**ctgcag	3	2	1	–	2	–	–	2	1	–	–	–	11	6
4	ccccctgcag **CACGTA**		**TATG**ctgcag	–	–	–	–	–	–	2	1	–	–	–	–	2	1
5	ccccctgcag **CACGTATG**		cag	–	–	–	–	–	2	–	–	–	–	–	–	3	2
6	c		g	–	–	–	–	–	2	–	–	–	–	1	–	3	2
7	ccccctgcag **CACGTA**		**ACGTATG**ctgcag	–	–	–	–	–	–	–	–	–	1	–	–	1	1
8	ccccctgcag **CACGTA**		tgcag	–	–	–	–	–	–	–	–	1	–	–	–	1	1
9	ccccctgcag **CACGT**		**ACGTATG**ctgcag	1	–	–	–	–	–	–	–	–	–	–	–	1	1
10	ccccctgcag **CACGTATG**		**CACGTATG**ctgcag	–	–	–	–	–	–	–	–	–	–	1	–	1	1
11	ccccctgcag **CACGTA**	GTG	**ACGTATG**ctgcag	–	–	–	–	1	–	1	–	–	–	–	–	2	2
12	ccccctgcag **CACG**	GTG	**ACGTATG**ctgcag	–	–	2	–	–	–	–	–	–	–	–	–	2	1
13	ccccctgcag **CACGTAT**	CAT	**CGTATG**ctgcag	–	–	–	–	–	1	–	–	–	–	–	–	1	1
14	ccccctgcag **CACGTAT**	CA	**ATG**ctgcag	–	–	1	–	–	1	–	–	–	–	–	–	2	2
15	ccccctgcag **CACGTATG**	ATACGTGCTGC AGGTG	**ACGTATG**ctgcag	–	–	–	–	–	–	–	8	–	1	–	–	9	2
16	ccccctgcag **CACGTA**	CG (Δ*Tip100*)	**CACGTATG**ctgcag	–	–	2	–	–	–	–	–	–	–	–	–	2	1
Number of clones sequenced				10	12	12	20	11	14	10	14	10	11	19	10	153	

[a]Footprint sequences found in the 12 transgenic plants transformed with pSUG1::*Tip100*. The original sequence shown in the *first row* represents the flanking sequences of *Tip100* in pSUG1::*Tip100*

[b]TSDs and their flanking sequences are shown in *bold upper case* and in *lower case letters*, respectively. Deleted nucleotides are omitted, and newly generated sequences within the footprints are shown in *the central column*

[c]The copy number of the *NPTII* gene in each transgenic plant is indicated in *parentheses*. N.D. indicates that the copy number was not determined. From each plant, at least 10 clones containing footprints were sequenced. The data represent the numbers of footprints found in each of the plants indicated

[d]N and T denote the total numbers of clones and transgenic plants carrying the given footprint, respectively

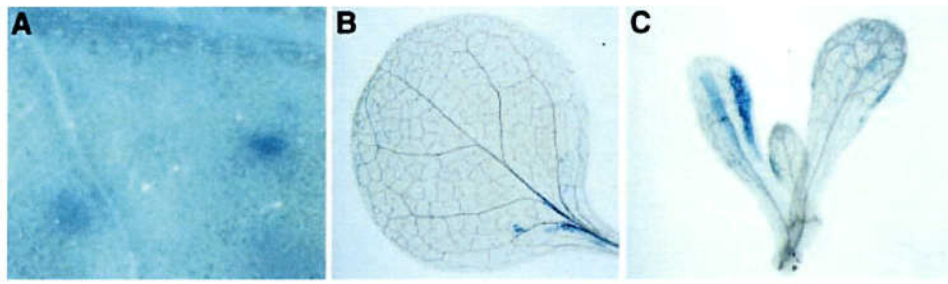


Fig. 3A–C GUS-positive patches observed in leaves of transgenic tobacco plants transformed with pSUG1::-*Tip100*. **A** T0 plant. **B, C** T1 plants

Table 2 Footprint sequences generated by *Tip100*-mediated excision of *dTip100*

Footprint[a]	Sequence			Number of footprints found			N[b]	T[b]
	5′ Flank	Insert	3′ Flank					
Original	ccccctgcag**CACGTATG**	*dTip100*	**CACGTATG**ctgcag	dTip1-1	dTip1-2	dTip1-3		
d1	ccccctgcag**CACGTAT**		**CACGTATG**ctgcag	2	4	2	8	3
d2	ccccctgcag**CACGTATG**		**ACGTATG**ctgcag	1	1	–	2	2
d3	ccccctgcag		**CACGTATG**ctgcag	–	1	–	1	1
Number of clones sequenced				3	6	2	11	

[a]Footprint sequences found in the three Hm^r calli derived from the transgenic plant dTip1 that were transformed with pSUG2::-*Tip100* , which carries one copy of the T-DNA (*NPTII* gene). The original sequence shown in the *top row* represents the flanking sequences of *dTip100* in pSUG1::*dTip100*

[b]N and T denote the total numbers of clones and Hm^r transformants carrying the given footprint, respectively. The data are displayed as in Table 1

in the Hm^r calli derived from the dTip1 plant, we chose three independent GUS-positive calli transformed with pSUG2::*Tip100*. All of them gave rise to the 450-bp fragments containing *dTip100* excision sites. We determined 11 footprints from these GUS-positive calli (Table 2) and found no significant difference between the footprints generated by excision of *dTip100* and those left behind on excision of *Tip100*. Therefore, we can conclude that excision of *Tip100* generates footprints characteristic of transposable elements belonging to the *Ac/Ds* family. A 450-bp fragment could not be amplified from the Hm^r calli transformed with pSUG2.

```
tcttattaGGCAAGAC : Tip100 : GGCAAGACaccatggg
aggtcatgGGTTTGGC : Tip100 : GGTTTGGCttttgggt
tgaacatcGTGACGGG : Tip100 : GTGACGGGttgagcag
atttaattCTCACGGC : Tip100 : CTCACGGCttaaaccc
aaatgtgcCATGTGCT : Tip100 : CATGTGCTgaaaggaa
ctacagccCTTAAGAC : Tip100 : CTTAAGACgaacatgc
tcaaagtcTTCGTCAT : Tip100 : TTCGTCATccactgct
tcttgaggCTGCTGTG : Tip100 : CTGCTGTGagagaatg
```

Fig. 4 Sequence analysis of *Tip100* reinsertion in transgenic tobacco plants. Newly generated TSDs and their flanking sequences are shown in *bold upper case* and *lower case letters*, respectively

Reintegration of *Tip100* into the tobacco genome

To examine whether the excised *Tip100* elements were reintegrated into new sites in the genomes of the transgenic tobacco plants, we characterized the flanking sequences of *Tip100* insertion sites that differed from the original site in the *GUS* gene. Since excision of *Tip100* occurs at rather low frequencies, the reintegration of *Tip100* was also expected to be a relatively rare event. We thus applied the IPCR procedure followed by the nested PCR amplification technique with appropriate primers (Fig. 2), and were able to identify eight different reintegration sites from four transgenic plants among the selfed T1 progeny derived from the GUS-positive T0 plants (Fig. 4). As expected, *Tip100* generates 8-bp TSDs upon reintegration.

Based on the results obtained in transgenic tobacco plants, we therefore conclude that *Tip100* is an autonomous element coding for an active transposase of 808 amino acids. Apart from the autonomous *Ac* element (Baker et al. 1986), several other plant transposable elements of the *Ac/Ds* family, including *Tam3* of snapdragon and *Tag1* of Arabidopsis, are regarded as autonomous because they are capable of somatic transposition in transgenic tobacco plants (Haring et al. 1989, 1991; Martin et al. 1989; Frank et al. 1997). To our knowledge, *Tip100* is the first autonomous element to be characterized among the various transposable elements known in *Ipomoea* species (Hoshino et al. 2001).

Tip100 and *Modulator* in the common morning glory

Based on the observations that (1) lines of the common morning glory carrying the alleles a^f and ygl^m tend to show

a striking similarity in both the timing and frequency of variegation in flowers and leaves, and (2) these variegated phenotypes are also generally heritable by their progeny, Kasahara (1956) postulated that there must be a factor (called *Modulator*) that acts on both a^f and vgl^m alleles, and that the timing and frequency of the variegation in flowers and leaves are determined by the state of the *Modulator* activity, which is generally heritable. Indeed, the *CHS-D* gene in the stable white-flower mutant carrying the mutable a^f allele without active *Modulator* is identical to that in mutable a^f lines displaying variegated flowers (Fig. 1), whereas the white-flower line carrying the stable *a* allele was found to have two copies of *Tip100* integrated into the *CHS-D* gene (Habu et al. 1998). Although the simplest explanation for these observations is that *Tip100* and *Modulator* represent a non-autonomous and an autonomous element, respectively, our present finding that *Tip100* is an autonomous element is clearly incompatible with this assumption. The genome of the common morning glory contains about 100 copies of *Tip100*-related elements and the majority of them appear to be structurally very similar to *Tip100* (N. Ishikawa, unpublished results), suggesting that a significant number of autonomous *Tip100* copies are present in the genome of *I. purpurea*. Similarly, it has been observed that most copies of *Tam3* in the genome of snapdragon have highly conserved structures of nearly the same size (Kishima et al. 1999). Interestingly, both *Tip100* and *Tam3* carry single ORFs for active transposase that lack intron sequences (Hehl et al. 1991; Habu et al. 1998).

A possible alternative hypothesis is that *Modulator* may be an element that affects the transposition activity of the autonomous *Tip100* element. In the *Ac/Ds* family of elements, *Stabilizer* in snapdragon and the *IAE* loci in Arabidopsis have been reported to control the transposition activity of the autonomous elements, *Tam3* and *Ac*, respectively (Carpenter et al. 1987; Jarvis et al. 1997). Recently, mutations in the *DDM1* gene (encoding a SWI2/SNF2 chromatin-remodeling factor) that lead to hypomethylation of DNA have also been reported to result in activation of DNA transposable elements and developmental abnormalities in Arabidopsis (Kakutani et al. 1996; Miura et al. 2001; Singer et al. 2001). Since no apparent developmental abnormalities have been observed in lines of the common morning glory that display highly variegated flowers, it is, however, unlikely that *Modulator* is directly related to the *DDM1* function.

Acknowledgements We thank Y. Habu for comments at the initial stage of this study and A. Hoshino for invaluable discussion. This work is supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology in Japan. Y. J-H. received a Research Fellowship awarded by the Japan Society for the Promotion of Science for Young Scientists.

References

Baker B, Schell J, Lörz H, Fedoroff N (1986) Transposition of the maize controlling element "*Activator*" in tobacco. Proc Natl Acad Sci USA 83:4844–4848

Barker EE (1917) Heredity studies in the morning glory (*Ipomoea purpurea*). Bull Cornell Univ Agric Exp Sta 392:1–38

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

Does MP, Dekker BMM, de Groot MJA, Offringa R (1991) A quick method to estimate the T-DNA copy number in transgenic plants at an early stage after transformation, using inverse PCR. Plant Mol Biol 17:151–153

Durbin ML, Denton AL, Clegg MT (2001) Dynamics of mobile element activity in chalcone synthase loci in the common morning glory (*Ipomoea purpurea*). Proc Natl Acad Sci USA 98:5084–5089

Epperson BK, Clegg MT (1987) Instability at a flower color locus in the morning glory. J Hered 78:346–352

Frank MJ, Liu D, Tsay YF, Ustach C, Crawford NM (1997) *Tag1* is an autonomous transposable element that shows somatic excision in both Arabidopsis and tobacco. Plant Cell 9:1745–1756

Habu Y, Hisatomi Y, Iida S (1998) Molecular characterization of the mutable *flaked* allele for flower variegation in the common morning glory. Plant J 16:371–376

Haring MA, Gao J, Volbeda T, Rommens CMT, Nijkamp HJJ, Hille J (1989) A comparative study of *Tam3* and *Ac* transposition in transgenic tobacco and petunia plants. Plant Mol Biol 13:189–201

Haring MA, Teeuwen-de Vroomen MJ, Nijkamp HJJ, Hille J (1991) *Trans*-activation of an artificial d*Tam3* transposable element in transgenic tobacco plants. Plant Mol Biol 16:39–47

Hehl R, Nacken WKF, 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

Hisatomi Y, Yoneda Y, Kasahara K, Inagaki Y, Iida S (1997) DNA rearrangements at the region of the dihydroflavonol 4-reductase gene for flower pigmentation and incomplete dominance in the morning glory carrying the mutable *flaked* mutation. Theor Appl Genet 95:509–515

Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA (1983) A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303:179–180

Hoshino A, Johzuka-Hisatomi Y, Iida S (2001) Gene duplication and mobile genetic elements in the morning glories. Gene 265:1–10

Iida S, Hoshino A, Johzuka-Hisatomi Y, Habu Y, Inagaki Y (1999) Floricultural traits and transposable elements in the Japanese and common morning glories. Ann New York Acad Sci 870:265–274

Imai Y, Tabuchi K (1935) Recurrent mutation in the *flaked* alleles of *Pharbitis purpurea*. J Genet 35:433–446

Izawa T, Miyazaki C, Yamamoto M, Terada R, Iida S, Shimamoto K (1991) Introduction and transposition of the maize transposable element *Ac* in rice (*Oryza sativa* L.). Mol Gen Genet 227:391–396

Jarvis P, Belzile F, Page T, Dean C (1997) Increased *Ac* excision (*iae*): *Arabidopsis thaliana* mutations affecting *Ac* transposition. Plant J 11:907–919

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: β-glucuronidase as a sensitive and vertile gene fusion marker in higher plants. EMBO J 6:3901–3907

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ (1996) Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci USA 93:12406–12411

Kasahara K (1956) Genetic studies on mutable genes in *Ipompea purpurea* (in Japanese). Res Lib Arts Hosei Univ 1:1–13

Kishima Y, Yamashita S, Martin C, Mikami T (1999) Structural conservation of the transposon *Tam3* family in *Antirrhinum majus* and estimation of the number of copies able to transpose. Plant Mol Biol 39:299–308

Kunze R, Saedler H, Lönnig WE (1997) Plant transposable elements. Adv Bot Res 27:331–470

Martin C, Prescott A, Lister C, MacKay S (1989) Activity of the transposon *Tam3* in *Antirrhinum* and tobacco: possible role of DNA methylation. EMBO J 8:997–1004

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

Shiokawa K, Inagaki Y, Morita H, Hsu TJ, Iida S, Noguchi H (2000) The functional expression of the *CHS-D* and *CHS-E* genes of the common morning glory (*Ipomoea purpurea*) in *Escherichia coli* and characterization of their gene products. Plant Biotechnol 17:203–210

Singer T, Yordan C, Martienssen RA (2001) Robertson's *Mutator* transposons in *A. thaliana* are regulated by the chromatin-remodeling gene *Decrease in DNA Methylation (DDM1)*. Genes Dev 15:591–602

Sugita K, Matsunaga E, Ebinuma H (1999) Effective selection system for generating marker-free transgenic plants independent of sexual crossing. Plant Cell Rep 18: 941–947

Acknowledgment

I deeply thank Professor Shigeru Iida for all the valuable discussions and critical comments. I thank Laurel Caitlin Coberly (Duke University) for providing seeds of the common morning glory collected from fields in North Carolina and Koichi Sugita (Nippon Paper Industries) for generating transgenic tobacco plants. I also thank Jeong-Doo Choi, Atushi Hoshino, Yasuyo Johzuka-Hisatomi, and Yoshiki Habu for discussions. I also acknowledge technical assistance of Miwako Matsumoto, Hisayo Asao, Tomoko Mori, and Kyoko Ikegaya. Last but not least, I thank my parents and my elder brother for supporting me all the time.