Plant Biotechnology 24, 467–479 (2007)

Original Paper

Expression analysis of gene trap lines and mapping of donor loci for *Dissociation* transposition in *Arabidopsis*

Tomohiko Kato[a,*], Satoshi Tabata, Shusei Sato

Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818, Japan
* E-mail: tomohiko-kato@ojipaper.co.jp Tel: +81-595-85-0122 Fax: +81-595-85-1528

Received July 11, 2007; accepted July 27, 2007 (Edited by D. Shibata)

Abstract In order to establish a system for characterizing gene function utilizing information obtained from genome sequences, we generated T-DNA insertion lines using a newly constructed binary vector. The vector carries a *uidA* [β-glucuronidase (*GUS*)] reporter gene which allows the promoter activity of the inserted genes to be monitored, a transposable element *Dissociation* (*Ds*) for targeted insertional mutagenesis, and the *cis* sequences required for *Agrobacterium*-mediated transformation. Approximately 8% of the 20,000 lines tested for GUS activity exhibited positive staining. Staining was detected in various organs including roots, leaves, stems, flowers, and siliques. These lines are therefore useful resources for analyzing tissue or organ specific gene expression. We have included the GUS expression patterns on our web site (http://www.kazusa.or.jp/ja/plant/GUS/). To establish a system for targeted insertional mutagenesis, integration sites of the T-DNAs in 140 lines with a single T-DNA insertion were mapped on the genome. The T-DNA contains the *Ds* element; therefore these 140 lines could be used as donor loci for *Ds* transposition. The integration sites were almost evenly distributed on all five chromosomes except for the nuclear organizer regions of chromosomes 2 and 4 and the centromeric regions. The *Ds* element in one line was transposed in combination with an *Activator* (*Ac*) element. In about half of the transposed lines, the *Ds* elements were reinserted within about 1 M bp from the donor locus. These results indicate that the donor loci for *Ds* transposition that were mapped in this study are a valuable resource for targeted mutagenesis throughout the *Arabidopsis* genome.

Key words: *Arabidopsis thaliana*, *Ds*, gene trap, T-DNA.

The sequencing of the Arabidopsis genome was completed in the year 2000, at which time 25,498 genes were assigned to the genome (Arabidopsis Genome Initiative 2000). The TAIR7 genome release has recently become available, and this contains 26,819 protein coding genes on the *Arabidopsis* genome (TAIR: http://www.arabidopsis.org/). The next challenge is to determine the function of thousands of the uncharacterized genes. One method of analyzing gene function using information obtained from the genome sequence is to use a reverse genetics approach. In plants, several methods have been adapted for loss of function mutations, and insertional mutagenesis is a widely used method to identify knock-out lines. If a large number of insertional lines are available, one can isolate the line carrying an insertion in any gene of interest by PCR-based screening (McKinney et al. 1995; Krysan et al. 1996; Winkler et al. 1998; Meissner et al. 1999) or using databases of the flanking sequences of the insertion sites (Parinov et al. 1999; Ito et al. 2002; Raina et al. 2002; Samson et al. 2002; Sessions et al. 2002; Alonso et al. 2003; Pan et al. 2003; Kuromori et al. 2004; Woody et al. 2007).

Insertion lines have been generated in *Arabidopsis* by the introduction of T-DNAs or transposable elements as mutagens. There are several reports describing the generation of a large number of insertional lines (Feldmann 1991; Wisman et al. 1998; Campisi et al. 1999; Krysan et al. 1999; Speulman et al. 1999; Tissier et al. 1999; Galbiati et al. 2000; Weigel et al. 2000; McElver et al. 2001; Marsch-Martinez et al. 2002; Rios et al. 2002; Nakazawa et al. 2003). Improved transformation of *Arabidopsis* using *Agrobacterium* (Bechtold and Pelletier 1998; Clough and Bent 1998) has allowed the production of a large number of T-DNA insertional lines. A transposable element is useful to disrupt neighboring genes because the elements tend to transpose near a region from the donor loci (Dooner and Belachew 1989; Jones et al. 1990; Dooner et al. 1991; Bancroft and Dean 1993; Carroll et al. 1995; Machida et

[a] Present address: Forestry Research Institute, Oji Paper Company Co. Ltd. Kameyama, Mie 519-0212, Japan

Abbreviations: *Ac*, *Activator*; BLAST, basic local alignment search tool; *Ds*, *Dissociation*; GUS, β-glucuronidase; IPCR, inverse polymerase chain reaction; LUC, luciferase; NAM, napthalene acetamide; NPTII, neomycin phosphotransferase II; rbcS3B, ribulose bisphosphate carboxylase small chain 3b; TAIL-PCR, thermal asymmetric interlaced polymerase chain reaction; X-gluc, 5-bromo-4-chloro-3-indolyl β-D-glucuronide

This article can be found at http://www.jspcmb.jp/

al. 1997). In *Arabidopsis* there are a large fraction of tandemly repeated genes (approximately 17%) and smaller genes whose sizes are less than 0.5 kb (10% to 20% of all genes). When a target gene is small in size or possesses tandem repeats, it is useful to use a transposable element near the target gene to disrupt it. Tantikanjana et al. (2004) reported the generation of double-knockout mutants of two tandemly duplicated cytochrome P450 genes using secondary transpositions. A targeted *Ds*-tagging strategy has also been reported to generate a high allelic diversity (Creff et al. 2006). If the donor loci for transposition are widely distributed on the genome, one can use the locus near the target gene to disrupt it. For this targeted insertional mutagenesis, *Dissociation* launch pads were mapped in the *Arabidopsis* genome (Muskett et al. 2003; Zhang et al. 2003). Recently the WiscDsLox collection was generated which can be used for reverse-genetic analysis (Woody et al. 2007).

The identification of the expression pattern of each gene provides information regarding the function of the gene. Gene, enhancer and promoter traps are attractive approaches to dissect gene function by the analysis of expression patterns (Springer 2000). In these approaches a reporter gene is randomly introduced into the plant genome and the expression pattern of the genes at the targeted site are analyzed by examining the localization of the reporter gene. In this way, characterization of the localization of novel genes can provide some insight into the function of the gene. Several groups have produced a large number of gene, enhancer and promoter trap lines in *Arabidopsis* (Fedoroff and Smith 1993; Topping et al. 1994; Klimyuk et al. 1995; Sundaresan et al. 1995; Campisi et al. 1999; Yamamoto et al. 2003; Alvarado et al. 2004). Many genes have been identified using these methods and a large scale analysis of expression patterns has been reported (Nakayama et al. 2005; Liu et al. 2005; Nagawa et al. 2006). Gene trap lines contain a reporter gene fused to a splice acceptor site, and in this situation the reporter gene is expressed if it inserts into the transcribed region of a gene with the proper orientation. In enhancer trap lines, a reporter gene is fused to a minimal promoter. If the reporter gene is integrated near enhancers, expression of the reporter gene is detected. Gene trap lines therefore have several advantages over enhancer trap lines. Firstly, a chromosomal gene that regulates the expression of the reporter gene can be identified by isolating flanking regions of the reporter gene. In this situation the reporter gene is only expressed when it is inserted in the transcribed region of the gene. In the case of enhancer trap lines, expression of the reporter gene is affected by an enhancer of a chromosomal gene. The gene next to the reporter gene does not always regulate the expression of the reporter gene. Secondly, if the expression of a reporter gene is detected in a gene trap line there is a possibility that the chromosomal gene regulating the expression is disrupted by the T-DNA insertion.

In the present study, we constructed a new T-DNA vector carrying the *GUS* reporter gene for gene trap analysis and the *Ds* element for targeted mutagenesis, and subsequently produced 51,000 gene trap lines. In addition to screening for T-DNA insertion lines by PCR, we analyzed the expression patterns of the *GUS* reporter gene in 20,000 gene trap lines. The GUS staining patterns detected in the gene trap lines can be used to identify genes with tissue- or organ-specific expression. Despite a large number of insertion populations, it is difficult to disrupt every *Arabidopsis* gene on the genome, especially small genes or tandem duplicated genes. We therefore introduced a *Ds* transposition system for targeted tagging. One hundred and forty integration sites of the T-DNAs in the gene trap lines were mapped on the genome. The T-DNA contains the *Ds* element therefore these lines can be used as donor loci for transposition. The integration sites of the lines were distributed almost evenly on all five chromosomes. The *Ds* element in one gene trap line was transposed in combination with an *Ac* element. Finally we discuss the use of our system for targeted insertional mutagenesis.

Materials and Methods

Construction of pGTAC-LUS

The promoter region of the nopaline synthase gene (*Pnos*) and neomycin phosphotransferase gene (*NPTII*) derived from pBI101 were introduced into the pBluescript vector. The *Ds* element, cauliflower mosaic virus 19S promoter (*19S*), *aph4* gene (*Hyg*) which confers hygromycin resistance, and the nopaline synthase terminator (*Tnos*) derived from *Ds*2 389-13 and constructed by Fedoroff and Smith (1993), were inserted between *Pnos* and the *NPTII* gene in the vector. The splice donor and acceptor sequence from the first intron of the *Arabidopsis* ribulose bisphosphate carboxylase small chain 3B precursor gene (*rbcS3B*: At5g38410) was fused to the *GUS* gene in the pSLG2 vector (Kato et al. 1991), and then the terminator region of the temperature-sensitive omega-3 fatty acid desaturase, the chloroplast precursor gene (At5g05580) and the phosphoribosylanthranilate isomerase gene (At5g05590) on the P1 clone, MOP10, was fused to the *GUS* gene. This *GUS* gene was inserted next to the *NPTII* gene in the vector with reverse orientation. *Pnos*, the *Ds* element, the *aph4* gene (*Hyg*), the *NPTII* gene and the *GUS* gene in the vector were introduced into the pYLTAC7 vector (Liu et al. 1999). The splice donor and acceptor sequence of the first intron in the *Arabidopsis rbcS3B* gene was fused to the luciferase gene (*LUC*) in the pDTGFIG28 (NIPPON GENE, Japan) vector, and then the terminator region of the pyruvate kinase gene (At5g08570) and the unknown protein gene (At5g08580) on the P1 clone, MAH20, were fused to the *LUC* gene. This *LUC* gene was introduced next to the hygromycin resistance gene (*Hyg*) in the newly constructed pYLTAC7

vector with the same orientation. The synthesized vector was named pGTAC-LUS.

Plant transformation and selection of transgenic Arabidopsis plants

Arabidopsis plants (Columbia accession or Ws accession) (T0) were grown at 22°C for about 1 month under a 16-h light, 8-h dark cycle in a growth chamber. Transgenic plants were produced using a vacuum infiltration protocol as described previously (Bechtold and Pelletier 1998), with *Agrobacterium tumefaciens* strain C58C1 (pMP90) harboring pGTAC-LUS. After transformation, the plants were grown for approximately one more month, and then the seeds (T1) were harvested from the plants. T1 seeds were sown on 0.8% agar plates containing 0.5X B5 medium (pH 5.7), 1% (w/v) sucrose, hygromycin B (15 μg ml^{-1}) and cefotaxime (250 μg ml^{-1}), and were grown at 22°C for 2 weeks under a 16-h light, 8-h dark cycle. Transgenic plants that had germinated normally on the plates were moved to 0.8% agar plates that did not contain hygromycin B for one further week, followed by growth at 22°C under a 16-h light, 8-h dark cycle in the growth chamber. Seeds (T2) were harvested from the T1 plants individually, and then stored in a desiccator.

GUS staining of the transgenic lines

About 30 T2 seeds from each line were sown on 0.8% agar plates containing 0.5X B5 medium (pH 5.7), 1% (w/v) sucrose and hygromycin B (10 μg ml^{-1}), and were grown at 22°C for 3 weeks under a 16-h light, 8-h dark cycle. Two plants from each line were stained with a GUS staining solution [6.38 mM $NaPO_4$ pH7.4, 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 0.3% Triton X-100, 20% Methanol, 1.9 mM 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-Gluc)] at 37°C for 40 h. Two other plants from each line were grown for 3 more weeks in the growth chamber. Two inflorescences and two siliques were cut from each plant and the siliques were split and subsequently stained with the GUS staining solution at 37°C for 40 h. After staining, the solution was replaced twice with 70% ethanol to remove the chlorophyll. The staining patterns of the plants were examined under a microscope.

Preparation of genomic DNA from the transgenic lines

Sixteen seeds were grown at 22°C for 3 weeks under a 16-h light, 8-h dark cycle in the growth chamber. Leaves from the 16 plants (100 mg) were collected in a microtube, and then genomic DNA was extracted using *Nucleon PhytoPure*, *plant and fungal DNA extraction kits* (Amersham, England) with a minor modification. Reagent I from the kit was initially added to the leaves, and then the tissues were ground with aluminum beads by vortexing for 7 min. DNA extraction was subsequently carried out according to the protocol.

Isolation and sequencing of the flanking region of the T-DNAs

The flanking regions of the T-DNAs were isolated using two methods: inverse PCR (IPCR) and TAIL-PCR. IPCR was performed as follows: genomic DNA (500 ng) from each line was digested with *Hind* III, and the digested fragments were ligated. The flanking regions of the T-DNAs were amplified by PCR using the primers D0046 and D0055 for the left border, and D0057 and D0053 for the right border. The primer sequences were as follows: D0046; AAGAAAATGCCGATACTTCATTGGC, D0055; CGGCCAATTAGGCCACAAACTATTC, D0057; ATCCAGACTGAATGCCCACAGGCCG, D0053; CCGTGTAATTCTAGAGATTACAGAC. The products of this amplification were reamplified using the primers D0047 and D0056 for the left border, and D0058 and D0054 for the right border. The sequences of these primers were as follows: D0047; GTACATTAAAAACGTCCGCAATGTG, D0056; TCTCTCTCACGTAGATCATCTGGTG, D0058; CTACAGGACGTAACATAAGGGACTG, D0054; GGTCACAATAGTTCCTGAAACTCTG. After amplification, the products were precipitated with PEG, and the pellets were suspended in 20 μl of sterile distilled water.

TAIL-PCR was performed as described previously (Liu et al. 1995) using specific primers for pGTAC-LUS. Genomic DNA (100 ng) from each line was used in the 1st round of PCR. The flanking regions of the T-DNAs were amplified using three primers D0046, D0047 and D0048 for the left border, and D0057, D0058 and D0043 for the right border. The sequences of D0048 and D0043 were as follows: D0048; AACGTCCGCAATGTGTTATTAAG, D0043; TTCCCTTAATTCTCCGCTCATGATC. After amplification, the products were purified using a QIAquick PCR Purification Kit (QIAGEN, Germany).

Sequence reactions were carried out using the primer D0048 for the left border or the primer D0043 for the right border, and sequences of the flanking regions were determined using type 377XL sequencers (Applied Biosystems, USA).

Transposition of the Ds elements and selection of transposed Arabidopsis plants

To mobilize the *Ds* element, crosses were performed between one gene trap line carrying the *Ds* element, KG1870, as male parents and the *Ac* T-DNA line constructed by Fedoroff and Smith (1993), Nae*Ac* 380-6, as female parents. Seeds (F1) from the crossed plants were harvested. F1 seeds were sown on 0.8% agar plates containing 0.5X B5 medium (pH 5.7), 1% (w/v) sucrose and hygromycin B (10 μg ml^{-1}), and were grown at 22°C for 2 weeks under a 16-h light, 8-h dark cycle. Plants that germinated normally were moved to 0.8% agar plates that did not contain hygromycin B for one more week, followed by growth at 22°C under a 16-h light, 8-h dark cycle in the growth chamber. Seeds (F2) were harvested from the F1 plants.

Approximately 60 F2 seeds from each line were sown on 0.8% agar plates containing 0.5X B5 medium (pH 5.7), 1% (w/v) sucrose, kanamycin (25 μg ml^{-1}), and napthalene acetamide (NAM) (10 μM), and were grown at 22°C for 1 week under a 16-h light, 8-h dark cycle. Plants that germinated normally were moved to 0.8% agar plates containing kanamycin (40 μg ml^{-1}) and hygromycin B (10 μg ml^{-1}) for two additional weeks. Plants grown normally were transferred to 0.8% agar plates that did not contain antibiotics for one more week, followed by growth at 22°C under a 16-h light, 8-h dark cycle in the growth chamber. Seeds (F3) were harvested from the F2 plants individually, and then stored in the desiccator.

Isolation and sequencing of the flanking region of the Ds elements

The flanking region of the *Ds* elements was isolated using IPCR. IPCR was performed as described below. Genomic DNA (500 ng) from each line was digested with *Hpa* I, and the digested fragments were ligated. The flanking regions of the T-DNAs were amplified by PCR using the primers D0061, D0063 for the left border, and D0065, D0067 for the right border, the sequences of which were as follows: D0061; GGTTCGAAATCGATCGGGATAAAAC, D0063; AACGCTGCTCTGCCTCTCTCCCAG, D0065; GTTAAATATGAAAATGAAAACGGTAGAGG, D0067; CGCCCAAGTAAGTATCCAGCTGTG. The products of this amplification were reamplified using the primers D0062, D0064 for the left border, and D0066, D0068 for the right border, the sequences of which were as follows: D0062; AAATCGGTTATACGATAACGGTCGG, D0064; GCCAGGCACCACACGAGTAACAGC, D0066; CGACCGTTACCGACCGTTTTCATCC, D0068; GAACGTACGTCCGAATTCCCGATCG. After amplification, the products were precipitated with PEG, and the pellets were suspended in 20 μl of sterile distilled water.

Sequence reactions were carried out using the primer D0062 for the left border or the primer D0066 for the right border, and sequences of the flanking regions were determined using type 377XL sequencers (Applied Biosystems, USA).

Luciferase assay of the transposed lines

About 20 F3 seeds from each line were sown on 0.8% agar plates containing 0.5X B5 medium (pH 5.7) and 1% (w/v) sucrose, and were grown at 22°C for 3 weeks under a 16-h light, 8-h dark cycle. Three plants from each line were sprayed with a 1 mM K-luciferin (Promega, USA) solution containing 0.1% Triton X-100. Luciferase activity was visualized using a Hamamatsu photon-counting video system controlled by Argus50 software.

Results and Discussion

Construction of a T-DNA vector and generation of gene trap lines

Figure 1 shows the structure of the T-DNA region in the newly constructed vector. The vector was designated pGTAC-LUS, and contains the coding region of the *GUS* gene fused to the first intron of the *rbcS3B* gene for the gene trap, the *Ds* element for targeted mutagenesis and the *cis* sequences required for *Agrobacterium*-mediated transformation. The vector also carries a hygromycin resistance gene (*Hyg*) driven by the cauliflower mosaic virus 19S promoter for use as a selectable marker, a kanamycin resistance gene (*NPTII*) which is expressed under the control of the nopaline synthase promoter (*Pnos*) when the *Ds* element is excised, and the coding region of the *LUC* gene fused to the first intron of the *rbcS3B* gene. This *LUC* gene can be used for gene trap analysis when the *Ds* element is transferred into other regions of the genome in combination with an *Ac* element.

The *GUS* gene is expressed when the T-DNA region of the vector is inserted in the transcribed region of a gene with the proper orientation. The *GUS* gene was fused to the intron including the splice donor and acceptor sequences. Two additional splice acceptor sequences were added to transcribe the *GUS* gene in every reading frame. If the T-DNA was inserted in the intron of a chromosomal gene, the transcript was spliced out at the splice acceptor sequences, and the *GUS* gene was subsequently transcribed. If the T-DNA was inserted in the exon of a gene, splicing occurs from the donor site to the acceptor site of the intron, and a fusion protein was produced. The *LUC* gene was also expressed when the transposed *Ds* element was inserted in the transcribed region of a gene with the proper orientation.

We produced a total of 51.000 gene trap lines. Genomic DNA from each line was prepared, and DNA was pooled to facilitate the screening of knock-out lines. Each DNA pool contained genomic DNA derived from 384 transgenic lines. Several knock-out mutants were isolated by PCR from the pooled DNA (data not shown).

Analysis of GUS expression patterns in gene trap lines

In order to comprehensively analyze gene expression, 20,000 gene trap lines were stained with X-gluc, a substrate for the GUS protein. Out of 20,000 lines, 1582 (about 8% of the lines tested), exhibited positive GUS staining. The frequency of staining was lower than that reported previously for gene trap lines using a

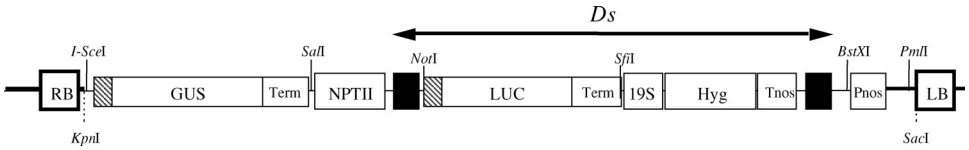

Figure 1. A schematic representation of the T-DNA region of pGTAC-LUS. Right border (RB) and left border (LB) regions are shown. The black boxes indicate the *Ds* elements, and transposed region are represented by an arrow. The splice donor and acceptor sequence of the 1st intron of the *rbcS3B* gene is represented by the hatched box, and was fused in front of the β-glucuronidase gene (*GUS*). The splice donor and acceptor sequence of the 1st intron of the *rbcS3B* gene was also fused in front of the luciferase gene (*LUC*). Terminator sequences (*Term*) are shown. The neomycin phosphotransferase gene (*NPTII*) is expressed under the control of the nopaline synthase promoter (*Pnos*) when the *Ds* element is excised. The *aph4* gene (*Hyg*) confers hygromycin resistance and is expressed under the control of the cauliflower mosaic virus 19S promoter (*19S*). The nopaline synthase terminator (*Tnos*) is indicated.

Table 1. Summary of GUS expression patterns.

organ	number of staining lines	number of staining lines restricted to[a]
root	359	82
hypocotyl	118	5
cotyledon	42	11
axillary bud	78	13
flower bud	348	33
leaf	693	154
stem	555	72
sepal	284	13
petal	144	18
stamen	245	51
carpel	454	54
silique	324	18
seed	234	102

[a] The number of lines exhibiting GUS staining restricted to one organ.

transposable element (26%) (Sundaresan et al. 1995). The lower frequency in our system may be due to the occurrence of rearrangement, tandem duplication, or deletion during the T-DNA integration. Another possibility is that the time of staining for GUS activity was different between the two analyses.

Table 1 shows the number of lines stained in the organs examined and the number of lines in which staining was restricted to the organ specified. Six hundreds twenty-six out of 1,582 lines showing GUS expression (39.6%) exhibited positive staining in one organ under our staining conditions. Analyzing those lines was useful for identifying tissue or organ specific promoter activities. Other lines showed staining in multiple organs. Many *Arabidopsis* genes are expressed in different organs or at different developmental stages. One possible explanation for this is that a gene is expressed in the same organelle in different organs. Another possibility is that a gene has the same function in different organs, for example a gene functional in transport or cell division. Several typical GUS expression patterns are shown in Figure 2. There were many lines where only one bud of the plant was stained with GUS, but we did not count those lines as positive staining; we also eliminated lines with faint staining patterns. We identified genes involved in vascular development using these gene trap lines (Nagawa et al. 2006), thus demonstrating that analyzing the gene trap lines is an effective method of isolating genes expressed in specific organs. GUS expression patterns in this study have been placed on our web site (http://www.kazusa.or.jp/ja/plant/GUS/).

We determined the flanking sequences of the T-DNAs in several GUS positive lines to identify genes which regulate GUS expression. In some lines, the *GUS* reporter gene was inserted in the transcribed region of a gene with the proper orientation, but in other lines, the *GUS* gene was inserted in the transcribed region of a gene with the reverse orientation or inserted in the intergenic region. Our recent results identified similar insertion patterns (Nagawa et al. 2006); these results have also been described in previous studies (Yamamoto et al. 2003; Nakayama et al. 2005; Stangeland et al. 2005). There is a possibility that the *GUS* gene was tagging an antisense transcript found in many annotated genes (Yamada et al. 2003). In the line that contained the reporter gene inserted in the intergenic region, the *GUS* gene may be tagging an unannotated gene or may be controlled by a 'criptic' promoter (Plesch et al. 2000). Another possible explanation is that the *GUS* gene may be inserted in a genome region for non-coding RNA such as a micro RNA precursor. Further analyses will be needed to elucidate the mechanism defining these observations.

Mapping of donor loci for Ds transposition

Despite the large number of insertion populations, it was difficult to disrupt all of the *Arabidopsis* genes, especially in the case of small genes or tandem duplicated genes. The transposon approach is suitable for targeted insertional mutagenesis because a transposon preferentially moves to closely linked sites. To establish a system for targeted mutagenesis, integration sites of the T-DNAs in the gene trap lines were mapped on the genome. The T-DNA contains a *Ds* element therefore the gene trap lines can provide the donor loci for *Ds* transposition. Southern blot analyses were initially performed using the DNA derived from the gene trap lines and the *GUS* gene or the *Hyg* gene as a probe. We selected the lines that showed a single band (data not shown) as these lines were thought to possess a single copy T-DNA in the single locus on the genome. One hundred and forty integration sites were determined by sequencing the flanking regions and these, together with the results from a BLAST search, are shown in Table 2. Figure 3 maps the positions of the inserted T-DNAs on the genome. These positions were almost evenly distributed on all five chromosomes except for the centromeric regions and the nuclear organizer regions of chromosomes 2 and 4. The 140 lines in which the integration sites of the T-DNA were mapped could therefore be used as donor loci for *Ds* transposition.

Transposition of the Ds elements

To test the feasibility of transposition, one line carrying the *Ds* element, KG1870 (male parent), was crossed with another line that contained the modified *Ac* element, Nae*Ac* 380-6 (female parent). This line contained the modified *Ac* element carrying a complete transposase gene, *NPTII*, and a 35S-*tms2* gene as a negative selectable marker. The *tms2* gene that was derived from the *Agrobacterium tumefaciens* Ti plasmid encodes an indoleacetamide hydrolase. In the presence of NAM, an

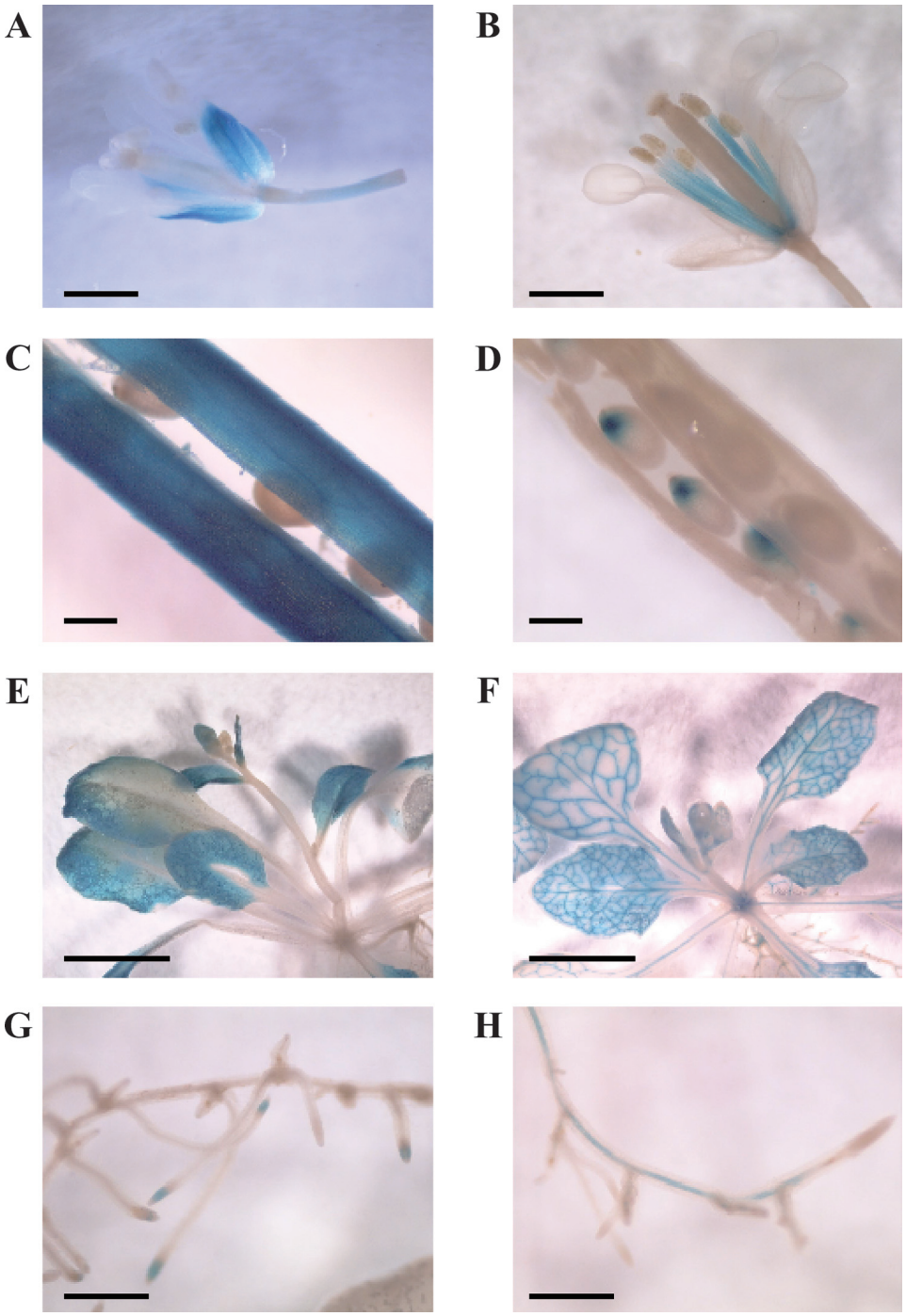

Figure 2. GUS expression patterns in gene trap lines. (A) The KG2360 line exhibits GUS staining in sepals. (B) The KG4208 line exhibits GUS staining in stamens. (C) The KG2932 line exhibits GUS staining in silique. (D) The KG4017 line exhibits GUS staining in seeds. (E) The KG4056 line exhibits GUS staining in leaves. (F) The KG5073 line exhibits GUS staining in veins. (G) The KG3574 line exhibits GUS staining in root tips. (H) The KG4032 line exhibits GUS staining in the middle of the root. Scale bars = 1 mm in (A, B), 200 μm in (C, D), 10 mm in (E, F) and 2 mm in (G, H).

analog of the growth hormone derivative indole-3-acetamide, the growth of seedlings possessing the *tms*2 gene was severely retarded. In the transposed lines, the appearance of an empty donor site (EDS) was judged by resistance to kanamycin, and the existence of a transposed *Ds* element was judged by resistance to hygromycin. NAM was used to segregate the *Ac* T-DNA line, because the *Ds* element might transpose in another region in the presence of the transposase.

We selected 50 plants that were resistant to kanamycin, hygromycin, and NAM. Integration sites of the *Ds* elements in 37 transposed lines were mapped on the genome by sequencing the flanking DNA. Figure 4A shows a map of the positions of the transposed *Ds*

Table 2. T-DNA integration sites in the gene trap lines.

Line No.	Accession No.	BAC clone	Position	Chr	Insertion site	Predicted protein	AGI code
KG1870	AC005990	F5O8	13308	1			
KG7313	AC009398	F20B24	35180	1	3′ UTR	expressed protein	At1g10660
KG7378	AL391143	T20K14	56972	5	Intron	expressed protein	At5g15540
KG7380	AC008153	F24K9	98319	3			
KG7387	AL138655	F24I3	80772	3	Exon	patatin-related	At3g57140
KG7562	AC021665	F17M19	32749	1	Exon	sucrose transporter	At1g71890
KG7655	AC006194	T17A11	78809	2	Exon	bZIP family transcription factor	At2g13150
KG7677	AL049876	T22B4	25519	4			
KG7689	AB010076	MUB3	80094	5			
KG7692	AC009243	F28K19	50675	1	Intron	MADS box transcription factor, putative	At1g77950
KG7711	AL161578	fragment No. 74	100814	4	Intron	expressed protein	At4g31150
KG7717	AB005231	MBB18	56316	5	Intron	cytochrome b-561	At5g38630
KG7719	AC006232	F10A12	48131	2	Intron	expressed protein	At2g27460
KG7732	AC003672	F16B22	95665	2			
KG7743	AC004261	T3K9	13594	2	Intron	expressed protein	At2g41200
KG7778	AC021666	T22A15	33026	1	Exon	gag-pol polyprotein -related	At1g36035
KG7786	AL078637	T22A6	39410	4			
KG7808	AL049483	F20B18	52330	4			
KG7809	AC000348	T7N9	36381	1	5′ upstream	hypothetical protein	At1g27060
KG7926	AC007019	F7D8	75221	2	Exon	hypothetical protein	At2g21920
KG7927	AB019229	MDC16	10913	3			
KG13054	AL162972	T1E3	45828	5	Intron	expressed protein	At5g04680
KG13075	AB023042	MUL3	18581	5	Intron	calcium-transporting ATPase	At5g57110
KG13078	AF195115	F5I10	13970	4			
KG13144	AL391148	T21H19	23900	5	5′ upstream	hypothetical protein	At5g16180
KG13185	AL161746	T10O8	35599	5	3′ UTR	expressed protein	At5g01360
KG13211	AC021199	F7F23	618	1			
KG13292	AC007017	F11F19	24461	2			
KG13317	AC005623	T20P8	45234	2			
KG13318	AL392174	F8L15	27289	5	3′ UTR	hydroxymethylbilane synthase	At5g08280
KG13324	AC006954	F25P17	53737	2			
KG13336	AL096859	T6H20	37932	3	Exon	subtilisin-like serine protease	At3g46850
KG13339	AC010704	T5M16	48714	1			
KG13384	AB024033	MBK21	48799	3	Exon	expressed protein	At3g12800
KG13407	AL049862	F18B3	85132	3	Exon	disease resistance protein , putative	At3g50950
KG13445	AL161946	F7A7	31849	5	5′ upstream	expressed protein	At5g01590
KG13463	AC010871	T16O11	77271	3	5′ upstream	pentatricopeptide repeat-containing protein	At3g08820
KG13474	AB026632	F15L12	20105	5			
KG13539	AC007887	F15O4	97045	1	3′ UTR	invertase -related	At1g35580
KG13541	AL096882	F8L21	75699	4			
KG13551	AL161541	fragment No. 41	115717	4			
KG13619	AL035656	T9A14	15026	4			
KG13692	AB026637	K16E14	8385	5	5′ upstream	cyclin-dependent protein kinase-related	At5g50860
KG13694	AC005724	MSF3	73478	2	Exon	reverse transcriptase family protein	At2g18820
KG13741	AC006551	F12K8	74367	1			
KG13755	AC002311	T26J12	2014	1	Exon	conserved hypothetical protein	At1g23230
KG13759	AC005169	F6F22	66553	2			
KG13839	AC069273	F28P5	26847	1			
KG13858	AC013258	F9E10	96525	1			
KG13860	AB008266	MHJ24	10173	5			
KG13917	AC005278	F15K9	65945	1	Exon	expressed protein	At1g03170
KG13920	AC003952	T13L16	26822	2			
KG14041	AC007627	F15F15	77072	5			
KG14050	AC006919	F1O11	18081	2	Intron	1,4-alpha-glucan branching enzyme	At2g36390
KG14052	AC011713	F23A5	60305	1			
KG14084	AB022215	MCB17	72199	3	Intron	imidazoleglycerolphosphate dehydratase	At3g22425
KG14120	AC005499	T6A23	32901	2			
KG14137	AL078470	F19B15	4135	4	Exon	expressed protein	At4g28990
KG14166	AB024031	K6A12	27772	5			
KG14207	AC025290	F9P14	29697	1	3′ UTR	elongation factor Tu family protein	At1g06220
KG14213	AC007654	T19E23	44722	1			
KG14235	AL035526	F28A21	13410	4	Exon	hypothetical protein	At4g18620

Table 2. Continued.

Line No.	Accession No.	BAC clone	Position	Chr	Insertion site	Predicted protein	AGI code
KG14332	AC068667	F15D2	44037	1	Exon	auxin-induced protein family	At1g29420
KG14345	AL049640	T20K18	61837	4	Exon	hydrolase, alpha/beta fold family	At4g12830
KG14358	AC006550	F10O3	46225	1	Exon	putataive transport protein	At1g03060
KG14387	AL162351	T20L15	63090	5			
KG14483	AL031326	F16G20	90680	4			
KG14513	AP001312	MYF5	11588	3	Intron	expressed protein	At3g27160
KG14547	AC006922	T1J8	93031	2	Exon	pentatricopeptide repeat-containing protein	At2g36980
KG14554	AL162295	T4C21	42886	3			
KG14577	AC007171	T8O18	68411	2	Exon	protein kinase family	At2g28590
KG14582	AC069551	T10F20	52394	1	Exon	mitogen-activated protein kinase, putative	At1g18150
KG14584	AL078465	T15N24	38667	4	Exon	calcineurin B-like protein 3	At4g26570
KG14615	AC009322	F18B13	88231	1	Intron	glyoxalase family protein	At1g80160
KG14663	AC004809	F13M7	82255	1			
KG14669	AC005106	T25N20	22004	1			
KG14670	AC008113	F12A21	63949	1			
KG14678	AL109787	T10C21	2751	4	5′ upstream	expressed protein	At4g30670
KG14708	AC003105	F18A8	43603	2			
KG14813	AL035525	F4I10	82817	4	Exon	myosin - like protein	At4g33200
KG14816	AL161539	fragment No. 39	121662	4	5′ upstream	expressed protein	At4g14690
KG14817	AC023912	F3E22	30322	3	Exon	polygalacturonase, putative	At3g06770
KG14839	AC069473	T21B14	23167	3	Exon	actin 11	At3g12110
KG14844	AB015479	MTE17	54054	5			
KG14864	AC004218	F12L6	85620	2	Exon	expressed protein	At2g39580
KG15016	AB007645	K8K14	57914	5			
KG15119	AL078467	F27G19	40692	4			
KG15184	AC022472	T20H2	45925	1			
KG15218	AC074360	F27M3	15149	1			
KG15237	AB009050	MDF20	30067	5			
KG15249	AC005936	F5O4	22161	2			
KG15278	AC079679	T6B12	16290	1			
KG15352	AC009325	F4P13	17828	3	Intron	carbonic anhydrase, chloroplast precursor	At3g01500
KG15382	AB006699	MDJ22	8153	5			
KG15390	AB016886	MCA23	27926	5			
KG15398	AC003033	T21L14	66995	2	Intron	AtZW10 protein	At2g32900
KG15419	AL138652	T18B22	52241	3			
KG15425	AC005700	T32F6	44922	2	Intron	homeodomain protein	At2g32370
KG15465	AC004667	T4C15	73655	2	3′ UTR	expressed protein	At2g35110
KG15515	AB010699	MSN9	20209	5			
KG15542	AC004561	F16P2	114126	2	Exon	pumilio-family RNA-binding protein, putative	At2g29190
KG15563	AC006072	T9J23	8274	2	Exon	glutaredoxin protein family	At2g47880
KG15583	AB011480	MPI7	39035	5	Intron	5-methyltetrahydropteroyltriglutamate—homocysteine S-methyltransferase	At5g17920
KG15627	AB024034	MDC11	52064	5			
KG15662	AC005896	F3G5	107500	2	Exon	expressed protein	At2g37480
KG15684	AC011765	F1M20	69176	1			
KG15693	AC004697	T16B24	23914	2	3′ UTR	jacalin lectin family	At2g39310
KG15706	AL022198	F6I18	40097	4			
KG16079	AL138642	F21F14	113116	3			
KG16158	AL133363	T20E23	77177	3			
KG16176	AC020665	T27F4	30405	1	5′ upstream	F-box protein family	At1g66320
KG16334	AC034256	F7H2	56278	1			
KG16341	AC016662	F2P9	100227	1	Exon	kinesin-related protein	At1g73860
KG16397	AB010073	MFB13	59001	5			
KG16480	AC007508	F1K23	32552	1			
KG16495	AC010793	F20B17	38200	1			
KG16533	AP002058	T1O13	49276	3			
KG16574	AC006593	F16D14	65121	2	Exon	transducin / WD-40 repeat protein family	At2g31300
KG16582	AB005232	MBG8	24884	5	Exon	tryptophan synthase, beta subunit 1	At5g54810
KG16592	AL161545	fragment No. 45	147754	4			
KG16615	AB025632	MQJ2	25677	5	Intron	RNA/ssDNA-binding protein - like	At5g58470
KG16630	AB028608	K20I9	34771	3			
KG16644	AC006259	F21J6	26822	5			

Table 2. Continued.

Line No.	Accession No.	BAC clone	Position	Chr	Insertion site	Predicted protein	AGI code
KG16657	AL391712	T5E8	102250	5	5′ upstream	proline-rich protein family	At5g09530
KG16663	AL035601	F6G17	51995	4			
KG16665	AC007067	T10O24	67878	1	3′ UTR	expressed protein	At1g10560
KG16732	AB011479	MNA5	84819	5	Intron	expressed protein	At5g65470
KG16946	AB025622	MFH8	29174	5			
KG16997	AL161543	fragment No. 43	69709	4	3′ UTR	expressed protein	At4g16143
KG17098	AL021749	F20O9	25553	4			
KG17110	AC000375	F19K23	14530	1	Exon	hypothetical protein	At1g62090
KG17114	AF160182	F17I23	75086	4			
KG17195	AC007190	F23N19	23519	1	3′ UTR	expressed protein	At1g62690
KG17214	AB010693	K21C13	73490	5	3′ UTR	disease resistance protein-related	At5g45050
KG17219	AC002535	T30B22	76100	2			
KG17252	AC079829	F28B23	58850	1	Intron	expressed protein	At1g26190
KG17259	AB024027	K18B18	7245	5			
KG17299	AB016871	K16L22	16436	5			
KG17301	AC007153	F3F20	95537	1			
KG17500	AC006223	F22D22	47365	2	5′ upstream	CCR4-associated factor -related	At2g32070

5′ upstream designates insertions within 300 bp upstream of the first ATG codon.

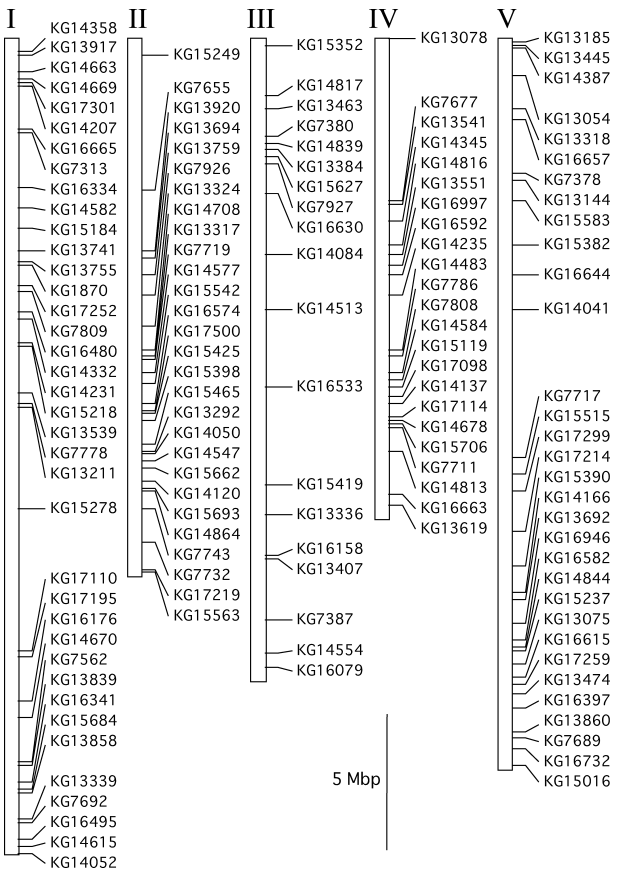

Figure 3. A map of the position of the inserted T-DNAs on the genome. Integration sites of the T-DNAs in gene trap lines were mapped by sequencing the flanking regions. The vertical open bars represent five chromosomes and each chromosome is identified by its number. The mapped positions of the inserted T-DNAs were arranged on chromosomes according to their integration sites. The numbers of the lines are given to the right of the chromosomes.

elements on the genome. The donor locus for *Ds* transposition in the line KG1870 is on chromosome I at approximately the 8.4 M bp position. The *Ds* elements in 21 out of 37 lines were transposed onto chromosome I, 2 were on chromosome II, 5 were on chromosome III, 6 were on chromosome IV, and 3 were on chromosome V. The integration sites of the *Ds* elements in 16 lines were within about 1M bp from the donor locus on the chromosome I (Figure 4B). The integration sites of the *Ds* elements in 4 lines were on the BAC clone, F5O8 where the donor locus was mapped (Figure 4B). The *Ds* elements transposed in both directions from the donor locus in roughly equal proportion, and the *Ds* elements in about half of the transposed lines preferentially moved to closely linked sites (within about 1 M bp). These results indicate that the 140 donor loci mapped for *Ds* transposition are useful resources for targeted tagging of *Arabidopsis* genes throughout the genome.

Interestingly, we isolated a late flowering mutant in the transposed lines (T239). In this line the *GIGANTEA* gene was disrupted by insertion of the *Ds* element; the *GIGANTEA* gene was mapped on BAC clone T22J18 which is near F5O8. In the line T443, the *Ds* element was inserted in the coding region of the At1g23720 gene. Disruption of the coding region of the At1g23720 gene is not found in the database T-DNA Express (http://www.signal.salk.edu/cgi-bin/tdnaexpress; Alonso et al. 2003). In the lines T498 and T548, the At1g23500 gene and the At1g23540 gene were disrupted by the *Ds* element. The T-DNA Express database includes one insertion line in which the coding region of each gene was disrupted. Our transposition system is therefore useful for isolating first mutant alleles as well as for obtaining additional mutant alleles to confirm existing phenotypes.

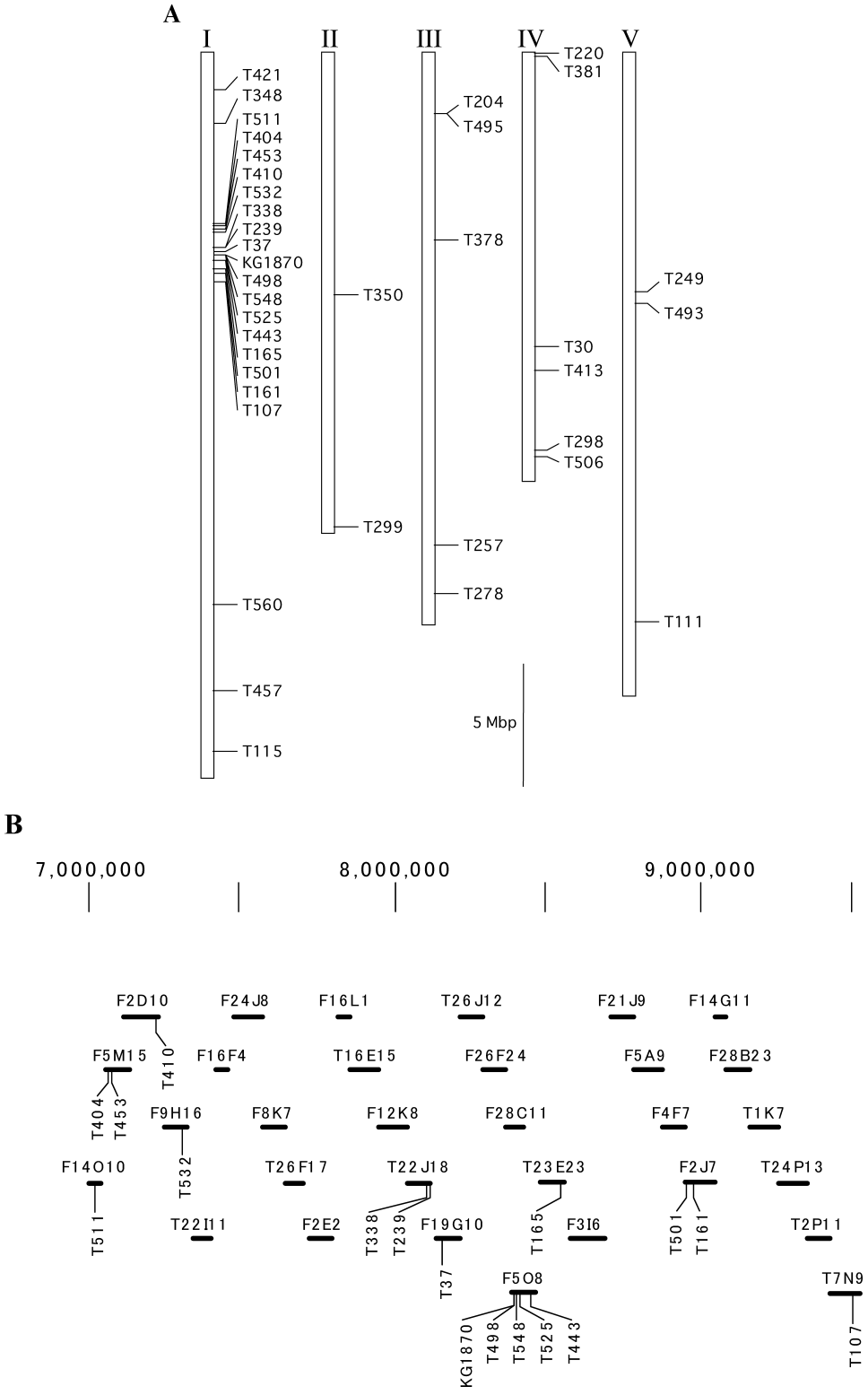

Figure 4. (A) A map of the positions of transposed *Ds* elements on the genome. Integration sites of the transposed *Ds* elements were mapped by sequencing the flanking regions. The vertical open bars represent five chromosomes and each chromosome is identified by its number. The mapped positions of the transposed *Ds* elements were arranged on chromosomes according to their integration sites. The numbers of the transposed lines are given to the right of the chromosomes. KG1870, the transposition donor locus, is shown on chromosome I. (B) Positions of the *Ds* elements transposed near the donor locus on chromosome I. The solid bars represent BAC clones, the numbers of which are shown on the upper side of the clones. The numbers representing the transposed lines and the donor locus for transposition, KG1870, are listed on the lower side of the BAC clones. The distance of the top arm from the telomeric site is indicated on the horizontal scale.

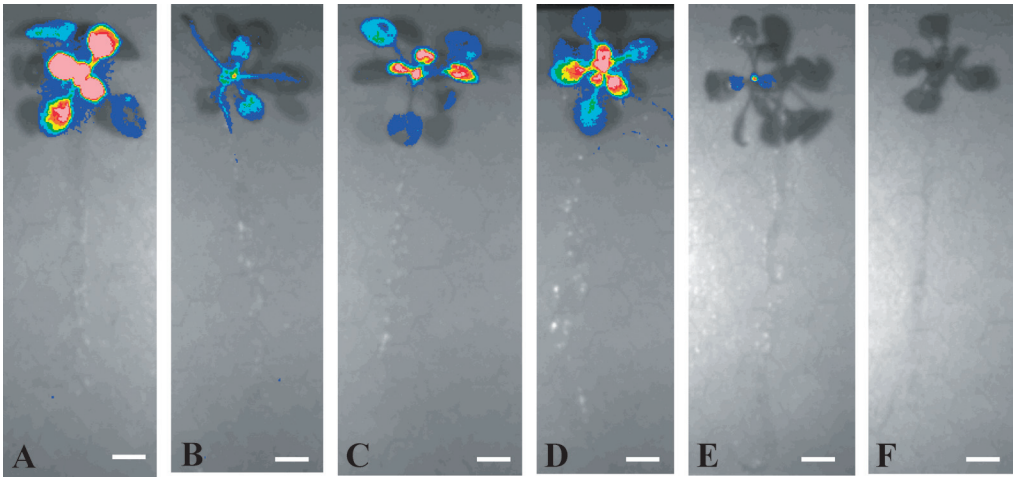

Figure 5. Images of light emission in the transposed lines. Light emission is represented in false colors overlayed onto photographs of the plants. (A)–(D) Light emission of 3-week-old plants in the transposed lines. Luciferase genes were expressed in leaves of the plants. (E) Light emission of a 3-week-old plant from the transposition donor line, KG1870. (F) Light emission of a 3-week-old Columbia plant. Scale bars = 10 mm.

Detection of luciferase activity in the transposed lines

The *Luc* gene in the vector can be used for gene trap analysis when the *Ds* element is transposed in other regions on the genome. We therefore analyzed luciferase activity in the transposed lines. Three plants from the transposed lines were sprayed with luciferin and the luminescence from each plant was analyzed in a photon-counting video system. Four out of 45 lines exhibited light emission in their leaves (Figure 5) whereas KG1870, the transposition donor plants, showed only a little light emission in their leaves, and wild type Columbia plants showed no light emission (Figure 5). These results indicate that gene trap mutagenesis using the transposed lines is also useful for expression analysis of the genes in targeted regions.

Acknowledgements

We are grateful to the Arabidopsis Biological Resource Center and Nina Fedoroff for providing seeds of the *Ac* and *Ds* T-DNA lines. We thank N. Nakazki, T. Wada, Y. Kishida, T. Akutsu, Y. Hasegawa, S. Kondo, C. Shitara, K. Sugawara, and H. Tsuruoka for excellent technical assistance. This work was supported by the Kazusa DNA Research Institute Foundation.

References

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657

Alvarado MC, Zsigmond LM, Kovacs I, Cseplo A, Koncz C, Szabados LM (2004) Gene trapping with firefly luciferase in Arabidopsis. Tagging of stress-responsive genes. *Plant Physiol* 134: 18–27

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815

Bancroft I, Dean C (1993) Transposition pattern of the maize element *Ds* in *Arabidopsis thaliana*. *Genetics* 134: 1221–1229

Bechtold N, Pelletier G (1998) *In planta Agrobacterium*-mediated transfomation of adult *Arabidopsis thaliana* plants by vacuum infiltration. *Meth Mol Biol* 82: 259–266

Campisi L, Yang Y, Yi Y, Heilig E, Herman B, Cassista AJ, Allen DW, Xiang H, Jack T (1999) Generation of enhancer trap lines in *Arabidopsis* and characterization of expression patterns in the inflorescence. *Plant J* 17: 699–707

Carroll BJ, Klimyuk VI, Thomas CM, Bishop GJ, Harrison K, Scofield SR, Jones JDG (1995) Germinal transpositions of the maize element *Dissociation* from T-DNA loci in tomato. *Genetics* 139: 407–420

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16: 735–743

Creff A, Leonard B, Desnos T (2006) Targeted *Ds*-tagging strategy generates high allelic diversity at the Arabidopsis *HY2* locus. *Plant Mol Biol* 61: 603–613

Dooner HK, Belachew A (1989) Transposition pattern of the maize element *Ac* from the *bz-m2*(*Ac*) allele. *Genetics* 122: 447–457

Dooner HK, Keller J, Harper E, Ralston E (1991) Variable patterns of transposition of the maize element *Activator* in tobacco. *Plant Cell* 3: 473–482

Fedoroff NV, Smith DL (1993) A versatile system for detecting transposition in *Arabidopsis*. *Plant J* 3: 273–289

Feldmann KA (1991) T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *Plant J* 1: 71–82

Galbiati M, Moreno MA, Nadzan G, Zourelidou M, Dellaporta SL (2000) Large-scale T-DNA mutagenesis in *Arabidopsis* for functional genomic analysis. *Funct Integr Genomics* 1: 25–34

Ito T, Motohashi R, Kuromori T, Mizukado S, Sakurai T, Kanahara H, Seki M, Shinozaki K (2002) A new resource of locally

transposed *Dissociation* elements for screening gene-knockout lines in silico on the Arabidopsis genome. *Plant Physiol* 129: 1695–1699

Jones JDG, Carland F, Lim E, Ralston E, Dooner HK (1990) Preferential transposition of the maize element *Activator* to linked chromosomal locations in tobacco. *Plant Cell* 2: 701–707

Kato T, Shirano Y, Kawazu T, Tada Y, Itoh E, Shibata D (1991) A modified β-glucuronidase gene: Sensitive detection of plant promoter activities in suspension-cultured cells of tobacco and rice. *Plant Mol Biol Rep* 9: 333–339

Klimyuk VI, Nussaume L, Harrison K, Jones JDG (1995) Novel GUS expression patterns following transposition of an enhancer trap *Ds* element in *Arabidopsis*. *Mol Gen Genet* 249: 357–365

Krysan PJ, Young JC, Tax F, Sussman MR (1996) Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport. *Proc Natl Acad Sci USA* 93: 8145–8150

Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in Arabidopsis. *Plant Cell* 11: 2283–2290

Kuromori T, Hirayama T, Kiyosue Y, Takabe H, Mizukado S, Sakurai T, Akiyama K, Kamiya A, Ito T, Shinozaki K (2004) A collection of 11800 single-copy *Ds* transposon insertion lines in *Arabidopsis*. *Plant J* 37: 897–905

Liu P-P, Koizuka N, Homrichhausen TM, Hewitt JR, Martin RC, Nonogaki H (2005) Large-scale screening of *Arabidopsis* enhancer-trap lines for seed germination-associated genes. *Plant J* 41: 936–944

Liu YG, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* 8: 457–463

Liu YG, Shirano Y, Fukaki H, Yanai Y, Tasaka M, Tabata S, Shibata D (1999) Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. *Proc Natl Acad Sci USA* 96: 6535–6540

Machida C, Onouchi H, Koizumi J, Hamada S, Semiarti E, Torikai S, Machida Y (1997) Characterization of the transposition pattern of the *Ac* element in *Arabidopsis thaliana* using endonuclease I-*Sce*I. *Proc Natl Acad Sci USA* 94: 8675–8680

Marsch-Martinez N, Greco R, Van Arkel G, Herrera-Estrella L, Pereira A (2002) Activation tagging using the *En-I* maize transposon system in Arabidopsis. *Plant Physiol* 129: 1544–1556

McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA, Tossberg J, Nickle T, Levin JZ, Law M, Meinke D, Patton D (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* 159: 1751–1763

McKinney EC, Ali N, Traut A, Feldmann KA, Belostotsky DA, McDowell JM, Meagher RB (1995) Sequence-based identification of T-DNA insertion mutations in *Arabidopsis*: Actin mutants *act2-1* and *act4-1*. *Plant J* 8: 613–622

Meissner RC, Jin H, Cominelli E, Denekamp M, Fuertes A, Greco R, Kranz HD, Penfield S, Petroni K, Urzainqui A, Martin C, Paz-Ares J, Smeekens S, Tonelli C, Weisshaar B, Baumann E, Klimyuk V, Marillonnet S, Patel K, Speulman E, Tissier AF, Bouchez D, Jones JJD, Pereira A, Wisman E, Bevan M (1999) Function search in a large transcription factor gene family in Arabidopsis: assessing the potential of reverse genetics to identify insertional mutations in R2R3 *MYB* genes. *Plant Cell* 11: 1827–1840

Muskett PR, Clissold L, Marocco A, Springer PS, Martienssen R, Dean C (2003) A resource of mapped *Dissociation* launch pads for targeted insertional mutagenesis in the Arabidopsis genome. *Plant Physiol* 132: 506–516

Nagawa S, Sawa S, Sato S, Kato T, Tabata S, Fukuda H (2006) Gene trapping in *Arabidopsis* reveals genes involved in vascular development. *Plant Cell Physiol* 47: 1394–1405

Nakayama N, Arroyo JM, Simorowski J, May B, Martienssen R, Irish VF (2005) Gene trap lines define domains of gene regulation in *Arabidopsis* petals and stamens. *Plant Cell* 17: 2486–2506

Nakazawa M, Ichikawa T, Ishikawa A, Kobayashi H, Tsuhara Y, Kawashima M, Suzuki K, Muto S, Matsui M (2003) Activation tagging, a novel tool to dissect the functions of a gene family. *Plant J* 34: 741–750

Pan X, Liu H, Clarke J, Jones J, Bevan M, Stein L (2003) ATIDB: *Arabidopsis thaliana* insertion database. *Nucl Acids Res* 31: 1245–1251

Parinov S, Sevugan M, Ye D, Yang WC, Kumaran M, Sundaresan V (1999) Analysis of flanking sequences from *Dissociation* insertion lines: a database for reverse genetics in Arabidopsis. *Plant Cell* 11: 2263–2270

Plesch G, Kamann E, Mueller-Roeber B (2000) Cloning of regulatory sequences mediating guard-cell-specific gene expression. *Gene* 16: 83–89

Raina S, Mahalingam R, Chen F, Fedoroff N (2002) A collection of sequenced and mapped *Ds* transposon insertion sites in *Arabidopsis thaliana*. *Plant Mol Biol* 50: 93–110

Rios G, Lossow A, Hertel B, Breuer F, Schaefer S, Broich M, Kleinow T, Jásik J, Winter J, Ferrando A, Farrás R, Panicot M, Henriques R, Mariaux J-B, Oberschall A, Molnár G, Berendzen K, Shukla V, Lafos M, Koncz Z, Rédei GP, Schell J, Koncz C. (2002) Rapid identification of *Arabidopsis* insertion mutants by non-radioactive detection of T-DNA tagged genes. *Plant J* 32: 243–253

Samson F, Brunaud V, Balzergue S, Dubreucq B, Lepiniec L, Pelletier G, Caboche M, Lecharny A (2002) FLAGdb/FST: A database of mapped flanking insertion sites (FSTs) of *Arabidopsis thaliana* T-DNA transformants. *Nucl Acids Res* 30: 94–97

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke JD, Cotton D, Bullis D, Snell J, Miguel T, Hutchison D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff SA (2002) A high-throughput Arabidopsis reverse genetics system. *Plant Cell* 14: 2985–2994

Speulman E, Metz PLJ, van Arkel G, te Lintel Hekkert B, Stiekema WJ, Pereira A (1999) A two-component *Enhancer-Inhibitor* transposon mutagenesis system for functional analysis of the Arabidopsis genome. *Plant Cell* 11: 1853–1866

Springer PS (2000) Gene traps: tools for plant development and genomics. *Plant Cell* 12: 1007–1020

Stangeland B, Nestestog R, Grini PE, Skrbo N, Berg A, Salehian Z, Mandal A, Aalen RB (2005) Molecular analysis of *Arabidopsis* endosperm and embryo promoter trap lines: reporter-gene expression can result from T-DNA insertions in antisense orientation, in introns and in intergenic regions, in addition to sense insertion at the 5′ end of genes. *J Exp Bot* 56: 2495–2505

Sundaresan V, Springer P, Volpe T, Haward S, Jones JDG, Dean C, Ma H, Martienssen R (1995) Patterns of gene action in plant

development revealed by enhancer trap and gene trap transposable elements. *Genes Dev* 9: 1797–1810

Tantikanjana T, Mikkelsen MD, Hussain M, Halkier BA, Sundaresan V (2004) Functional analysis of the tandem-duplicated P450 genes *SPS/BUS/CYP79F1* and *CYP79F2* in glucosinolate biosynthesis and plant development by *Ds* transposition-generated double mutants. *Plant Physiol* 135: 840–848

Tissier AF, Marillonnet S, Klimyuk V, Patel K, Torres MA, Murphy G, Jones JDG (1999) Multiple independent defective *Suppressor-mutator* transposon insertions in Arabidopsis: a tool for functional genomics. *Plant Cell* 11: 1841–1852

Topping JF, Agyeman F, Henricot B, Lindsey K (1994) Identification of molecular markers of embryogenesis in *Arabidopsis thaliana* by promoter trapping. *Plant J* 5: 895–903

Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrandiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia Y, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J (2000) Activation tagging in Arabidopsis. *Plant Physiol* 122: 1003–1013

Winkler RG, Frank MR, Galbraith DW, Feyereisen R, Feldmann KA (1998) Systematic reverse genetics of transfer-DNA-tagged lines of Arabidopsis. Isolation of mutations in the cytochrome P450 gene superfamily. *Plant Physiol* 118: 743–750

Wisman E, Hartmann U, Sagasser M, Baumann E, Palme K, Hahlbrock K, Saedler H, Weisshaar B (1998) Knock-out mutants from an *En-1* mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. *Proc Natl Acad Sci USA* 95: 12432–12437

Woody ST, Austin-Phillips S, Amasino RM, Krysan PJ (2007) The *WiscDsLox* T-DNA collection: an arabidopsis community resource generated by using an improved high-throughput T-DNA sequencing pipeline. *J Plant Res* 120: 157–165

Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, Pham P, Cheuk R, Karlin-Newmann G, Liu SX, Lam B, Sakano H, Wu T, Yu G, Miranda M, Quach HL, Tripp M, Chang CH, Lee JM, Toriumi M, Chan MM, Tang CC, Onodera CS, Deng JM, Akiyama K, Ansari Y, Arakawa T, Banh J, Banno F, Bowser L, Brooks S, Carninci P, Chao Q, Choy N, Enju A, Goldsmith AD, Gurjal M, Hansen NF, Hayashizaki Y, Johnson-Hopson C, Hsuan VW, Iida K, Karnes M, Khan S, Koesema E, Ishida J, Jiang PX, Jones T, Kawai J, Kamiya A, Meyers C, Nakajima M, Narusaka M, Seki M, Sakurai T, Satou M, Tamse R, Vaysberg M, Wallender EK, Wong C, Yamamura Y, Yuan S, Shinozaki K, Davis RW, Theologis A, Ecker JR (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* 302: 842–846

Yamamoto YY, Tsuhara Y, Gohda K, Suzuki K, Matsui M (2003) Gene trapping of the *Arabidopsis* genome with a firefly luciferase reporter. *Plant J* 35: 273–283

Zhang S, Raina S, Li H, Li J, Dec E, Ma H, Huang H, Fedoroff NV (2003) Resources for targeted insertional and deletional mutagenesis in *Arabidopsis*. *Plant Mol Biol* 53: 133–150