CMLS, Cell. Mol. Life Sci. 61 (2004) 2632–2645
1420-682X/04/202632-14
DOI 10.1007/s00018-004-4284-8

CMLS Cellular and Molecular Life Sciences

Research Article

Single-copy T-DNAs integrated at different positions in the *Arabidopsis* genome display uniform and comparable β-glucuronidase accumulation levels

S. De Buck [a], P. Windels [b], M. De Loose [b] and A. Depicker [a,*]

[a] Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology (VIB), Ghent University, Technologiepark 927, 9052 Gent (Belgium), Fax: +32 9 3313809, e-mail: ann.depicker@psb.ugent.be
[b] Department of Plant Genetics and Breeding, Center of Agricultural Research-Gent, Caritasstraat 21, 9090 Melle (Belgium)

Received 29 June 2004; accepted 19 August 2004

Abstract. This study aimed at determining whether transgene expression variability is observed in single-copy T-DNA plants and whether it can be correlated with the T-DNA integration position. Among a population of 135 *Arabidopsis thaliana* transformants, selected on the basis of antibiotic resistance marker expression, 21 single-copy T-DNA transformants were identified and characterized. In 19 of these 21 lines, 35S-β-glucuronidase transgene expression, measured in two subsequent generations, was similar. This observation means that the intra-transformant variability was as high as the inter-transformant variability. Integration into an intergenic or genic region, into an exon or intron, in sense or antisense orientation, did not result in differential transgene expression. Remarkably, single-copy transformants were not always the highest expressers, implying that low transgene expression is not always induced by multicopy transformants. In only 2 of the 21 single-copy plants was the transgene expression more than 20-fold lower. However, characteristics of the insertion position in one of these lines did not differ significantly when compared to high-expressing lines. In the remaining line, methylation of the transgene was clearly demonstrated. In conclusion, screening for single-copy T-DNA transformants greatly enriches for stable and high transgene expression, because the integration position is not a major determinant of transgene expression variability in *Arabidopsis*.

Key words. Position effect; T-DNA integration; transgene expression.

Over the past 20 years, *Agrobacterium tumefaciens*-mediated transformation of mono- and dicotyledonous plants has become a useful tool both for basic research and for many commercial breeding programs [1–5]. For example, the use of transgenic plants for the production of high-value proteins is becoming increasingly important [6]. Transgene loci obtained after *A. tumefaciens* transformation are considered to be less complex than those obtained after direct gene transfer, but integration of multiple T-DNA copies into direct and inverted repeats is fairly common [7–12]. In plants, these repeated transgenes have been correlated with transgene expression variability and induction of transgene silencing [8, 13–17]. Although repetitiveness is an important feature in the induction of silencing, single-copy transgenes can be silenced, either by the influence of trans-acting mechanisms or by the neighboring heterochromatic DNA, the so-called 'position effect' [18–23].

Currently, strategies are being developed to diminish the influence of the chromatin environment on transgenes

* Corresponding author.

and/or to limit the number of transgene copies in plants. One approach, employed in rice and tobacco, is to integrate the transgene at a fixed chromosomal position through site-specific recombination [20, 24]. In rice, the transgene is expressed at a more or less consistent level [24], whereas in tobacco, nearly half the insertions show variability of transgene expression [20]. Furthermore, Srivastava et al. [25] demonstrated successful resolution of a complex integration pattern into a single-copy transgene using the Cre/lox recombination system. Another approach is to protect the transgene from the influence of the local chromatin environment by the addition of matrix attachment regions (MARs) [17, 26, 27]. However, the functional relevance of MARs for stabilizing transgene expression remains controversial [21, 28]. Finally, by using *Arabidopsis thaliana* post-transcriptional gene silencing mutants as a genetic background for transgene expression analysis, Butaye et al. [29] showed that these mutants confer a significantly higher transgene expression level, which is stably maintained in consecutive generations.

To date, the correlation between integration position and transgene expression has not been examined. This study investigated whether the transgene expression level in single-copy T-DNA transformants resulted from the single T-DNA integration itself or whether it was influenced by the integration position within the genome. The aim was to determine whether selecting solely for single-copy T-DNA transformants results in transgenes with high and stable expression. An additional aim was to establish whether the integration position or purely the presence of a single T-DNA copy is more important for stable expression. To discriminate between the two possibilities, the β-glucuronidase (GUS) accumulation levels were determined in a population of 100 transformants. Subsequently, the correlation between the T-DNA integration position and GUS accumulation levels was analyzed in 21 single-copy *Arabidopsis* T-DNA transformants.

Materials and methods

Constructs and strains

For the construction of the plasmids pK2L610 (K), pH610 (H), pHsb610 (Hsb), and pSingleGus (Ksb), the reader is referred to De Buck et al. [11, 30]. The K and Ksb T-DNAs had the kanamycin resistance gene under control of the nopaline synthase (*nos*) promoter, whereas this promoter drove the hygromycin resistance gene in the H and Hsb T-DNAs (fig. 1). All T-DNAs harbored the *gus* gene under control of the cauliflower mosaic virus promoter (P35S). The pK2L610 and pH610 plasmids contained borders in the natural octopine context with inner and outer border regions, and the pHsb610 and pSingleGus plasmids possessed the outer border regions only [30].

Generation of transgenic plants

Two main groups of transgenic plants were analyzed. From the group of CK2L transformants, obtained after root transformation of *A. thaliana* (L.) Heynh. (C24 ecotype) harboring the K T-DNA from pK2L610 (CK transformants [31]), only the transformants with a single K T-DNA [32] were retained. The other group of F transformants was obtained from five different floral dip (F) experiments of *Arabidopsis* (ecotype Columbia) with an *A. tumefaciens* strain that contained either the K, the H, the Hsb, or the Ksb T-DNA. The floral dip procedure was performed as described by Clough and Bent [33]. Seeds of the dipped plants were harvested and sown for selection on K1 medium supplemented with kanamycin (50 mg l^{-1}) or hygromycin (20 mg l^{-1}). Because 25 mg of *Arabidopsis* seeds correspond to approximately 1 000 seeds, the number of seeds and the transformation frequency were approximated. A segregation ratio analysis to obtain the number of T-DNA loci was performed as described elsewhere [10]. T1 plants are defined as plants grown from seeds harvested from primary transformants obtained after *Arabidopsis* floral dip or root transformations.

Plant DNA preparation and DNA gel blot analysis

DNA of *Arabidopsis* leaf material was prepared as described previously [10] and the *gus*, neomycin phosphotransferase (*npt*), and hygromycin phosphotransferase (*hpt*) probes according to De Buck et al. [11]. Per lane, 1 µg of *Arabidopsis* DNA was loaded. A non-radioactive system (GeneImages random prime-labeling module and GeneImages CDP star detection module; Amersham Biosciences, Little Chalfont, UK) was used to label and detect the DNA.

T-DNA fingerprinting analysis and characterization of the T-DNA/plant junctions

The T-DNA copy number in the different lines was identified by the T-DNA fingerprint technique [32]. From each transgenic line, a T-DNA fingerprint was generated with *Mse*I- and *Bfa*I-digested DNA as template. Subsequently, in the single-copy lines, the sequence of the T-DNA/plant junction was determined after amplification [34]. Homology between the amplified junction and the T-DNA plasmid was screened by pairwise sequence alignment [34], whereas homology between the T-DNA junction fragment and the *Arabidopsis* genome was determined with the BLAST algorithm [35] against the *Arabidopsis* database (http://www.arabidopsis.org/BLAST [36]). The T-DNA-surrounding target sequences were analyzed with the MARSCAN software to determine the presence of MAR recognition signatures (MRSs) [37]. These MRSs are composed of two degenerate sequence

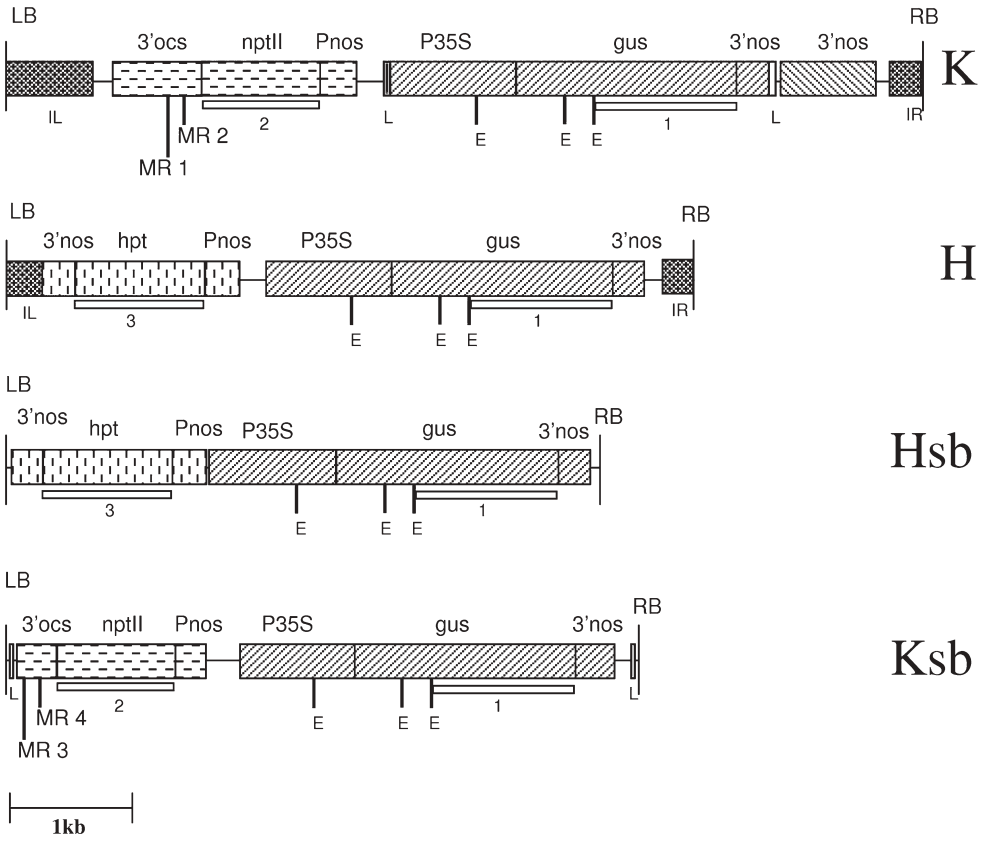

Figure 1. Schematic outline of the different T-DNAs. The different T-DNAs are presented with indication of the transgenes, i.e., kanamycin and hygromycin resistance genes, and the *gus* reporter gene. The 25-bp RB and LB repeats are represented as vertical black lines. The K and H T-DNAs contained the repeats in a natural octopine T-DNA context, whereas in the Hsb and the Ksb T-DNAs, the inner flanking regions of these borders were replaced by a random sequence. The K T-DNA was present in the CK2L, FK, and F2K transformants, the H T-DNA in the FH transformants, and the Ksb and Hsb T-DNAs in the F2Ksb and F2Hsb transformants. The MAR recognition sequences (MR) are: MR1, 5′cgttactatcgtattt3′; MR2, 5′aataacaa3′; MR3, 5′gttactatcgtatttt3′; and MR4, 5′aataacaa3′. The extents of the *gus*, *nptII,* and *hpt* probes are indicated by the white bars 1, 2, and 3, respectively, below the T-DNAs. Abbreviations: 3′*nos*, 3′ end of the nopaline synthase gene; 3′*ocs*, 3′ end of the octopine synthase gene; E, *Eco*RV; *gus*, *β*-glucuronidase gene; *hpt*, hygromycin phosphotransferase gene; IL, inner left border region; IR, inner right border region; L, recognition site of *Cre* recombinase; LB, left border; MR, MAR recognition sequence; nptII, neomycin phosphotransferase gene; P35S, cauliflower mosaic virus 35S promoter; Pnos, nopaline synthase gene promoter; RB, right border.

motifs situated in close proximity to each other. The MARSCAN software specifically searches for the presence of these two sequence motifs. The integration sites of the Ksb T-DNAs into F2Ksb5 and F2Ksb18 were not positioned, because the border regions could not be amplified.

Determination of GUS activity by fluorimetric measurement and histochemical staining

Three or 6 weeks after sowing on selective medium, three rosette leaves of each T1 or T2 plant were harvested and ground in 100 µl buffer containing 50 mM phosphate buffer (pH 7), 10 mM *β*-mercaptoethanol, 10 mM Na_2-EDTA, and 0.1% Triton X-100. The mixture was centrifuged (15,000 g) twice at 4 °C for 10 min to remove insoluble material. GUS activity, expressed as units of GUS protein relative to the total amount of soluble extracted protein (U GUS protein mg^{-1}) was determined as described elsewhere [15, 38].

Two-week-old plants (five per transformant) were stained with GUS staining buffer (0.1 M sodium phosphate buffer, pH 7, 10 mM EDTA, 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 0.5 g/l 5-bromo-4-chloro-3-indolyl-*β*-D-glucuronide in dimethylsulfoxide, and 1% dimethylsulfoxide) for 18 h. Tissues were cleared in 90% ethanol.

Results

Determination of GUS activity in a population of transgenic plants obtained after floral-dip transformation

Four T-DNAs, containing the same chimeric 35S-*gus*-3′*nos* gene (fig. 1) were used to obtain five series of transgenic plants after floral-dip transformation (see Materials and methods). In experiment F, *A. tumefaciens* strains containing either the K or H T-DNA (fig. 1) were used, resulting in FK and FH transformants. In experi-

ment F2, strains containing the K, Hsb or Ksb T-DNAs were used, resulting in F2K, F2Hsb and F2Ksb transformants (fig. 1). In this manner, a total of 100 transformants were selected: 21 FK, 31 FH, 13 F2K, 17 F2Ksb, and 18 F2Hsb transgenic plants (table 1). The floral-dip transformation frequencies varied between 0.08 and 0.1% (table 1; see Materials and methods), which was low compared with the 0.5–3% obtained by Clough and Bent [33].

To evaluate the transgene expression in a population of transformants, GUS activity was measured in five T1 progeny plants of all 100 self-fertilized segregating primary transformants (see Materials and methods). Figure 2A shows the GUS activity in 3-week-old T1 plants for 100 different transformants. Figure 2B–E displays the activity levels in five T1 plants from transformants with the same T-DNA from the various floral-dip experiments. At the moment of testing, locus, copy number, and zygosity of the T-DNA insert(s) were not known. To classify transformants with low, 'normal', or high GUS activity, we chose the arbitrary limit of 140 U GUS mg protein^{-1} based on previous results obtained with the same 35S-*gus* expression cassette [15]; the amount of GUS mg protein^{-1} from identical *gus* transgenes in an inverted repeat (line KH15) was lower than 20 U and increased to a value higher than 140 U after one *gus* transgene had been deleted from the inverted repeat (derived deletion line KH15d6) [15]. With this criterion, the population of transformants could be divided into three major groups. In the first group, representing 23% of the transformants, transgene expression was high in all T1 plants. The variation observed between progenies of every transformant in this category may be caused by different zygosity of the insert or different T-DNA copy numbers per seedling. In the second group, representing 42% of the transformants, all five seedlings had low transgene expression. In the last group, consisting of the remaining 35%, high transgene expression was observed in some, but not all, T1 plants. Again, different copy number and T-DNA configuration per seedling or the zygosity of the insert are a likely explanation for this difference. The comparable expression profiles of the subpopulations of transformants with a different T-DNA (fig. 2B–E) containing the same chimeric 35S-*gus*-3′*nos* gene at a similar position and orientation implies that the various components of the different T-DNAs do not influence the transgene expression.

Identification of the single-copy T-DNA transformants in a population of transgenic plants

To analyze the extent of variation in transgene expression, single-copy T-DNA transformants in the different populations of transgenic plants were selected both by DNA gel blot analysis and by T-DNA fingerprinting (see Materials and methods). A segregation ratio analysis of the progenies of 100 transformants revealed the presence of the T-DNA(s) in one or multiple loci (data not shown). One or more T-DNAs integrated at one single locus were found in 12 FK, 18 FH, 9 F2K, 8 F2Ksb, and 11 F2Hsb transgenic plants (table 1). All these single-locus plants were screened by DNA gel blot analysis for single left and right T-DNA/plant junction fragments with the methylation-insensitive enzyme *Eco*RV (fig. 1). To visualize the right border (RB) and left border (LB) T-DNA/plant junctions, the *gus* and *nptII* (for K and Ksb) or *hpt* (for H and Hsb) probes were used (fig. 1). Multiple right and left T-DNA/plant DNA junction fragments showed that the majority of the transformants obtained after floral-dip contained multiple integrated T-DNAs (data not shown). The transformants with only one left and one right T-DNA/plant hybridizing junction were further characterized with two additional enzymes (*Sph*I and *Hin*dIII for FH and FK/F2K, and *Sph*I and *Nhe*I for F2Hsb and F2Ksb transformants) in combination with the same probes (data not shown). In this manner, one FK, two FH, two F2K, two F2Ksb and four F2Hsb single-copy T-DNA transformants could be identified (table 1). Transformant FH48 was not retained because no signal was obtained for the RB region, probably because of a truncation event. The frequency with which single-copy transformants were obtained after floral dip transformation varied between 4.8% and 22%, which is much lower than the 52% single-copy T-DNA transformants obtained after *Arabidopsis* root transformation [32].

Table 1. Number of transgenic plants and single-copy transformants after floral-dip transformation.

Floral dip	Seeds	Transformants (frequency)	Single-locus transformants (%)[a]	Single-copy transformants (%)	Name of single-copy transformants
FK	27,012	21 (0.078%)	12/19 (63%)	1/21 (4.8%)	FK24
FH	31,080	31 (0.099%)	18/27 (67%)	2/31 (6.4%)	FH33, FH48
F2K	16,172	13 (0.080%)	9/11 (82%)	2/13 (15%)	F2K3, F2K16
F2Ksb	21,096	17 (0.080%)	8/13 (61%)	2/17 (12%)	F2Ksb5[b], F2Ksb18[b]
F2Hsb	21,720	18 (0.083%)	11/16 (69%)	4/18 (22%)	F2Hsb20, F2Hsb21, F2Hsb22, F2Hsb31

[a] For some transformants, the locus number could not be determined.
[b] The integration position could not be localized due to technical problems fingerprinting the borders.

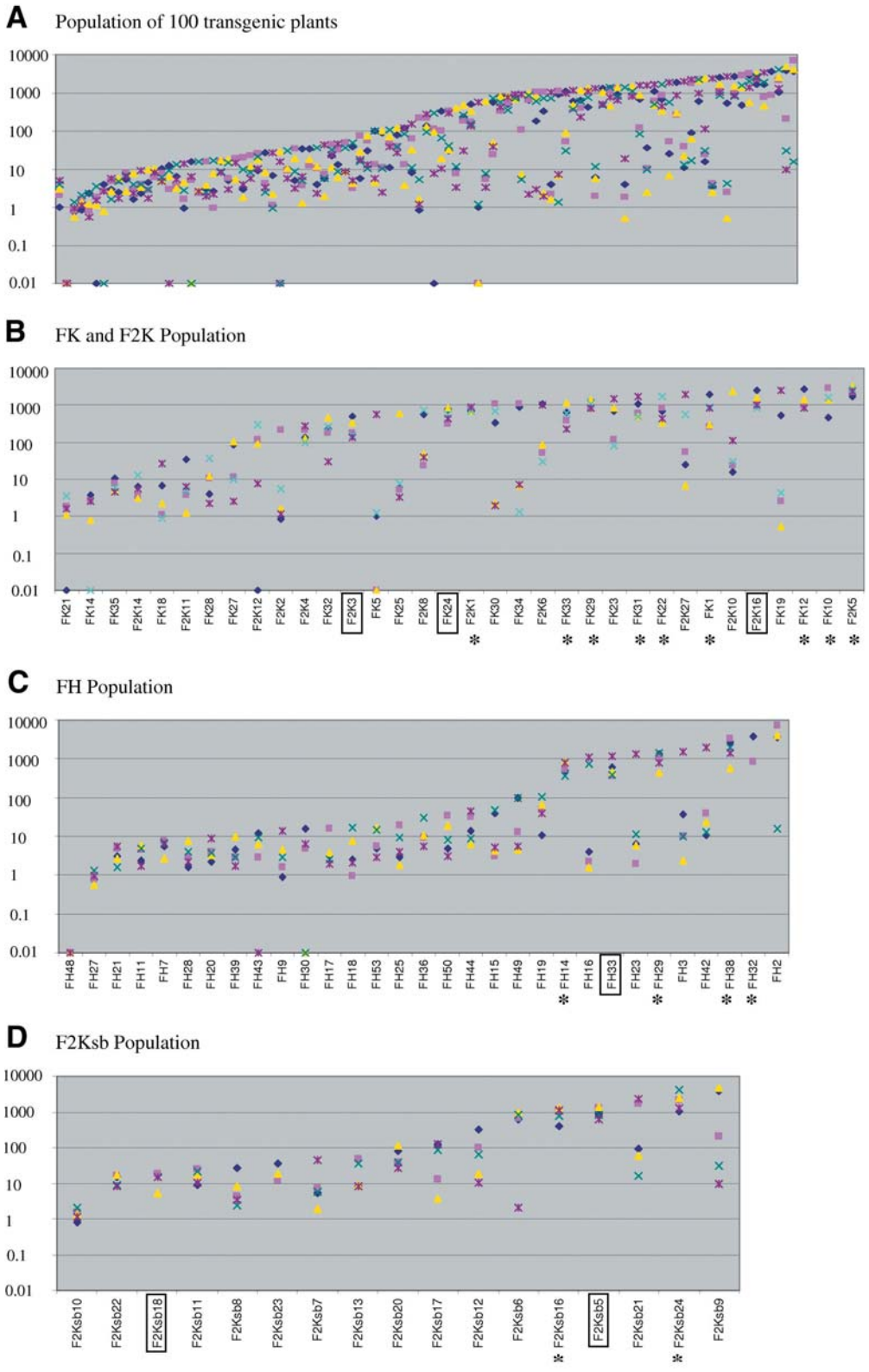

Figure 2. GUS activity analysis in populations of transgenic plants. The entire population of 100 transformants (*A*) consisted of 21 FK and 13 F2K transformants (*B*), 31 FH transformants (*C*), 17 F2Ksb transformants (*D*), 18 F2Hsb transformants (*E*), and 11 single-copy transformants (*F*). GUS activity was measured in leaves of five different 3-week-old T1 plants per transformant. The single-copy transformants are boxed and transformants marked with an asterisk had high transgene expression in all five T1 seedlings (indicated by different symbols). The transformants are ordered based on their highest expressers. The GUS activity levels are given as units GUS mg^{-1} of total soluble protein and plotted on a logarithmic scale.

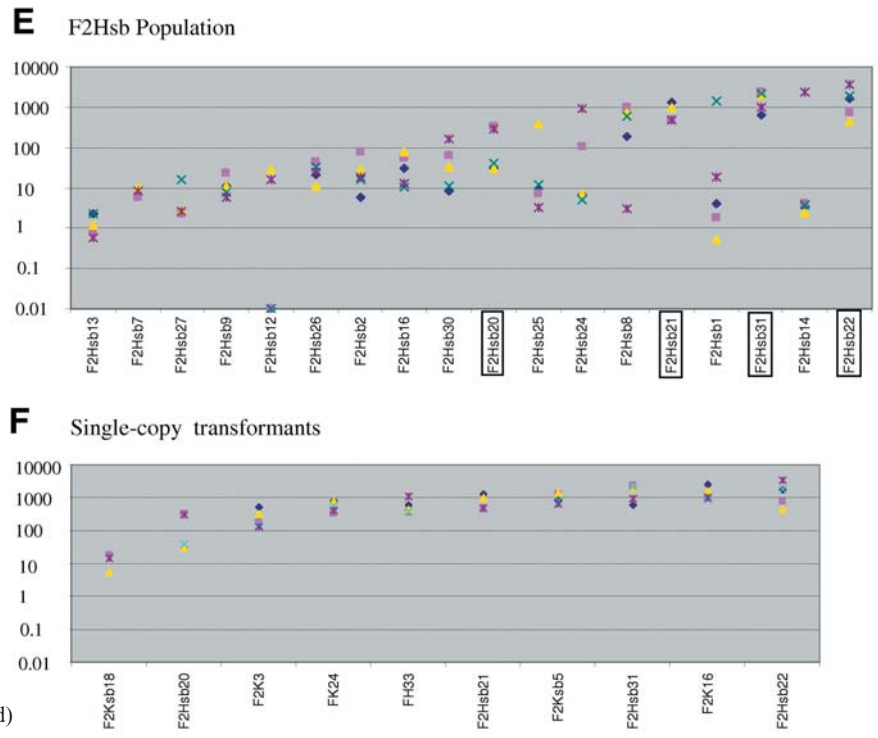

Figure 2. (continued)

Characterization of the T-DNA integration position in the single-copy transformants

To correlate the level of transgene expression in single-copy T-DNA plants with characteristics of the genomic sequences flanking the T-DNA, the integration positions of single-copy T-DNAs were analyzed in 8 identified single-copy transformants (as described above) and in 11 single-copy T-DNA CK2L lines, which contained the K T-DNA (fig. 1) [31, 32]. The sequences of LB and RB T-DNA/plant junctions were characterized (see Materials and methods [34]), and by alignment with the *Arabidopsis* genome sequence, the exact integration position of the T-DNAs on the different chromosomes could be established (table 2, fig. 3; see Materials and methods). In 10 of the 19 transformants, both LB and RB T-DNA/plant junctions could be positioned and the target deletion upon T-DNA integration could be calculated. Comparison with the annotation data [39] allowed us to determine whether the integration took place within a gene, in an exon, in an intron, or in an intergenic region. The AT percentage and the presence of tandemly repeated DNA sequences were assayed by analyzing 300 bp of the neighboring plant DNA of the T-DNA insert. Additionally, 3000 bp plant DNA bordering the T-DNA insert was also examined for the presence of MARs or MAR-like sequences (see Materials and methods).

Of the 19 single-copy transformants, seven T-DNAs were integrated into chromosome 1, three into chromosome 2, three into chromosome 3, two into chromosome 4, and two into chromosome 5 (table 2, fig. 3). For two transformants, as recently observed by Forsbach et al. [40], the LB and RB mapped to two different chromosomes. For transformants FH33 and CK2L129, the LB and RB regions were fused to sequences of chromosomes 4 and 1, and 1 and 2, respectively. This fusion may be caused by (i) a translocation of chromosome arms upon T-DNA integration, (ii) a translocation of a segment of DNA from its original site of insertion into a new location, or (iii) a replication-mediated copying of genetic information from another chromosome.

Nine of the single-copy T-DNAs were integrated into an intergenic region (FH33, F2K3, F2K16, F2Hsb31, CK2L6, CK2L7, CK2L102, CK2L107, and CK2L129) and ten within a transcribed annotated gene of which seven were in an exon (FK24, F2Hsb22, CK2L36, CK2L72, CK2L94, CK2L111, and CK2L148) and three were in an intron (F2Hsb20, F2Hsb21, and CK2L70) (table 2). In four (FK24, F2Hsb20, F2Hsb21, and CK2L 72) and five (F2Hsb22, CK2L36, CK2L70, CK2L94, and CK2L148) transformants, the *gus* gene was oriented in antisense and sense orientation relative to the transcription orientation of the endogenous gene, respectively. The AT content of the neighboring plant DNA varied between 40% and 75% and did not significantly differ from that of the T-DNA border regions (41–67%). The T-DNAs of six transformants (FH33, F2Hsb20, CK2L94, CK2L102, CK2L111, and CK2L129) were integrated close to MAR sequences. In addition, two sequences of 8 and 16 bp in the 3′*ocs* region of the kanamycin selection marker in both the K and Ksb T-DNAs (fig. 1) were identified when

Table 2. Overview of the T-DNA integration positions in single-copy T-DNA *Arabidopsis* plants.

Plant	Chromosome	Target clone	Target deletion[a] (bp)	AT (%)	MAR	Intergenic region or in a gene (exon or intron)[b]	Sense or antisense orientation[b]
FK24	1	F2D10, At1g20830	12	62 (LB/RB)	no	exon	antisense
FH33[c]	1 and 4	F14M19 (LB), AP2 contig (RB)	ND	74 (LB), 75 (RB)	+2146 (LB), +2, 190 and 1874 (RB)	intergenic	–
F2K3[d]	1	T4K22	ND	72 (RB)	no	intergenic	–
F2K16[d]	2	F14B2	ND	74 (LB)	no	intergenic	–
F2Hsb20	4	F8B4, At4g32490	ND	52 (LB)	yes	intron	antisense
F2Hsb21	2	F13A10, At2g46530	ND	63 (RB)	no	intron	antisense
F2Hsb22	5	F12B17, At5g10460	ND	60 (LB)	no	exon	sense
F2Hsb31	4	L23H3	ND	75 (RB)	no	intergenic	–
CK2L6	5	4F5A18	11,317	72 (LB), 65 (RB)	no	intergenic	–
CK2L7	1	MQM1	51	68 (LB/RB)	no	intergenic	–
CK2L36	3	F12P19, At1g65880	14	62 (LB/RB)	no	exon	sense
CK2L70	1	F5E6, At3g06530	11	57 (LB/RB)	no	intron	sense
CK2L72	3	T18A20, At1g53800	103	58 (LB/RB)	no	exon	antisense
CK2L94	1	F4F15, At3g52030	39	61 (LB/RB)	yes	exon	sense
CK2L102	3	F14J22	23	68 (LB/RB)	yes, 2	intergenic	–
CK2L107	1	F21A14	83	40 (LB/RB)	no	intergenic	–
CK2L111	1	F14J16[e]	ND	58 (LB), 64 (RB)	no (LB), yes (RB)	exon	ND
CK2L129[c]	1 and 2	F2I9[f](LB), T22J18 (RB)	ND	67 (LB), 70 (RB)	yes (LB), no (RB)	intergenic	–
CK2L148	2	T30B22, At2g47430	29	60 (LB/RB)	no	exon	sense

[a] When both left and right T-DNA/plant junction positions could be determined, the target deletion upon T-DNA integration was calculated.
[b] For the T-DNAs that were integrated into a gene (exon or intron), the orientation of the T-DNA (sense or antisense) relative to the plant endogenous gene was determined.
[c] The integration position of these transgenic lines was characterized on the basis of the results of different clones. For transformant CK2L129, the LB and RB regions were fused to sequences of chromosomes 2 and 1, respectively, whereas for transformant FH33, to sequences of chromosomes 4 and 1, respectively.
[d] Vector backbone sequences were detected at the LB region in F2K3 and at the RB region in F2K16.
[e] In CK2L111, both RB and LB regions were localized in clone F14J16. The integration of the RB occurred in At1g55930 (position 70573). However, for the LB region, sequence analysis revealed that the insertions could have occurred at two positions with the same probability, namely on 93,801 or 88,146. Therefore, neither the integrated LB region nor the target deletion could be positioned exactly.
[f] The LB region of CK2L129 integrated into At2g01430 and the RB region into At1g22650.
ND, not determined.

screening for the presence of a bipartite sequence element associated with MAR sequences [37]. Vector backbone sequences were linked to the LB and RB T-DNA repeats in transformants F2K3 and F2K16, respectively (table 2). Finally, the flanking T-DNA sequences in transformant CK2L107 aligned with clone F21A14, which was located in the neighborhood of the 5S RNA repeat and lay directly next to the centromere (fig. 3). The distance between the T-DNA insertion position and the centromeric region was estimated to be 217 kb.

In conclusion, the integration position was unambiguously determined for 19 of the 21 single-copy T-DNAs. Nine of them integrated into an intergenic region and 10 within a transcribed annotated gene (table 2). Furthermore, all 19 single-copy T-DNAs integrated into euchromatin and none of them had tandemly repeated DNA sequences in the flanking regions.

Determination of the GUS activity in 21 single-copy T-DNA transformants

Before the transgene expression variability among the 21 different single-copy transformants was assayed, the variation in GUS activity caused by experimental factors was determined. The differences in GUS activity using three

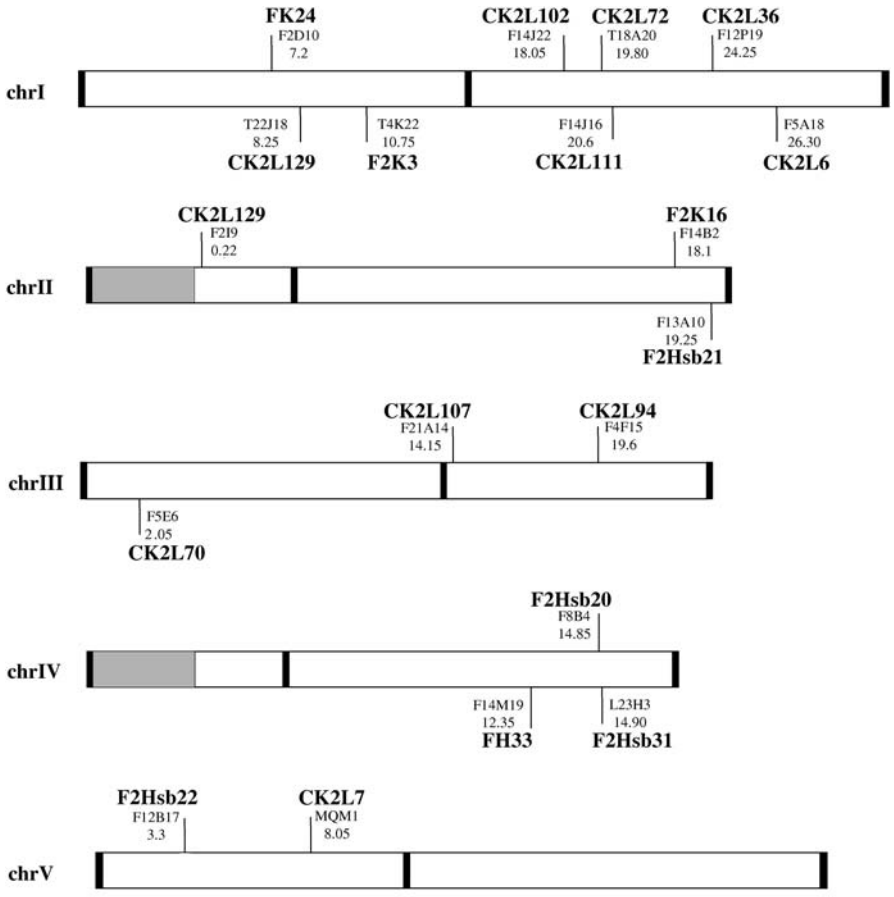

Figure 3. Distribution of the T-DNA insertions on the five chromosomes of the *Arabidopsis* genome. Centromeres and telomeres are indicated in vertical black boxes. The name of each T-DNA insertion is given in bold and grey boxes represent rDNA repeats. For each insertion, the sequence clone is mentioned.

independent measurements of protein extract ranged from a factor of 1.6 to a factor of 2.75 (data not shown). The variation in average expression of a sample can also be characterized by the coefficient of variation [CV= (standard deviation/average)×100%], which ranged from 23 to 51%.

In tobacco [20] and petunia [41], variation in transgene expression was observed between leaves of the same plant and even within an individual leaf. Therefore, we also investigated whether the single-copy lines had the same pattern of spatial expression by using histochemical GUS staining on 2- and 3-week-old seedlings (fig. 4 and data not shown). In all lines, except F2Hsb20/5, F2Ksb 18/1, and the silenced line KH15 [15], the GUS staining was homogeneous (fig. 4 and data not shown). The lines F2Hsb20/5, F2Ksb18/1, and the control line KH15, which exhibited low GUS activity, had a non-uniform *gus* staining (figs 4, 5). To determine whether this variation in staining was also reflected in the quantitative *gus* measurements, GUS activity was measured in different leaves of a silenced (KH15) and a non-silenced (CK2L7) transformant and compared with that obtained when a mixture of all leaves or only three leaves per plant were used to make the protein extract (table 3). In each case, a 2.1- to 3-fold difference in GUS activity between the individual leaves was observed, which equals a CV ranging from 24.85 to 43.95%. Such a CV is comparable to that for experimental errors, notwithstanding the great variation in histochemical GUS staining between the leaves of the KH15 plant. This result indicates that, especially for plants that were silenced [15], the intensity of GUS staining is not necessarily a good measure for GUS activity. When a mixture of all leaves of one plant was compared to a separate mixture of three leaves, the GUS activity was very analogous to that in the individual leaves; therefore, the results obtained in this analysis were not subjected to the random chance of selecting a leaf that had high or low expression.

To monitor the variability of transgene expression in single-copy T-DNA transformants of *Arabidopsis*, GUS accumulation levels were measured in five T1 and T2 plants derived from the 21 single-copy transformants (fig. 5A, B). For each plant, homozygosity or hemizygosity for the T-DNA insert was determined.

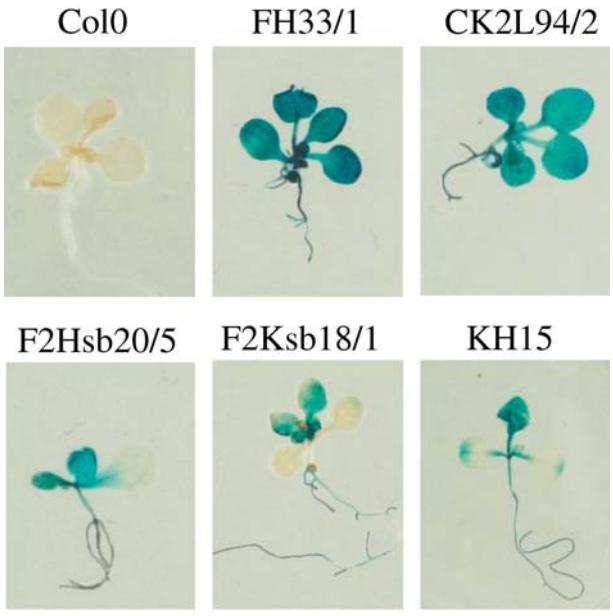

Figure 4. Histochemical GUS staining in 2-week-old seedlings. All single-copy T-DNA transformants that showed high GUS activity, such as FH33/1 and CK2L-94/2, were uniformly stained blue, whereas in transformants displaying low GUS activity (F2Hsb20/5, F2Ksb18/1, and the silenced plant KH15), a patchy pattern was observed. For both Col0 and C24 (data not shown), no blue staining was detected.

Table 3. Comparison of GUS activity levels measured in different leaves of silenced (KH15) and non-silenced (CK2L7) transformants.

Sample[a]	KH15[b]	CK2L7[b]
Leaf a	2.4	284
Leaf b	1.1	315
Leaf c	2.3	278
Leaf d	1.5	391
Leaf e	0.8	186
Leaf f	NA	236
Leaf g	NA	220
Three leaves of one plant in one extract	4	467
All leaves of one plant in one extract	1.65	ND

For the individual leaves of KH15, the mean value ± SD was 1.62 ± 0.7. For the individual leaves of CK2L7, the mean value ± SD was 273 ± 68. The CV was 43.95% for KH15 and 24.85% for CK2L7. NA, not analyzed, because this KH15 had only five leaves; ND, not determined.

[a] Three different sample types were compared. A sample was made from only one leaf (leaf a–g), from a mixture of three leaves per plant, or from all leaves of one plant.

[b] GUS activity is expressed as units GUS protein relative to the total amount of soluble extracted protein (U GUS mg protein^{-1}).

In eight transformants obtained after floral dip, GUS activity was high, varying from 420 to 6800 U GUS mg protein^{-1} in both the T1 and T2 generation (fig. 5A, B). Low GUS activity, below 10 U GUS mg^{-1}, was observed in the two remaining transformants, F2Ksb18 and F2Hsb20 (fig. 5A). All T1 and T2 F2Ksb18 plants had a low expression phenotype. Molecular analysis with the methylation-sensitive enzyme *Tae*I showed that in this line, the *gus*-coding sequences and more downstream sequences were strongly methylated (data not shown). In transformant F2Hsb20, low GUS activity was observed in the T1 and in some, but not all, T2 plants (fig. 5B). In this transformant, no methylation in the transcribed region could be detected (data not shown).

In parallel with the F transformants, 11 single-copy CK2L transformants, obtained after root explant transformation, were analyzed for their GUS accumulation in 3-week-old T1 plants (fig. 5C) and in 6-week-old T2 plants (fig. 5D). In all T1 progenies of the 11 transformants, high GUS activity was measured. Consistent and even higher expression was found for almost all transformants in the next generation (fig. 5B). None of the lines exhibited either promoter methylation or methylation in the transcribed sequences (data not shown).

To determine the variation based on physiological and environmental factors (intra-transformant variability), GUS activity was measured in isogenic homozygous T2 plants (fig. 5B, D). For the F transformants, the GUS activity levels of T2 plants of the same transformant differed from 1.6-fold (FH33/4) to 2.7-fold (F2Hsb31/3) (except for transformant F2K16/5, which had a 6.6-fold difference) (fig. 5B). The CV for the five isogenic plants in the seven homozygous lines ranged from 18.5 to 47.22%. Similar results were obtained for the CK2L T2 plants with a 1.2-fold (CK2L7/5 and CK2L111/4) to 2.5-fold (CK2L 70/1) difference or a CV varying between 9.08 and 39.73% (fig. 5D). These results indicate that the intra-transformant variability overlaps with the experimental variation.

In this context, the variability of GUS activity levels among the different transformants, the inter-transformant variability, can be considered. The average GUS activity in the homozygous F transformants varied between 1227 U GUS mg protein^{-1} (FK24/1) and 4468 U GUS mg protein^{-1} (F2Hsb31/3), which represents a 3.6-fold difference, whereas in the CK2L transformants, it varied from 1620 U GUS mg protein^{-1} (CK2L111/4) to 3236 U GUS mg protein^{-1} (CK2L94/2). Thus, the intra-transformant and inter-transformant variabilities are in the same range and comparable. These results suggest that the integration position has no major influence on the expression level of a 35S-driven transgene in a single-copy *Arabidopsis* transformant.

Transgene expression in a population of multicopy versus single-copy T-DNA plants

Figure 2F summarizes the GUS activity in 3-week-old plants of single-copy T-DNA transformants obtained after

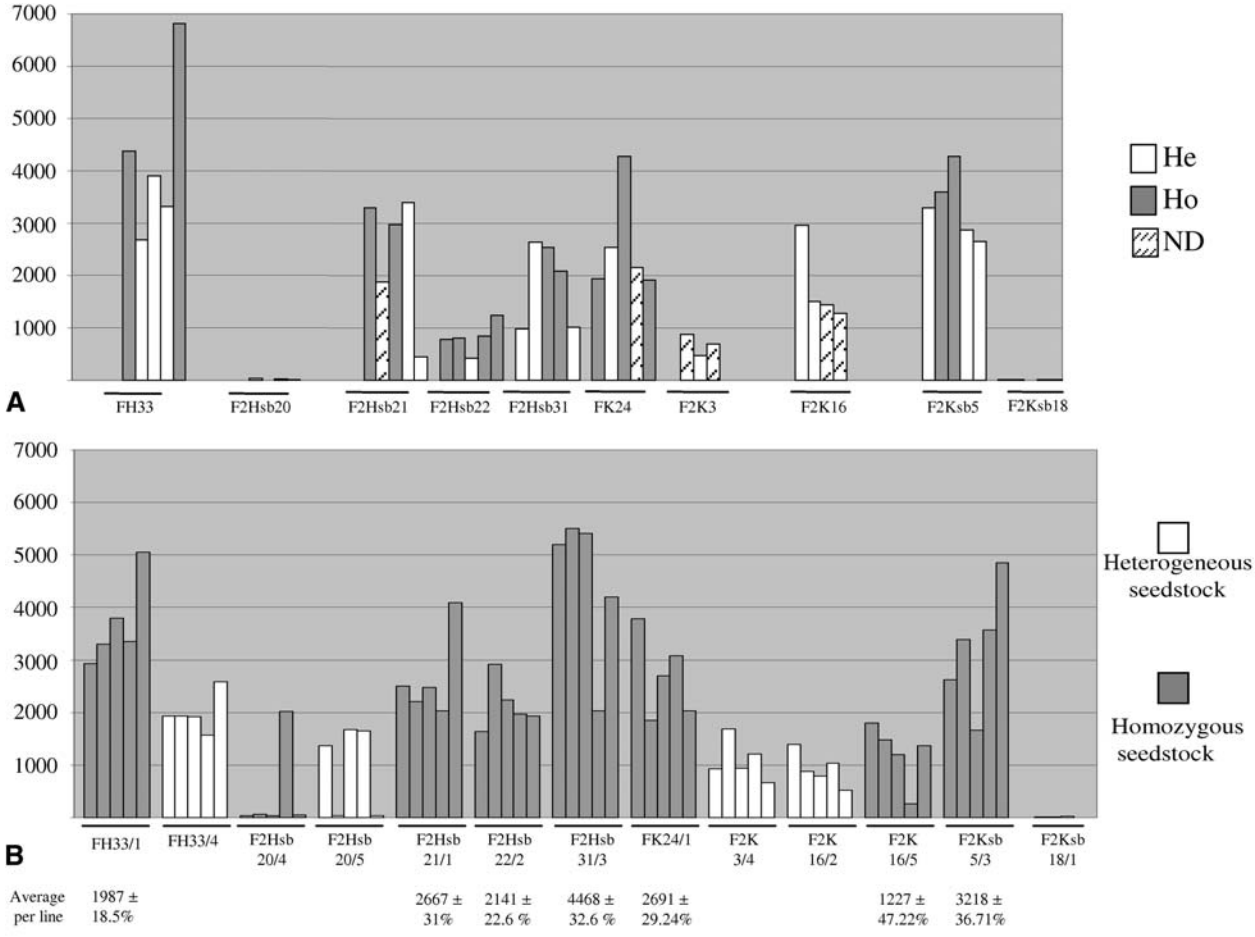

Figure 5. GUS activity analysis in the T1 and T2 generation of the single-copy T-DNA transformants. The FH, FK, F2K, F2Hsb, and F2Ksb transformants in the T1 (*A*) and T2 (*B*) generations and CK2L single-copy T-DNA transformants in the T1 (*C*) and T2 (*D*) generations are presented. In the T1 generation, GUS activity was measured in leaves of five 6-week-old T1 F plants (*A*) or 3-week-old T1 CK2L plants (*C*) per transformant. In the next generation, homozygosity (Ho) or hemizygosity (He) of the T-DNA insert was determined for each plants except for those marked by ND. In the T2 generation (*B*, *D*), GUS activity was measured in the leaves of five 6-week-old plants per transformant. Plants that were homozygous for the T-DNA insertions and those derived from a heterogeneous seed stock are represented by grey and white bars, respectively. The GUS activity levels are given as units GUS mg^{-1} of total soluble protein.

floral dip (see above). The majority of single-copy T-DNA transformants had a high and stable transgene expression. In contrast, 78 of the 100 transformants (fig. 2A) contained progeny plants with low transgene expression, confirming earlier observations that screening for single-copy T-DNA plants enriches for plants with high and stable transgene expression. However, not only single-copy plants exhibited high and stable expression. Progeny plants of 15 other transformants (asterisk in fig. 2B–E) also had high GUS activity, often even higher than in single-copy transformants. Six of these transformants (FK31, FK22, FK1, FK12, FH14, and F2Ksb24) had two T-DNAs at one locus, two had three or more T-DNA copies integrated at one locus (FK29 and FH38), and two contained T-DNAs at two different loci (FH29 and F2K1) (data not shown).

Discussion

Over the last 15 years, many studies have been devoted to the analysis of transgene expression in transgenic plants. Both variability and absence or strong reduction of transgene expression in a significant number of transformants are two obvious features described in the literature [42]. Because transgene repeat configuration and T-DNA arrangements are strongly correlated with the induction of gene silencing, screening for single-copy T-DNA transformants in a transgenic population may enrich for transformants with stable and high transgene expression [8]. To date, research has focused on the correlation between transgene expression and locus structure [8, 13–15, 43], and not integration position. Consequently, our study aimed at determining the transgene expression variability in different single-copy T-DNA transformants and at as-

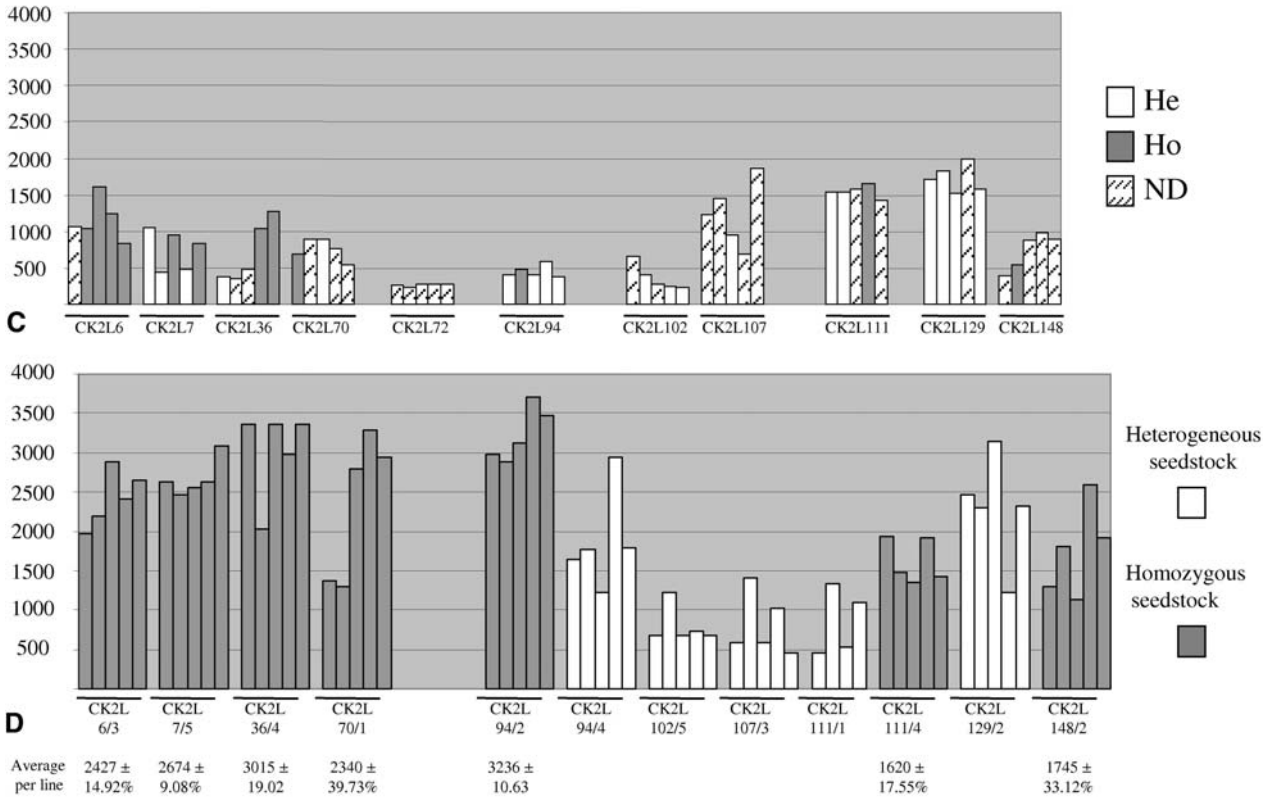

Figure 5 (continued)

sessing the impact of the integration position. Transgene expression was measured in a population of 100 transformants with single and multiple T-DNA copies integrated into one or several loci. In 78% of the transformants, low expressers were observed in one or more of the five analyzed T1 plants. From this population, 10 single-copy transformants were obtained. In addition, 11 single-copy CK2L transformants, obtained previously [31], were analyzed. Only 2 of the 21 single-copy transformants (9%) displayed low transgene expression in at least one of the five assessed plants per transformant. In the 19 transformants with high GUS activity, variation in *gus* expression was observed by a factor of 2 and 3.6, for CK2L and F transformants, respectively. Experimental errors and intra-transformant variability ranged from 1.6- to 2.7-fold (and even 6.6-fold). This value should be taken into account when assessing the inter-transformant variability. Based on this result, one can conclude that the integration position has a minor effect on transgene expression. The very low transgene expression in 2 of the 21 single-copy lines suggests that still unidentified characteristics of the integration position or other events during transformation can, at a low frequency, induce transgene silencing. The large difference in frequencies (9% compared to 78%) of low expression in single-copy versus multicopy T-DNA transformants shows that screening for single-copy T-DNA transformants strongly enriches for transformants with a stable and high transgene expression. In addition, the variation in transgene expression observed in various populations of transgenic plants is largely due to homology-dependent gene-silencing mechanisms that are triggered at a high rate in multicopy transformants [8, 29].

The literature shows that, even in single-copy T-DNA plants, many factors are responsible for the induction of low transgene expression. Beside the T-DNA integration position itself, linkage to repetitive sequences [44], the presence of neighboring highly methylated sequences [45], and the absence of MAR sequences near the integration site [26] are implicated as causative factors for the induction of low transgene expression. Additionally, an abrupt change in AT content [19] and the presence of vector backbone sequences [46] may reduce transgene expression. However, the present study did not provide evidence for a correlation with any of these factors.

High expressing single-copy T-DNAs were inserted randomly into the *Arabidopsis* genome (fig. 3, table 2). Nine of these T-DNAs integrated into intergenic regions and 10 within transcribed annotated genes (table 2). Of those inserted into annotated genes, seven inserted into exons and three into introns (table 2). The inserted *gus* genes were oriented both in antisense and sense orientation relative to the transcriptional orientation of the endogenous genes,

based on four and five transformants, respectively (table 2). Therefore, opposite transcription, resulting from the different orientation of transgene and endogenous genes does not greatly influence transgene expression nor does it induce gene silencing in single-copy T-DNA transformants.

In both transformants F2K3 and F2K16, in which vector backbone sequences are linked to the LB and RB, respectively, high and stable transgene expression was observed. Consequently, our results are in agreement with those of Meza et al. [47] and Kunz et al. [48], who demonstrated that binary vector sequences have no negative influence on transgene expression.

Transformants that possessed a T-DNA in the proximity of an endogenous MAR sequence (CK2L94, CK2L102, CK2L111, CK2L129, and FH33) did not differ in their expression profiles from those that did not (e.g., CK2L6, CK2L7, F2Hsb22, and F2Hsb31). Therefore, it is not possible to conclude whether integration in close proximity to endogenous MAR sequences is necessary for high transgene expression.

Finally, transgenes in *Drosophila melanogaster*, when integrated in the vicinity of heterochromatin, are known to become silenced by heterochromatization, a process called position effect variegation [49]. This heterochromatin silencing, acting as a global silencing mechanism, involves large-scale modifications of chromatin structure [50]. Of our work, the T-DNA in CK2L107 (fig. 3) inserted in the proximity of a centromere (fig. 3). However, transgene expression did not differ from that observed in the other high-expressing transformants. A single-copy *gus* transgene, integrated into pericentromeric heterochromatin, was recently reported to be stably expressed [40]. In parallel, integration in the vicinity of heterochromatic regions and the presence of moderately repetitive retroelement remnants do not necessarily result in low transgene expression [48]. However, these observations should be considered with caution because transformants, described in this analysis, were recovered based on the expression of the selectable marker, which lies on the same T-DNA. In addition, the *Arabidopsis* genome, when compared to other species, has a low amount of heterochromatic DNA [51]. In contrast to *Arabidopsis*, one-third of the *D. melanogaster* genome consists of centric heterochromatin [52]. The low number of heterochromatic regions in *Arabidopsis* may be one of the reasons why position effects were not observed.

In conclusion, we show that the position, and especially the T-DNA insertion within a gene instead of between genes, does not necessarily contribute to the transgene expression variability observed in single-copy transgenic *Arabidopsis* plants. Additionally, single-copy transformants are not always the highest expressers and multicopy transformants do not always trigger gene silencing. Of the 23 high expressers in the F population, only 8 were single-copy transformants and the remaining harbored multiple T-DNAs integrated into one or multiple loci. Indeed, high GUS activity was found in all seedlings of a significant number of transformants with two or even three integrated T-DNAs. On the other hand, the presence of more than two T-DNA copies, both under hemizygous or homozygous conditions, was recently reported to trigger silencing [43]. Therefore, the relationship between high copy number and low transgene expression seems to be far more complex than previously recognized.

Acknowledgments. The authors thank R. Whitford for critical reading and many editorial suggestions, G. Marjanac, I. Peck, G. Van der Auwera, G. Van Haute, and B. Van Droogenbroek for helpful comments, A. Blommaert, G. Tenea and A. Jacobs for technical assistance, S. Rombauts and Y. Van de Peer for help in determining the T-DNA integration position in the genome, and M. De Cock for help in preparing the manuscript. This work was supported by grants from the European Union BIOTECH program (QLRT-2000-00078), the European Community (18687-2001-11), and the Fund for Scientific Research-Flanders (G.0118.01).

1 Gelvin S. B. (2003) *Agrobacterium*-mediated plant transformation: the biology behind the 'gene-jockeying' tool. Microbiol. Mol. Biol. Rev. **67:** 16–37
2 Veluthambi K., Gupta A. K. and Sharma A. (2003) The current status of plant transformation technologies. Curr. Sci. **84:** 368–380
3 Komari T., Ishida Y. and Hiei Y. (2004) Plant transformation technology: *Agrobacterium*-mediated transformation. In: Handbook of Plant Biotechnology, vol. 1, pp. 233-261, Christou P. and Klee H. (eds), J. Wiley, Chichester
4 Kohli A., Twyman R. M., Abranches R., Wegel E., Stoger E. and Christou P. (2003) Transgene integration, organization and interaction in plants. Plant Mol. Biol. **52:** 247–258
5 Cheng M., Lowe B. A., Spencer T. M., Ye X. and Armstrong, C. L. (2004) Factors influencing *Agrobacterium*-mediated transformation of monocotyledonous species. In Vitro Cell. Dev. Biol.-Plant **40:** 31–45
6 Giddings G., Allison G., Brooks D. and Carter A. (2000) Transgenic plants as factories for biopharmaceuticals. Nat. Biotechnol. **18:** 1151–1155
7 De Block M. and Debrouwer D. (1991) Two T-DNA's co-transformed into *Brassica napus* by a double *Agrobacterium tumefaciens* infection are mainly integrated at the same locus. Theor. Appl. Genet. **82:** 257–263
8 Hobbs S. L. A., Warkentin T. D. and DeLong C. M. O. (1993) Transgene copy number can be positively or negatively associated with transgene expression. Plant Mol. Biol. **21:** 17–26
9 Cheng M., Fry J. E., Pang S., Zhou H., Hironaka C. M., Duncan D. R. et al. (1997) Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. Plant Physiol. **115:** 971–980
10 De Neve M., De Buck S., Jacobs A., Van Montagu M. and Depicker A. (1997) T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from ligation of separate T-DNAs. Plant J. **11:** 15–29
11 De Buck S., Jacobs A., Van Montagu M. and Depicker A. (1999) The DNA sequences of T-DNA junctions suggest that complex T-DNA loci are formed by a recombination process resembling T-DNA integration. Plant J. **20:** 295–304
12 Somers D. A. and Makarevitch I. (2004) Transgene integration in plants: poking or patching holes in promiscuous genomes? Curr. Opin. Biotechnol. **15:** 126–131

13 Jorgensen R. A., Cluster P. D., English J., Que Q. and Napoli C. A. (1996) Chalcone synthase cosuppression phenotypes in petunia flowers: comparison of sense vs. antisense constructs and single-copy vs. complex T-DNA sequences. Plant Mol. Biol. **31:** 957–973
14 Muskens M. W. M., Vissers A. P. A., Mol J. N. M. and Kooter J. M. (2000) Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. Plant Mol. Biol. **43:** 243–260
15 De Buck S., Van Montagu M. and Depicker A. (2001) Transgene silencing of invertedly repeated transgenes is released upon deletion of one of the transgenes involved. Plant Mol. Biol. **46:** 433–445
16 James V. A., Avart C., Worland B., Snape J. W. and Vain P. (2002) The relationship between homozygous and hemizygous transgene expression levels over generations in populations of transgenic rice plants. Theor. Appl. Genet. **104:** 553–561
17 Vain P., James V. A., Worland B. and Snape J. W. (2002) Transgene behaviour across two generations in a large random population of transgenic rice plants produced by particle bombardment. Theor. Appl. Genet. **105:** 878–889
18 Elmayan T. and Vaucheret H. (1996) Expression of single copies of a strongly expressed *35S* transgene can be silenced post-transcriptionally. Plant J. **9:** 787–797
19 Matzke A. J. M. and Matzke M. A. (1998) Position effects and epigenetic silencing of plant transgenes. Curr. Opin. Plant Biol. **1:** 142–148
20 Day C. D., Lee E., Kobayashi J., Holappa L. D., Albert H. and Ow D. W. (2000) Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced. Genes Dev. **14:** 2869–2880
21 Meyer P. (2001) Gene silencing in plants. In: Encyclopedia of Life Sciences, pp. 1–5, Nature, New York, www.els.net
22 Meza T. J., Kamfjord D., Hakelien A.-M., Evans I., Godager L. H., Mandal A. et al. (2001) The frequency of silencing in *Arabidopsis thaliana* varies highly between progeny of siblings and can be influenced by environmental factors. Transgen. Res. **10:** 53–67
23 Qin H., Dong Y. and von Arnim A. G. von (2003) Epigenetic interactions between *Arabidopsis* transgenes: characterization in light of transgene integration sites. Plant Mol. Biol. **52:** 217–231
24 Srivastava V., Ariza-Nieto M. and Wilson A. J. (2004) Cre-mediated site-specific gene integration for consistent transgene expression in rice. Plant Biotechnol. J. **2:** 169–179
25 Srivastava V., Anderson O. D. and Ow D. W. (1999) Single-copy transgenic wheat generated through the resolution of complex integration patterns. Proc. Natl. Acad. Sci. USA **96:** 11117–11121
26 Holmes-Davis R. and Comai L. (2002) The matrix attachment regions (MARs) associated with the *Heat Shock Cognate 80* gene (*HSC80*) of tomato represent specific regulatory elements. Mol. Genet. Genom. **266:** 891–898
27 Mankin S. L., Allen G. C., Phelan T., Spiker S. and Thompson W. F. (2003) Elevation of transgene expression level by flanking matrix attachment regions (MAR) is promoter dependent: a study of the interactions of six promoters with the RB7 3′ MAR. Transgen.Res. **12:** 3–12
28 De Bolle M. F. C., Butaye K. M. J., Coucke W. J. W., Goderis I. J. W. M., Wouters P. F. J., van Boxel N. et al. (2003) Analysis of the influence of promoter elements and a matrix attachment region on the inter-individual variation of transgene expression in populations of *Arabidopsis thaliana*. Plant Sci. **165:** 169–179
29 Butaye K. M. J., Goderis I. J. W. M., Wouters P. F. J., Pues J. M.-T. G., Delauré S. L., Broekaert W. F. et al. (2004) Stable high-level transgene expression in *Arabidopsis thaliana* using gene silencing mutants and matrix attachment regions. Plant J. **39:** 440–449
30 De Buck S., De Wilde C., Van Montagu M. and Depicker A. (2000) T-DNA vector backbone sequences are frequently integrated into the genome of transgenic plants obtained by *Agrobacterium*-mediated transformation. Mol. Breed. **6:** 459–468
31 De Buck S., De Wilde C., Van Montagu M. and Depicker A. (2000) Determination of the T-DNA transfer and the T-DNA integration frequencies upon cocultivation of *Arabidopsis thaliana* root explants. Mol. Plant-Microbe Interact. **13:** 658–665
32 Theuns I., Windels P., De Buck S., Depicker A., Van Bockstaele E. and De Loose M. (2002) Identification and characterization of T-DNA inserts by T-DNA fingerprinting. Euphytica **123:** 75–84
33 Clough S. J. and Bent A. F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743
34 Windels P., De Buck S., Van Bockstaele E., De Loose M. and Depicker A. (2003) T-DNA integration in *Arabidopsis* chromosomes: presence and origin of filler DNA sequences. Plant Physiol. **133:** 2061–2068
35 Altschul S. F., Gish W., Miller W., Myers E. W. and Lipman D. J. (1990) Basic local alignment search tool. J. Mol. Biol. **215:** 403–410
36 Huala E., Dickerman A. W., Garcia-Hernandez M., Weems D., Reiser L., LaFond F. et al. (2001) The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res. **29:** 102–105
37 van Drunen C. M., Sewalt R. G. A. B., Oosterling R. W., Weisbeek P. J., Smeekens S. C. M. and van Driel R. (1999) A bipartite sequence element associated with matrix/scaffold attachment regions. Nucleic Acids Res. **27:** 2924–2930
38 Breyne P., De Loose M., Dedonder A., Van Montagu M. and Depicker A. (1993) Quantitative kinetic analysis of β-glucuronidase activities using a computer-directed microtiter plate reader. Plant Mol. Biol. Rep. **11:** 21–31
39 Schoof H., Zaccaria P., Gundlach H., Lemcke K., Rudd S., Kolesov G. et al. (2002) MIPS *Arabidopsis thaliana* database (MAtDB): an integrated biological knowledge resource based on the first complete plant genome. Nucleic Acids Res. **30:** 91–93
40 Forsbach A., Schubert D., Lechtenberg B., Gils M. and Schmidt R. (2003) A comprehensive characterization of single-copy T-DNA insertions in the *Arabidopsis thaliana* genome. Plant Mol. Biol. **52:** 161–176
41 Van Leeuwen W., Ruttink T., Borst-Vrenssen A. W. M., van der Plas L. H. W. and Van der Krol A. R. (2001) Characterization of position-induced spatial and temporal regulation of transgene promoter activity in plants. J. Exp. Bot. **52:** 949–959
42 De Wilde C., Van Houdt H., De Buck S., Angenon G., De Jaeger G. and Depicker A. (2000) Plants as bioreactors for protein production: avoiding the problem of transgene silencing. Plant Mol. Biol. **43:** 347–359
43 Lechtenberg B., Schubert D., Forsbach A., Gils M. and Schmidt R. (2003) Neither inverted repeat T-DNA configurations nor arrangements of tandemly repeated transgenes are sufficient to trigger transgene silencing. Plant J. **34:** 507–517
44 Pröls F. and Meyer P. (1992) The methylation pattern of chromosomal integration regions influence gene activity of transferred DNA in *Petunia hybrida*. Plant J. **2:** 465–475
45 Meyer P., Heidmann I. and Niedenhof I. (1993) Differences in DNA-methylation are associated with a paramutation phenomenon in transgenic petunia. Plant J. **4:** 89–100
46 Iglesias V. A., Moscone E. A., Papp I., Neuhuber F., Michalowski S., Phelan T. et al. (1997) Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco. Plant Cell **9:** 1251–1264
47 Meza T. J., Stangeland B., Mercy I. S., Skårn M., Nymoen D. A., Berg A. et al. (2002) Analyses of single-copy *Arabidopsis* T-DNA-transformed lines show that the presence of vector

backbone sequences, short inverted repeats and DNA methylation is not sufficient or necessary for the induction of transgene silencing. Nucleic Acids Res. **30:** 4556–4566

48 Kunz C., Narangajavana J., Jakowitsch J., Park Y.-D., Delon T. R., Kovarik A. et al. (2003) Studies on the effects of a flanking repetitive sequence on the expression of single-copy transgenes in *Nicotiana sylvestris* and in *N. sylvestris-N. tomentosiformis* hybrids. Plant Mol. Biol. **52:** 203–215

49 Schotta G., Ebert A., Dorn R. and Reuter G. (2003) Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. Semin. Cell Dev. Biol. **14:** 67–75

50 Avramova Z. V. (2002) Heterochromatin in animals and plants: similarities and differences. Plant Physiol. **129:** 40–49

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

52 Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D., Amanatides P. G. et al. (2000) The genome sequence of *Drosophila melanogaster*. Science **287:** 2185–2195