Molecular Breeding **12:** 71–81, 2003.

Effects of genomic position and copy number of Acyl-ACP thioesterase transgenes on the level of the target fatty acids in *Brassica napus* L.

Jihong Tang, Rachael Scarth* and Brian Fristensky
*Department of Plant Science, University of Manitoba, Winnipeg, R3T 2N2, MB, Canada; *Author for correspondence (e-mail: Rachael_Scarth@umanitoba.ca; phone: 204-474-6082; fax: 204-474-7525)*

Received 13 June 2002; accepted in revised form 10 February 2003

Key words: Brassica napus L., Copy number, Fatty acid, Gene dosage, Genomic position, Thioesterase

Abstract

The effects of genomic position and copy number of acyl-acyl carrier protein (ACP) thioesterase (TE) transgenes on the major target fatty acid, either lauric acid (C12:0) or palmitic acid (C16:0) depending on the TE, in transgenic *Brassica napus* seed oil were investigated. Four transgenic parental lines, transformed individually with the bay-TE (*Uc FatB1*), elm-TE (*Ua FatB1*), nutmeg-TE (*Mf FatB1*) and *Cuphea*-TE (*Ch FatB1*) transgenes, were crossed with the non-transgenic recipient genotypes '212/86' or 'QO4'. Bay-TE and *Cuphea*-TE F_1 seeds, which carry half the number of the construct copies compared to the self-pollinated seeds of the transgenic parents, showed significantly lower levels of the target fatty acid. Doubled haploid (DH) lines were developed through microspore culture from F_1 hybrids with the elm-TE or the *Cuphea*-TE transgenes. DH lines carrying one to five copies of the *Cuphea*-TE transgene displayed a positive correlation between transgene copy number and the target fatty acid C16:0 level ($r = 0.77^{**}$). DH lines with elm-TE transgene copies at four different loci showed different C16:0 levels, with one of the loci (E-II) leading to significantly higher C16:0 levels. This study supports the importance of the selection of high transgene copy number and/or the optimum genomic integration site in order to achieve maximum expression levels of the target fatty acid in transgenic oil quality modification.

Abbreviations: ACP – acyl carrier protein, C12:0 – lauric acid, C16:0 – palmitic acid, DH – doubled haploid, DIG – digoxigenin, SP – self-pollinated, TE – acyl-acyl carrier protein thioesterase

Introduction

The fatty acid composition of seed oil of conventional canola cultivars (*Brassica napus* L.) is made up of more than 90% C18 fatty acids including stearic acid (C18:0), oleic acid (C18:1), linoleic acid (C18:2) and linolenic acid (C18:3), with traces of lauric acid (C12:0) and ca. 4% palmitic acid (C16:0) (Eskin et al. 1996; McVetty and Scarth 2002; Voelker et al. 1996). Transformation of canola plants with special TE genes cloned from several non-crop plant species has resulted in transgenic lines accumulating significant levels of C12:0 or enhanced levels of C16:0 in the seed oil (Jones et al. 1995; van der Hoeven et al. 1994; Voelker et al. 1997). The TE hydrolyzes the thioester bond between the elongating fatty acyl group and the co-factor acyl carrier protein (ACP), leading to premature termination of the fatty acid biosynthetic process and accumulation of fatty acids with a shorter chain length in the seed oil.

Wide variation in the levels of the fatty acids targeted by TE has been frequently observed among different transformants from the same transformation experiments. Among transformants with the bay-TE transgene (*Uc FatB1*), for example, the level of the target fatty acid, lauric acid (C12:0), varied from the background level (near zero) to up to 40% (Voelker et al. 1996). The C18:0 level of different transformants carrying a FatA TE (*Garm FatA1*), cloned from a mangosteen seed cDNA library, varied from the background level (ca. 2% C18:0) to 22% C18:0 (Hawkins and Kridl 1998). Wide variation was also

observed among plants transformed with other genes involved in the plant fatty acid biosynthetic pathway, e.g., genes coding for desaturases (Hitz et al. 1995; Knutzon et al. 1992).

The effects of genomic position and copy number on transgene expression have been reported in plants (Allen et al. 2000; De Neve et al. 1997; Gendloff et al. 1990; Hobbs et al. 1993; McCabe et al. 1999). Since foreign DNA integrates almost exclusively at random, non-homologous sites during transformation, some integrations may occur in transcriptionally active chromatin environments, others in condensed, transcriptionally inert chromatin regions (Mengiste and Paszkowski 1999). It is believed that transgenes in heterochromatic areas such as those surrounding centromeres are prone to silencing and give rise to reduced and/or variable expression (Allen et al. 2000; Maqbool and Christou 1999; Mengiste and Paszkowski 1999; Weiler and Wakimoto 1995). There are conflicting reports about the relationship between copy number and expression level (Hobbs et al. 1993; McCabe et al. 1999). The two variables have been shown to be negatively correlated (Cervera et al. 2000; Hobbs et al. 1993; Mannerlöf et al. 1997), not correlated (Bauer et al. 1998; Hobbs et al. 1993; McCabe et al. 1999), or positively correlated (Gendloff et al. 1990; Hobbs et al. 1993; McCabe et al. 1999; van der Hoeven et al. 1994; Voelker et al. 1996). In many reports, the results were mainly based on analyses of primary transformants (Cervera et al. 2000; Gendloff et al. 1990; Hobbs et al. 1990; Mannerlöf et al. 1997).

In the present study, the effects of genomic position and copy number of TE transgenes on the levels of the fatty acids targeted by the TE were studied by comparing self-pollinated (SP) seeds and F_1 seeds carrying the same TE for each of four TE transgenes, as well as by analyses of DH lines with different transgenic loci and copy numbers for each of two TE transgenes.

Materials and methods

Parental genotype and production of F_1 and SP seeds

The original seeds of four transgenic parental lines TL1, TL3, TL5 and TL6, as well as two non-transgenic *B. napus* breeding lines 212/86 and QO4, were kindly provided by Calgene Inc. (USA). TL1 and TL5 were homozygous for one transgene locus; TL3 and TL6 had three or more transgene loci (M. Sovero, pers. comm.). TL1, developed by *Agrobacterium*-mediated transformation with the bay-TE transgene *Uc FatB1* (Voelker et al. 1996, 1992), showed ca. 45% C12:0 in the seed oil of the parental plants (Table 1). TL3, TL5 and TL6, transformed with the elm-TE *Ua FatB1*, nutmeg-TE *Mf FatB1* (Voelker et al. 1997), and *Cuphea*-TE *Ch FatB1* (Jones et al. 1995), respectively, showed 27–36% C16:0. The breeding line 212/86 was the recipient genotype used in the transformation for the development of TL1, TL3 and TL6. QO4 was the recipient genotype for TL5. Like 212/86 and QO4, the three non-transgenic *B. napus* cultivars, Apollo (Scarth et al. 1995a), AC Excel (Rakow 1993) and Mercury (Scarth et al. 1995b) all have less than 5% C16:0 with no accumulation of C12:0 in the seed oil.

The four transgenic parental lines together with 212/86 and QO4 were grown in a growth room with day/night temperatures of 25/20 °C and a 16-h photoperiod and a 580 μE m^{-2} s^{-1} light intensity. F_1 seeds of four crosses were produced by crossing each of the four transgenic parents as the female with the male being the non-transgenic genotype used in the original transformation of the transgenic lines, 212/86 or QO4. F_1 seeds and SP seeds were harvested from individual parental plants at maturity and the fatty acid composition was determined as described below.

DH line development and planting

DH lines were developed by microspore culture followed by chromosome doubling based on the procedure described by Ferrie and Keller (1995). Microspores were isolated from unopened flower buds of transgenic hybrid plants originated from crosses of the elm-TE or the *Cuphea*-TE parental line as the female parent and Apollo, AC Excel or Mercury as the male parent as described above. Embryogenesis was induced by culturing microspores in induction media at 32.5 °C for 3 d. DH_1 plants, transgenic or non-transgenic, were self-pollinated to produce DH_2 seeds. DH_2 seeds were tested for the seed oil fatty acid composition. Transgenic lines carrying the TE transgenes, as characterized by enhanced levels of C16:0 (> 8.3% C16:0), were grown in a greenhouse and self-pollinated to produce DH_3 seeds.

To study the effects of genomic position and copy number, DH_3 seeds of 128 transgenic DH lines, including 101 *Cuphea*-TE transgenic DH lines, 27

Table 1. Transgenic and non-transgenic *B. napus* parental genotypes.

Genotype	Transgene/GenBank accession no.[2]	Transgene donor	Recipient genotype	Transgenic loci[3]	Target fatty acid[1]	
					C12:0(%)	C16:0(%)
Transgenic line:						
TL1	*Uc FatB1* (bay-TE)/M94159	Bay tree (*Umbellularia californica*)	212/86	1	45.2(1.5)	3.1(0.1)
TL3	*Ua FatB1* (elm-TE)/U65644	Elm tree (*Ulmus americana*)	212/86	3 or more	< 1	31.6(1.1)
TL5	*Mf FatB1* (nutmeg-TE)/U65642	Nutmeg (*Myristica fragrans*)	QO4	1	< 1	27.7(0.9)
TL6	*Ch FatB1* (*Cuphea*-TE)/U17076	Cuphea (*Cuphea hookeriana*)	212/86	3 or more	< 1	35.7(1.5)
Non-transgenic line/cultivar:						
212/86	–	–	–	–	< 1	5.1(0.3)
QO4	–	–	–	–	< 1	4.0(0.6)
Apollo	–	–	–	–	< 1	4.4(0.1)
AC Excel	–	–	–	–	< 1	3.9(0.3)
Mercury	–	–	–	–	< 1	3.0(0.1)

[1] The mean and the standard deviation (in parenthesis) of the fatty acid level is based on ca. ten individual plants grown in a growth room with controlled temperatures (20/15 °C at day/night) and a 16-h photoperiod since growing conditions of TE transgenic plants showed a significant effect on the fatty acid composition in a study about the stability of TE transgene expression (results not shown). [2]For details of the transgenes and transformation, see references (Voelker et al. 1996, 1992) for the bay-TE; (Voelker et al. 1997) for the elm-TE and the nutmeg-TE; and (Jones et al. 1995) for the *Cuphea*-TE. [3]The number of transgenic loci in the parental lines was based on M. Sovero (personal comm.).

elm-TE transgenic DH lines, and eight non-transgenic DH lines (with < 5.0% C16:0 in the seed oil, included as the control,) were grown in 15-cm pots, with one plant from each line, in a growth room. The growth room had day/night (16 h/8 h) temperatures of 20/15 °C and light intensity as described above. SP DH_4 seeds were harvested from each of the DH_3 plants and tested for the fatty acid composition, locus and copy number, as described below.

Preparation of DNA probes

The probes were prepared using a digoxigenin (DIG)-labeling procedure with the PCR DIG Probe Synthesis Kit (Roche, Germany). To prepare the *Cuphea*-TE probe, the TE transgene was amplified by PCR from genomic DNA of the transgenic parental line TL6 with the left primer PT1, 5′-ATTAGAGCCTCGGCTTCACTC-3′ and the right primer PT2, 5′-GGATCCCATTGGATGATCTTT-3′ designed based on the published sequence of the TE. The amplified DNA fragment was cloned with the pGEM®-T Easy Vector and JM109 Competent Cells (Promega, USA). A positive white colony was grown overnight in LB Broth medium and the plasmid DNA carrying the *Cuphea*-TE gene was extracted (Ausubel et al. 1995). With the plasmid DNA having the *Cuphea* TE gene as the template, a 1.1-kb internal fragment of the *Cuphea*-TE transgene was produced and DIG-labeled by PCR under the presence of DIG-dUTP. With the same procedure, a 0.9-kb probe for the elm-TE transgene was prepared with the primer E3, 5′-TCCACAACAGCACCATCATT-3′ and primer E2, 5′-CTTGCTGCAATCACGACTGT-3′, for the PCR DIG labelling reaction.

Plant genomic DNA isolation and Southern blotting

Plant genomic DNA was extracted from ca. 3 g of cotyledon and young leaves using a CTAB procedure (Ausubel et al. 1995; Kidwell and Osborn 1992). Genomic DNA (ca. 5 µg) was digested with a restriction enzyme *Nsi*I, separated by electrophoresis on 0.8% agarose gel, and blotted onto a nylon hybridization transfer membrane (Hybond-N+, Amersham). *Nsi*I has an unique site located in the napin promoter in the T-DNA region. The T-DNA engineered with the elm-TE or the *Cuphea*-TE gene was ca. 7.6 and 7.8

kb long, respectively (Jones et al. 1995; Voelker et al. 1997). The T-DNA region from the *Nsi*I site to the left border, which contained the *Cuphea*-TE or the elm-TE transgene, was 6.0 and 5.9 kb, respectively. Hybridization and detection was conducted following a DIG non-radioactive hybridization and chemiluminescent detection procedure as described in the product instructions (Roche, Germany) as follows. Hybridization was done overnight in DIG Easy Hyb solution at 50 °C in the presence of the TE probe, with stringency washes at 68 °C in a 2 × wash solution (2 × SSC, containing 0.1 SDS) for 2 × 30 min. After washing and blocking, the membrane was incubated with a dilution of Anti-Digoxigein-Fab fragment conjugated to alkaline phosphatase (AP). Followed by treatment with an 1:100 dilution of the AP substrate CDP-Star, the luminescent signal was recorded on X-ray film. Since the restriction enzyme *Nsi*I had only one site at one end of the TE transgene in the T-DNA region, a plant genomic restriction site was necessary to enable the production of a *Nsi*I fragment carrying the transgene copy. Thus the number of bands should indicate the copy number (Voelker et al. 1996). The sizes of the bands were estimated based on the 1 KB PLUS DNA Ladder (GibcoBRL, Canada) running alongside the digested DNA samples on the agarose gel.

Fatty acid analysis

The fatty acid composition of seed oils was determined by gas chromatography of the methyl ester derivatives of the fatty acids (Hougen and Bodo 1973). A sample of 10 seeds was picked randomly from each of the plants to be tested, and the oil of the 10 seeds was extracted overnight with 1 ml heptane. Then, 300 µl of 0.5 M sodium methoxide was added for methyl ester derivitization. The oven temperature was programmed to increase from 190 to 230 °C. The level of a fatty acid was reported as the percentage of the total fatty acids in the seed oil.

Multiple comparisons among mean fatty acid levels and correlation analysis between copy number and fatty acid level were performed as described by Ott (Ott 1993).

Table 2. Mean levels (%) of the target fatty acids in F_1 seeds from crosses of transgenic parental lines TL1, TL3, TL5 and TL6 with non-transgenic *B. napus* genotypes 212/86 or QO4 and in self-pollinated (SP) seeds of the transgenic parents.

Cross	Parameter[1]	SP seeds	F_1 seeds	Transgene
TL1 × 212/86	C12:0,%	26.3 a[2]	16.2 b	Bay-TE
	rep.	2	3	
TL3 × 212/86	C16:0,%	24.7 a	23.3 a	Elm-TE
	rep.	4	4	
TL5 × QO4	C16:0,%	14.4 a	12.2 a	Nutmeg-TE
	rep.	4	7	
TL6 × 212/86	C16:0,%	25.2 a	16.5 b	*Cuphea*-TE
	rep.	4	5	

[1]For SP seeds, a parental plant was a replicate (rep.); for F_1 seeds, a cross between a male and a female parental plant in pair was a rep. [2]The means followed by the same letter within the same row are not significantly different according to Fisher's LSD test at the 0.05 level.

Results

Effect of copy number on the level of the target fatty acids in SP and F_1 seeds

The mean level of the target fatty acid, C12:0 for the bay-TE and C16:0 for the other three TE transgenes, was determined for SP seeds of the four transgenic parental lines and for F_1 seeds of the four crosses between the four transgenic parental lines and the non-transgenic breeding lines (Table 2). The number of transgene copies in F_1 seeds was half that in the SP seeds because the F_1 seeds were produced by cross-pollinating the transgenic lines with non-transgenic plants. Therefore, a comparative analysis of the SP and F_1 seeds provides evidence for the effect of transgene copy number on the level of the target fatty acids of the TE.

SP seeds carrying the bay-TE or the *Cuphea*-TE showed a significantly higher level of the target fatty acid, C12:0 or C16:0, respectively, than the corresponding F_1 seeds carrying the same TE (Table 2). Since the F_1 and SP seeds of the same TE had the same recipient genotype, the difference in the target fatty acid level between the SP and the F_1 seeds should be caused by the different numbers of TE transgene copies. For the elm-TE and the nutmeg-TE transgenes, no significant differences were detected between the SP seeds and the F_1 seeds, indicating that the extra copies in the SP seeds did not affect the production of the target fatty acid C16:0.

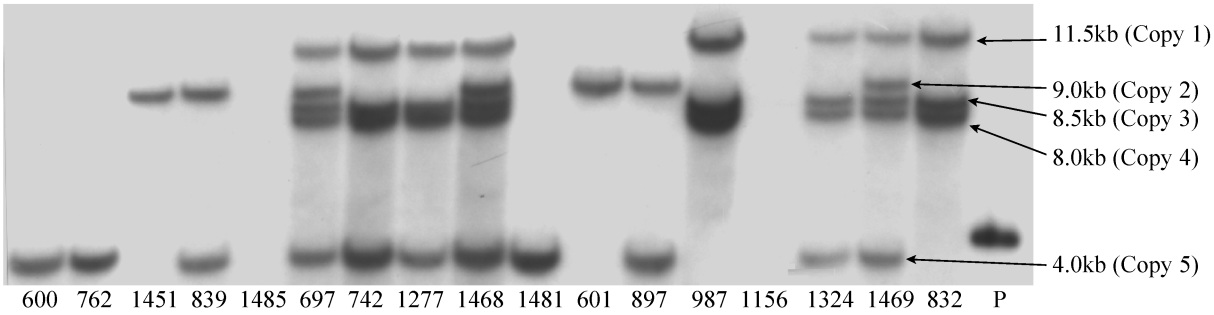

Figure 1. Southern blotting analysis of DH lines developed from crosses between the *Cuphea*-TE transgenic parental line TL6 and non-transgenic cultivars Apollo, AC Excel or Mercury. The identification no. of the DH lines are presented below the lanes. P, a digested plasmid carrying the *Cuphea*-TE transgene. The estimated sizes of the bands are shown in kilobase (kb).

Segregation of locus and copy number in Cuphed-*TE DH lines*

Southern blotting analysis was conducted for 109 DH lines developed from crosses between the *Cuphea*-TE parental line TL6 and the non-transgenic cultivars (Apollo, AC Excel and Mercury), which included 101 lines accumulating enhanced levels of palmitic acid (C16:0) targeted by the *Cuphea*-TE transgene. The C16:0 levels ranged from 9.4% to 34.5%. Since non-transgenic canola plants have only ca. 4% C16:0, these lines were putatively transgenic. As well, eight lines with 3.3–5.1% C16:0 were tested as negative control for the transgene.

As expected, all the 101 lines with enhanced levels of C16:0 showed at least one band with the *Cuphea*-TE gene probe, and all the negative control lines showed no band, e.g., plants No. 1485 and 1156 (Figure 1). Five bands of different lengths were detected with the transgenic lines, which were approximately 11.5, 9.0, 8.5, 8.0, 4.0 kb long, and were referred to as Copies 1, 2, 3, 4 and 5, respectively. The 4.0-kb band was smaller than the expected minimum length of approximately 6.0 kb since this T-DNA copy was incomplete, which showed the *Cuphea*-TE but not the selectable marker kanamycin resistance gene by Southern blotting and PCR analyses (J. Tang et al. in review). Copies 1, 3 and 4 co-segregated; e.g., DH lines 697, 742, 1277 all displayed these three copies. None of the 101 transgenic DH lines showed only one or two of these three copies. Therefore it can be concluded that these three copies were closely linked or tandem repeats at a genomic site. The putative site with the three TE transgene copies was designated Locus C-I. Copies 2 and 5 segregated from each other as well as from the other three copies as shown by the DH lines with only Copy 2 (e.g., DH line No. 1451 and 601) or only Copy 5 (No. 600 and 762). This indicates that Copies 2 and 5 were located at two additional genomic sites, designated Loci C-II and C-III, respectively.

Segregation analysis was conducted to determine whether the three transgene loci segregated independently or were linked. Assuming Loci C-I, C-II and C-III were completely independent, the population of DH plants developed from hybrid plants hemizygous for the three loci should be made up of eight genotypic classes of plants, with one non-transgenic and seven transgenic genotypic classes. The seven transgenic classes were: 1) lines with C-I; 2) lines with C-II; 3) lines with C-III; 4) lines with C-I and C-II; 5) lines with C-I and C-III; 6) lines with C-II and C-III; and 7) lines with C-I, C-II and C-III. In addition, DH plants of each of the transgenic genotypes should occur in an equal frequency in the population. The observed frequencies of the seven transgenic classes among 75 DH lines that originated from a single transgenic parental plant (TL6.12), were 10, 11, 5, 7, 17, 13, 12, respectively. These frequencies were not significantly different from the expected equal frequency with three independently segregating loci by χ^2-test ($p = 0.19$), supporting the conclusion that the three TE transgene loci were segregating independently.

Based on the Southern blotting analysis and segregation analysis, it can be concluded that the five *Cuphea*-TE transgene copies, detected in the *Cuphea*-TE DH lines, were present at three independently segregating loci, designated as Locus C-I, C-II and C-III, with Locus C-II and C-III each having only one copy of the *Cuphea*-TE transgene. Locus C-I had three closely linked or tandem repeated transgene copies, which behaved as a single locus.

Table 3. Mean palmitic acid (C16:0) level (%) of transgenic DH lines with different loci and copy numbers of the *Cuphea*-TE transgene developed from crosses of the *Cuphea*-TE transgenic parental line TL6 with non-transgenic cultivars Apollo or AC Excel.

Class No.	1	2	3	4	5	6	7	8
Locus [1]	C-I	C-II	C-III	C-I + C-II	C-I + C-III	C-II + C-III	C-I + C-II + C-III	Control
No. of copies	3	1	1	4	4	2	5	0
No. of DH lines	11	15	7	6	16	18	10	8
C16:0 (%)[2]	19.4 cd	15.5 e	16.1 de	21.3 bc	22.8 b	19.8 c	26.1 a	4.1

[1] The three distinct loci, C-I, C-II and C-III, were defined based on Southern blotting analyses. Locus C-1 had three copies, Copies 1, 3 and 4; Locus C-II had Copy 2; Locus C-III had Copy 5. [2]The means followed by the same letters were not significantly different according to Fisher's LSD test at the 0.05 level. Non-transgenic control plants were not included in the statistical test.

Effect of genomic position in Cuphea-*TE DH lines*

A total of 83 DH lines, developed from crosses with Apollo or AC Excel as the non-transgenic parent, were divided into seven transgenic genotypic classes based on the Southern blotting results, and the mean C16:0 level for each class was determined (Table 3). Of the 83 DH lines, the mean C16:0 level of the 56 DH lines from Apollo was 20.2%, not significantly different from that of the 27 DH lines from AC Excel (19.9% C16:0), indicating that the phenotypic influence of Apollo and AC Excel as non-transgenic parent on the expression of the *Cuphea*-TE transgene was the same.

The expression of the *Cuphea*-TE transgene at Locus C-II was not significantly different from the expression at Locus C-III based on the C16:0 level of the DH lines (Table 3). The mean C16:0 level of the 15 transgenic DH lines with one copy at Locus C-II was the same as the mean C16:0 level of the 7 DH lines with one copy at Locus C-III (15.5% and 16.1% C16:0, respectively). DH lines in Classes 4 and 5 had four transgene copies but with different distributions in the genome. Lines in Class 4 each had a copy at Locus C-II and lines in Class 5 each had a copy at Locus C-III, with the lines in the both classes having the three copies at Locus C-1. These two classes had similar mean levels of C16:0.

The DH lines with only Locus C-I showed a higher mean level of C16:0 (19.4%) than the DH lines with only Locus C-II and the lines with only Locus C-III (Table 3). However, it could not be determined, based on the comparison here, whether the distinct genomic position of Locus C-I was the reason for the higher level of C16:0 or whether the higher level was the result of the two extra copies at this locus.

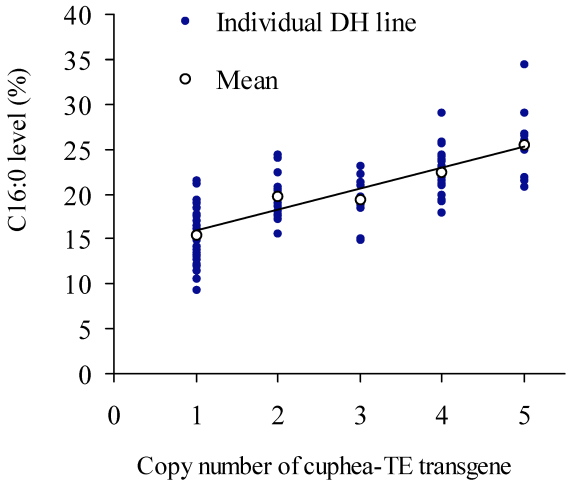

Figure 2. Palmitic acid (C16:0) level (%) of DH lines with different copy numbers of the *Cuphea*-TE transgene developed from crosses of the *Cuphea*-TE transgenic parental line TL6 with non-transgenic cultivars Apollo or AC Excel.

Effect of copy number in Cuphea-*TE DH lines*

Individual DH lines with the same copy numbers of the *Cuphea*-TE transgene showed wide variation in the C16:0 level as shown by the scatter plot of the DH lines (Figure 2). The C16:0 levels of the 22 single-copy DH lines, i.e., DH lines in Class 2 and 3 (Table 3), ranged from 10.6% to 21.2%, a two-fold difference. DH lines with 2 to 5 copies showed a ca. 1.5-fold difference among lines with the same copy number (Figure 2).

Although there was wide variation among the DH lines within each class of copy number, the mean C16:0 level of the DH lines increased as the copy number increased (Figure 2). The DH lines with one copy showed a mean C16:0 level of 15.7%, the weighted mean of DH lines in Class 2 and 3 (Table 3). The mean level of C16:0 was increased to 19.8%, 22.4% (the weighted mean of Classes 4 and 5 in Ta-

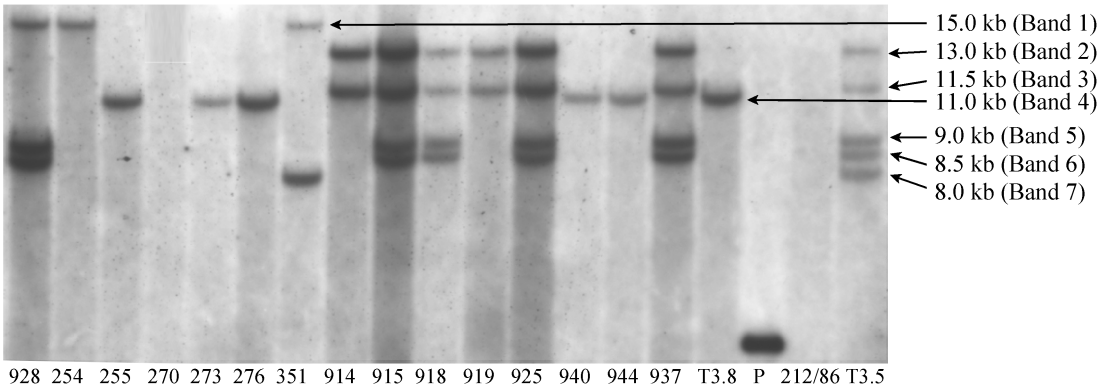

Figure 3. Southern blotting analysis of DH lines developed from crosses of the elm-TE transgenic parental line TL3 with non-transgenic cultivars Apollo, AC Excel or Mercury. Identification numbers of the DH lines are presented under the lanes. P, a plasmid cloned with the elm-TE transgene; 212/86, a non-transgenic breeding line. The estimated sizes of the bands are shown in kilobase (kb).

ble 3) and 26.1%, respectively, with two, four and five copies. Correlation analysis between copy number and C16:0 level of the 83 DH lines showed that the two variables were significantly positively correlated (r = 0.77**).

In addition, the maximum C16:0 level increased with increased copy numbers in the DH lines (Figure 2). The highest C16:0 level found among the 22 DH lines with one copy was 21.2%. The highest C16:0 level increased to 24.5, 25.8 and 34.5%, respectively, in DH lines with two, four and five copies.

Segregation of locus and copy number in elm-TE DH lines

A total of 27 DH lines with the elm-TE transgene were tested for the locus and copy number by Southern blotting analysis. Seven distinct bands, 15.0, 13.0, 11.5, 11.0, 9.0, 8.5 and 8.0 kb in size, were detected, and were referred to as band 1–7, respectively (Figure 3). The non-transgenic genotype 212/86 did not show any of these bands. Based on the banding patterns, the 27 DH lines were divided into eight groups with distinct transgene genotypes: 1) three lines with band 1; 2) three lines with bands 2 and 3; 3) four lines with bands 5 and 6; 4) one line with bands 1, 5 and 6; 5) six lines with bands 2, 3, 5 and 6; 6) eight lines with band 4; 7) one line with band 7; and 8) one line with bands 1 and 7.

The seven bands represented five elm-TE transgene loci. Band 1 segregated from the other bands, with three lines displaying only this band, e.g., plant No. 254 (Figure 3), indicating that the elm-TE transgene copy represented by this band was at a distinct genomic site, and was designated Locus E-I. Bands 2 and 3 were closely linked, showing no segregation. Nine lines showed these two bands (Groups 2 and 5) and the others did not show either band. This was interpreted as a second genomic site, designated Locus E-II. Based on similar analyses, the copy represented by band 4 was concluded to be at a third genomic site, designated Locus E-III; bands 5 and 6 represented linked copies at a fourth genomic site designated Locus E-IV. Only two DH lines showed band 7 (Groups 7 and 8), with one line (Group 7) displaying only this band, indicating the transgene copy represented by this band was at a different site, designated Locus E-V. Because there were no replications for the single lines in Groups 4, 7 and 8, these three lines were not included in the subsequent analyses.

Effect of genomic position in the elm-TE DH lines

The mean C16:0 levels of five groups of elm-TE DH lines are presented in Table 4. All the lines in Groups 1, 2, 3, 5 originated from the cross between the elm-TE transgenic parental line TL3 and the non-transgenic parent AC Excel. Of the six lines in Group 6, four lines originated from Apollo as the non-transgenic parent and two lines from AC Excel as the non-transgenic parent. The mean C16:0 level of the four lines from Apollo was not significantly different from that of the two lines from AC Excel (25.8% and 23.7%, respectively).

All of the DH lines with only one of the four loci, Locus E-I, E-II, E-III or E-IV, showed higher C16:0 level than the non-transgenic control (ca. 4%), e.g.,

Table 4. Mean palmitic acid (C16:0) level for five groups of transgenic DH lines with different loci and copy numbers of the elm-TE transgene.

Group	Locus[1]	Copy number	C16:0 (%)[2]	No. of DH lines
1	E-I	1	10.8 c	3
2	E-II	2	36.7 a	3
3	E-IV	2	28.7 b	4
5	E-II + E-IV	4	35.2 a	6
6	E-III	1	25.2 b	6

[1]The four distinct loci, E-I to E-IV, were defined based on the Southern blotting analyses. E-I and E-IV had one copy of the elm-TE transgene; E-II and E-III had two copies. [2]The means followed by the same letter are not significantly different according to Fisher's LSD test.

the three DH lines with Locus E-I had 10.8% C16:0 on average, indicating that the elm-TE transgene was expressed at each of the four loci.

The influence of genomic position on the target fatty acid was observed in the elm-TE DH lines. DH lines with the TE transgene at different genomic positions showed significantly different levels of C16:0 (Table 4). The mean C16:0 level of the DH lines with only Locus E-I (10.8%) was significantly lower than the mean C16:0 level of the DH lines with only Locus E-III (25.2% on average) although each of the two loci had one copy of the elm-TE transgene. In addition, the mean C16:0 level of the DH lines with only Locus E-II (Group 2) was significantly different from that of the DH lines with only Locus E-IV (Group 3) although the two loci had the same copy number.

Effect of copy number in elm-TE DH lines

Transgenic lines with two elm-TE transgene copies showed higher levels of C16:0 than lines with one copy on average (Table 4). The six lines with two copies at either locus E-II or E-IV had a mean C16:0 level of 32.1%, which was much higher than the mean C16:0 level of the nine lines (20.51%) with one copy at either Locus E-I or E-III. However, it was not clear whether the higher levels of C16:0 in the DH lines with two copies at Locus E-II or E-IV, compared to single copy lines, were caused by the extra transgene copies or by the different genomic positions.

When the copy number of the elm-TE transgene increased from two to four, the C16:0 level did not necessarily increase, depending on the loci involved (Table 4). The mean C16:0 level of the DH lines with four copies at Locus E-II and E-IV (35.2% on average) was higher than the mean C16:0 level of the lines with only two copies at Locus E-IV (28.7%), but was not different from that of the DH lines with two copies at Locus E-II (36.7%). This indicated that when a locus leading to a higher C16:0 level (e.g., Locus E-II) was present, addition of extra copies of the elm-TE transgene had no effect on the production of the target fatty acid. When a locus with a relatively lower expression level (e.g., Locus E-IV) was present, additional elm-TE transgene copies could increase the C16:0 level.

Discussion

Effect of genomic position

Three independently segregating *Cuphea*-TE transgene loci were identified in the *Cuphea*-TE transgenic DH lines. Two of the loci (Locus C-II and C-III), each having a single copy of the *Cuphea*-TE transgene, were not different based on comparative analysis of the level of the target fatty acid in the transgenic DH lines carrying Locus C-II or C-III. However, the DH lines with the elm-TE transgene at different genomic positions showed different levels of the target fatty acid. Positional effect has been widely associated with variation in transgene expression as frequently observed among primary transformants from the same transformation experiment in a number of plant species (Allen et al. 2000; De Neve et al. 1997; Gendloff et al. 1990). The chromatin structure (heterochromatin vs. euchromatin), existence of endogenous regulator sequences such as enhancers, and discrepancy in the GC content between a transgene and the flanking sequences at the integration site are believed to influence transgene expression (Allen et al. 2000; Chandler and Vaucheret 2001; Fagard and Vaucheret 2000). Significant differences in the level of the fatty acid targeted by the elm-TE transgene between DH lines with different transgenic loci suggested a positional effect on the expression, possibly

due to different flanking sequences at transgene insertion sites. The observation that the DH lines with the *Cuphea*-TE transgene inserted at different genomic sites had similar levels of the target fatty acid does not exclude the possibility that this gene is influenced by positional effects. It is possible that each of the insertion sites studied had comparable chromatin properties, leading to similar effects on the transgenes. This hypothesis could reconcile conflicting reports in the literature, in which positional effect was reported to influence the level of C12:0 targeted by the bay-TE transgene in *B. napus* (Voelker et al. 1996), while no significant positional effect was detected in some studies with other transgenes in other plant species (Gendloff et al. 1990; Hobbs et al. 1990).

Effect of copy number

Comparisons between SP seeds and F_1 seeds of the bay-TE or the *Cuphea*-TE transgenes, among transgenic DH lines with different numbers of the *Cuphea*-TE transgene, as well as among lines with different numbers of the elm-TE transgene, all showed the effect of transgene copy number on the target fatty acid.

Dosage effect of transgene copies has been demonstrated by comparing hemizygous and homozygous transgenic plants, where homozygotes are expected to have double the number of transgene copies compared to the hemizygotes in some studies in several plant species (Azhakanandam et al. 2000; Beaujean et al. 1998; Hobbs et al. 1993; Tenllado and Diaz Ruiz 1999). For example, transgenic tobacco (*Nicotiana tabacum*) F_1 plants, originated from cross-fertilization between two different homozygous transformants with high expression inserts, showed an expression level equal to the parents. Individual F_2 plants expressed the transgenes at 50%, 100%, 150% and 200% of parent values as the copy number in the whole genome increased from 1 to 4, whether the copies were allelic or non-allelic (Hobbs et al. 1993). In the present study, the higher levels of the target fatty acids in SP seeds than F_1 seeds suggested a positive effect of extra TE transgene copies on the level of the target fatty acids.

The mean level of the target fatty acid C16:0 in the *Cuphea*-TE DH lines increased as the copy number increased from one to five copies per haploid genome. Analysis based on 83 DH lines with one to five copies of the *Cuphea*-TE transgene showed a positive correlation between copy number and the target fatty acid level ($r = 0.77^{**}$), indicating that copy number accounted for nearly 60% ($R^2 = 0.593$) of the total variation in the level of the target fatty acid among the *Cuphea*-TE DH lines. This observation is supported by previous reports regarding the relationship between copy number and expression level with other transgenes (Gendloff et al. 1990; Hobbs et al. 1993; McCabe et al. 1999; van der Hoeven et al. 1994). Transgenic tobacco plants showed increased expression levels as the copy number of the chloramphenicol acetyl transferase (CAT) transgene increased from 1 to 4, although some plants with one copy showed a higher CAT activity than those with more copies (Gendloff et al. 1990). The variation in the target fatty acid level which could not be explained by copy number of the TE transgenes might be due to plant to plant variation within the same DH lines. Only one plant was tested for the fatty acid level for each DH line in this experiment, although there were replicates for DH lines with the same locus and copy number. Plant to plant variation within the same TE transgenic DH lines was demonstrated in a study focused on the stability of TE transgene expression (Tang 2002). As well, variation in transgene expression was observed between asexually propagated transgenic tobacco plants (Bhattacharyya et al. 1994).

DH lines carrying the elm-TE transgene displayed a complex relationship between copy number and the level of the target fatty acid C16:0. When the target fatty acid level was high (e.g., 36.7% C16:0 for the lines with two copies at Locus E-II), additional transgene copies did not further increase the target fatty acid level. When the target fatty acid level was relatively low (e.g., 28.7% C16:0 for the lines with two copies at Locus E-IV), the target fatty acid level increased by additional transgene copies. Thus, there appeared to be an upper limit, ca. 36.7% C16:0 in this study and 33% C16:0 in the primary transformants (Voelker et al. 1997), for the level of C16:0 in transgenic lines with the elm-TE transgene.

Studies with *B. napus* bay-TE transgenic plants showed that the activity of the enzyme had a linear positive correlation with the level of C12:0 targeted by the enzyme until the level reached 40% (Voelker et al. 1996). Above 40%, the linear relationship was lost and the C12:0 level increased at a lower and decreasing rate as the enzyme activity increased, with an upper limit of ca. 60% C12:0 in bay-TE transgenic lines (Voelker et al. 1996). Other enzymes, e.g. β-ketoacyl ACP synthase (KAS), which catalyzes the elongation of fatty acyl chain-ACP and acyl-ACP

acyltransferase (LPAAT), which assembles fatty acids into triacylglycerols, have been proposed as the factors limiting further increase in the target fatty acid level in TE transgenic plants (Eccleston and Ohlrogge 1998; Hawkins and Kridl 1998; Voelker et al. 1997; Wiberg et al. 1997), although there is no experimental support. The positive linear correlation observed in this study between the copy number of the *Cuphea*-TE transgene and the level of the target fatty acid C16:0 could represent a similar relationship between the enzyme activity and the target fatty acid level as was observed with the bay-TE transgenic lines. All the DH lines with the *Cuphea*-TE transgene had less than 40% C16:0 in this study; thus, the activity of the *Cuphea*-TE enzyme was still a limiting factor. Further study is needed to determine if an increase in the enzyme activity by expression of more copies could result in higher levels of C16:0. However, for the elm-TE transgenic lines, the upper limit of the target fatty acid was lower than that of the bay-TE and *Cuphea*-TE transgenic lines. Expression of the elm-TE transgene at an appropriate genomic position with a limited number of copies (e.g., Locus E-II with two copies) produced sufficient C16:0 to reach the limiting level. Additional elm-TE transgene copies did not significantly increase the fatty acid level, illustrating the important relationship of transgene copy number and genomic position in determining the expression level. The practical significance of this study is that it is possible to maximize the expression of a transgenic phenotype such as fatty acid level by selecting transgenic lines with multiple transgene copies or with different transgene loci. An increase in transgene copy numbers could also be achieved by crossing different transformants with a lower transgene copy number.

References

Allen G.C., Spiker S. and Thompson W.F. 2000. Use of matrix attachment regions (MARs) to minimize transgene silencing. Plant Mol. Biol. 43: 361–376.

Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A. et al. 1995. Current Protocols in Molecular Biology. Vol. 1. John Wiley & Sons, Inc., New York.

Azhakanandam K., McCabe M.S., Power J.B., Lowe K.C., Cocking E.C. and Davey M.R. 2000. T-DNA transfer, integration, expression and inheritance in rice: effects of plant genotype and *Agrobacterium* super-virulence. J. Plant Physiol. 157: 429–439.

Bauer M., Libantova J., Moravcikova J. and Bekesiova I. 1998. Transgenic tobacco plants constitutively expressing acidic chitinase from cucumber. Biologia 53: 749–758.

Beaujean A., Sangwan R.S., Hodges M. and Sangwan-Norreel B.S. 1998. Effect of ploidy and homozygosity and transgene expression in primary tobacco transformants and their androgenetic progenies. Mol. Gen. Genet. 260: 362–371.

Bhattacharyya M.K., Stermer B.A. and Dixon R.A. 1994. Reduced variation in transgene expression from a binary vector with selectable markers at the right and left T-DNA borders. Plant J. 6: 957–968.

Cervera M., Pina J.A., Juárez J., Navarro L. and Peña L. 2000. A broad exploration of a transgenic population of citrus: stability of gene expression and phenotype. Theor. Appl. Genet. 100: 670–677.

Chandler V.L. and Vaucheret H. 2001. Gene activation and gene silencing. Plant Physiol. 125: 145–148.

De Neve M., De Buck S., Jacobs A., Van Montagu M. and Depicker A. 1997. T-DNA integration patterns in cotransformed plant cells suggest that T-DNA repeats originate from cointegration of separate T-DNAs. Plant J. 11: 15–29.

Eccleston V.S. and Ohlrogge J.B. 1998. Expression of lauroyl-acyl carrier protein thioesterase in *Brassica napus* seeds induces pathways for both fatty acid oxidation and biosynthesis and implies a set point for triacylglycerol accumulation. Plant Cell 10: 613–622.

Eskin M., McDonald B.E., Przybylski R., Malcolmson L.J., Scarth R., Mag T. et al. 1996. Canola oil. In: Hui Y.H. (ed.), Edible Oil and Fat Products: Oil and Oil Seeds. Vol. 2. John Wiley & Sons, Inc., New York, pp. 1–96.

Fagard M. and Vaucheret H. 2000. (Trans)gene silencing in plants: how many mechanisms. Annu. Rev. Plant Physiol. Plant Mol. Biol. 51: 167–194.

Ferrie A.M.R. and Keller W.A. 1995. Microspore culture for haploid plant production. In: Gamborg O.L. and Phillips G.C. (eds), Plant Cell, Tissue and Organ Culture: Fundamental Methods. Springer-Verlag, Berlin, pp. 155–164.

Gendloff E.H., Bowen B. and Buchholz W.G. 1990. Quantitation of chloramphenicol acetyl transferase in transgenic tobacco plants by ELISA and correlation with gene copy number. Plant Mol. Biol. 14: 575–583.

Hawkins D.J. and Kridl J.C. 1998. Characterization of acyl-ACP thioesterases of mangosteen (*Garcinia mangostana*) seed and high levels of stearate production in transgenic canola. Plant J. 13: 743–752.

Hitz W.D., Mauvis C.J., Ripp K.G. and Reiter R.J. 1995. The use of cloned rapeseed genes for the cytoplasmic fatty acid desaturases and the plastid acyl-ACP thioesterases to alter relative levels of polyunsaturated and saturated fatty acids in rapeseed oil. In: Proceedings of the Ninth International Rapeseed Congress: Rapeseed Today and Tomorrow. Cambridge, UK, pp. 470–472.

Hobbs S.L.A., Kpodar P. and DeLong C.M.O. 1990. The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. Plant Mol. Biol. 15: 851–864.

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.

Hougen F.W. and Bodo V. 1973. Extraction and methanolysis of oil from whole or crushed rapeseed for fatty acid analysis. J. Am. Oil Chem. Soc. 50: 230–234.

Jones A., Davies H.M. and Voelker T.A. 1995. Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant cell 7: 359–371.

Kidwell K.K. and Osborn T.C. 1992. Simple plant DNA isolation procedures. In: Beckmann J.S. and Osborn T.C. (eds), Plant Genomes: Methods for Genetic and Physical Mapping. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 1–13.

Knutzon D.S., Thompson G.A., Radke S.E., Johnson W.B., Knauf V.C. and Kridl J.C. 1992. Modification of *Brassica* seed oil by antisense expression of a stearoyl-acyl carrier protein desaturase gene. Proc. Natl. Acad. Sci. USA 89: 2624–2628.

Mannerlöf M., Tuvesson S., Steen P. and Tenning P. 1997. Transgenic sugar beet tolerant to glyphosate. Euphytica 94: 83–91.

Maqbool S.B. and Christou P. 1999. Multiple traits of agronomic importance in transgenic indica rice plants: analysis of transgene integration patterns, expression levels and stability. Mol. Breed. 5: 471–480.

McCabe M.S., Mohapatra U.B., Debnath S.C., Power J.B. and Davey M.R. 1999. Integration, expression and inheritance of two linked T-DNA marker genes in transgenic lettuce. Mol. Breed. 5: 329–344.

McVetty P.B.E. and Scarth R. 2002. Breeding for improved oil quality in *Brassica* oilseed species. J. Crop. Production 5: 345–369.

Mengiste T. and Paszkowski J. 1999. Prospects for the precise engineering of plant genomes by homologous recombination. Biol. Chem. 380: 749–758.

Ott R.L. 1993. An Introduction to Statistical Methods and Data Analysis. 4th edn. Duxbury Press, Belmont.

Rakow G. 1993. AC Excel summer rape. Can. J. Plant Sci. 73: 183–184.

Scarth R., Rimmer S.R. and McVetty P.B.E. 1995a. Apollo low linolenic summer rape. Can. J. Plant Sci. 75: 203–204.

Scarth R., McVetty P.B.E. and Rimmer S.R. 1995b. Mercury high erucic acid low glucosinolate summer rape. Can. J. Plant Sci. 75: 205–206.

Tang J. 2002. Effects of Genotype and Environment on the Level of the Target Fatty Acids of Acyl-ACP Thioesterase Transgenes in *Brassica napus* L. PhD Thesis.

Tenllado F. and Diaz Ruiz J.R. 1999. Complete resistance to pepper mild mottle tobamovirus mediated by viral replicase sequences partially depends on transgene homozygosity and is based on a gene silencing mechanism. Transgen Res. 8: 83–93.

van der Hoeven C., Dietz A. and Landsmann J. 1994. Variability of organ-specific gene expression in transgenic tobacco plants. Transgen Res. 3: 159–165.

Voelker T.A., Hayes T.R., Cranmer A.M., Turner J.C. and Davies H.M. 1996. Genetic engineering of a quantitative trait: metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. Plant J. 9: 229–241.

Voelker T.A., Jones A., Cranmer A.M., Davies H.M. and Knutzon D.S. 1997. Broad-range and binary-range acyl-acyl carrier protein thioesterases suggest an alternative mechanism for medium-chain production in seeds. Plant Physiol. 114: 669–677.

Voelker T.A., Worrell A.C., Anderson L., Bleibaum J., Fan C., Hawkins D.J. et al. 1992. Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257: 72–74.

Weiler K.S. and Wakimoto B.T. 1995. Heterochromatin and gene expression in *Drosophila*. Annu. Rev. Genet. 29: 577–605.

Wiberg E., Banas A. and Stymne S. 1997. Fatty acid distribution and lipid metabolism in developing seeds of laurate-producing rape (*Brassica napus* L.). Planta 203: 341–348.