JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Detecting and Quantifying the Adventitious Presence of Transgenic Seeds in Safflower, *Carthamus tinctorius* L.

JED CHRISTIANSON,[†,‡] MARC MCPHERSON,[§] DEBORAH TOPINKA,[§] LINDA HALL,[§] AND ALLEN G. GOOD*,[†]

Deparment of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada, Department of Agricultural, Food, and Nutritional Science, University of Alberta, Edmonton, Alberta T6G 2P5, Canada, and CSIRO Division of Plant Industry, Black Mountain, Canberra, ACT 2601, Australia

Safflower (*Carthamus tinctorius* L.) is currently being developed as a platform for the production of novel proteins. Methods for detecting and quantifying transgenic safflower are needed to ensure seed quality and to monitor for its adventitious presence. We developed and compared three methods of assaying for transgenic safflower presence in conventional seedlots: field bioassays, enzyme-linked immunosorbent assays (ELISA), and quantitative polymerase chain reaction (Q-PCR). Limits for reliable quantification for both ELISA and Q-PCR are ~0.1%, although levels at least as low as 0.02% can be detected by Q-PCR. Levels of quantification for the field bioassay are limited only by space and resources available. Multiple sampling methods to detect and quantify transgenic safflower presence at levels lower than 0.1% were used on field collected samples from a pollen outcrossing experiment to quantify the adventitious presence of transgenic safflower. Taking into account the potential utility and relative advantages or disadvantages of each detection method, it is recommended that the initial testing for the adventitious presence of transgenic seed be carried out using an antibody-based test if available and that Q-PCR-based assays to quantify transgenic proportion be used when it is necessary to identify specific transgenic constructs or if antibody-based assays are not readily available.

KEYWORDS: Safflower; Carthamus tinctorius L.; Q-PCR; ELISA; adventitious presence; transgenic; genetically modified organism

INTRODUCTION

The advent of transgenic plants for use in agriculture has resulted in an increased awareness of the potential for genes to move and spread through populations (*1*). Gene flow can occur both spatially and temporally, across geographical distances and down through generations. The spread of transgenes can lead to commingling of transgenic and nontransgenic seeds, leading to an adventitious presence in material intended for food or feed. For plants, the main vectors of gene flow are through seed and pollen dispersal. In this age of international commodities trades, seeds containing engineered traits can be shipped and spread throughout the world. Persistence in the seed bank of seeds lost during harvesting that contain engineered traits can lead to commingling with subsequent crops. To facilitate the coexistence of field production of crops for nontraditional uses such as the production of plant made pharmaceuticals (PMPs) and food or feed, it is critical to be able to detect and quantify commingling (*2*).

A number of recent publications has detailed the development and optimization of methods for assessing transgene presence in processed foodstuffs (*3–8*). Less attention has been devoted to assessing transgene presence in whole grain or bulk commodity samples (*9–11*) and even less to oilseed crops, which present a number of unique challenges in optimizing nucleic acid extraction and in storage or mixing of reference flour samples (*12*). Sampling for transgenic material in bulk seedlots carries with it a particular set of challenges. Unlike a processed food sample, the amount of material to be tested can be quite large, and potential transgenic contaminants may not be distributed randomly; therefore, an appropriate sampling strategy must be employed to achieve results that are representative of the bulk lot. Sampling strategies may involve means to thoroughly homogenize a seedlot prior to sampling or when this is not feasible may require taking spatially distributed samples to reduce the chances of nonhomogenous distribution of transgenic material to bias estimates of quantity. Recent publications have addressed theoretical concerns and statistical procedures for obtaining representative samples from potentially

* Corresponding author. E-mail: allen.good@ualberta.ca.
[†] Deparment of Biological Sciences, University of Alberta.
[‡] CSIRO Division of Plant Industry.
[§] Department of Agricultural, Food, and Nutritional Science, University of Alberta.

10.1021/jf800683g CCC: $40.75
Published on Web 06/26/2008

nonhomogenous populations very thoroughly (*13, 14*). In addition, procedures for analyzing transgene content in seed samples need to be optimized for the particular species to be studied since the amount of sample that can feasibly be processed will be dependent on seed size, and DNA extraction procedures will need to be tailored to seed composition.

Safflower (*Carthamus tinctorius*) is utilized as a platform for the production of PMPs (SemBioSys Genetics Inc.; www.sembiosys.com). Safflower is capable of outcrossing, with self-pollination rates reported to vary between 9.3 and 81.5% depending on genotype (*15*). This appears to be highly dependent on the distance between plants, as other studies have observed outcrossing rates of 10% for plants grown side by side decreasing to between 1.5−2.3% for plants grown 13.7 m away (*16*), 0.8% for plants grown 0.6 m apart, and between 0−0.12% for plants ranging between 3 and 47 m apart (*17*). In experiments using insect exclusion cages, safflower pollen was not observed to move beyond 1.2 m, suggesting that safflower pollination over distances may be insect-mediated. In support of insect-mediated pollination in safflower, bees have been observed foraging for pollen in safflower flowers (*18*). Pollen or seed movement from transgenic safflower intended for molecular farming to neighboring fields of the same crop grown for food or feed could lead to local scale commingling. Since PMP-containing crop varieties need to be strictly controlled and monitored, it is vital that appropriate methods for detecting and quantifying material from these varieties be developed.

This study focused on methods to detect PMP safflower seed. However, in this study, we chose to track the phosphinothricin acetyltransferase (PAT) gene based on a number of careful considerations. The PAT protein is expected to be expressed in most tissues at consistent levels and be expressed as a soluble protein in the cytoplasm, making it easier to detect using antibody-based methods. All PMP constructs for expression in safflower, which are currently under development, will use PAT as a selectable marker; however, the specific gene product in each construct may differ, as may the tissue and/or developmental expression of the PMP transgene. Thus, tracking PAT would be a logical first step in any monitoring program, as it will be present in all SemBioSys constructs. Should it be necessary, PCR primers and/or PMP specific antibody-based methods specific to each individual construct and insertion event can then be used for further analysis.

We assessed two methods, an ELISA-based assay and quantitative PCR, for detecting and quantifying the presence of transgenic safflower seeds in bulk seedlots. Estimates of transgene frequency derived by these methods were compared to those derived from a field-based bioassay, using a set of safflower seed bulks harvested from a field trial established to determine the frequency of outcrossing in safflower.

MATERIALS AND METHODS

Plant Material. Seeds of safflower (*C. tinctorius*) variety S-317, homozygous for a transgene construct that contained the PAT gene expressed under the control of the CaMV35S promoter as a selectable marker and seeds of nontransgenic S-317 were obtained from SemBioSys Genetics Inc. (Calgary, AB, Canada). These seedlots were tested for purity and to ensure that the transgene was homozygous, by progeny screening in a greenhouse, prior to being used in any experiments. Segregation analysis and sequencing of the insertion site indicated that the transgenic line used in these experiments contained a single, homozygous insert (S. Zaplachinski, personal communication).

Samples with unknown proportions of transgenic seed content were obtained from a field trial that was grown in 2002 in the municipality of El Bosque, in the province of Santiago, Chile to determine outcrossing frequency. A transgenic source plot (10 m × 11.2 m) was surrounded by eight blocks of nontransgenic safflower, 2.8 m wide, planted radially to form spokes extending from the transgenic source. Seeds were harvested from 16 plots (1.2 m × 1.2 m) spaced equidistant along the block (spoke) to form the bulked seed samples used for analysis.

Table 1. Composition of Seed Pools with Known Transgenic Content for Validation of Q-PCR and ELISA Quantification Methodologies

proportions[a]	transgenic seeds	nontransgenic seeds	sample size[b]	total seed sampled
		No Sampling		
1 in 10	30	270	300	all
1 in 100	3	297	300	all
1 in 1000	3	2997	3000	all
1 in 5000	3	14997	15000	all
		Sampling		
1 in 250	10	2490	150	750
1 in 500	10	4990	300	1500
1 in 1500	10	14990	1000	5000
1 in 2500	10	24990	1250	6250
1 in 5000	12	59988	3500	17500

[a] Proportion of transgenic to nontransgenic seeds. [b] For sampling methodology testing, either the whole seed pool or a sample of the pool was used for quantification.

Preparation of Seed Samples. To validate the ELISA and Q-PCR methods, seedlots containing known proportions of transgenic to nontransgenic seeds were prepared in the following manner. For the ELISA and initial quantitative real-time PCR validations, pools of 15000, 3000, and two pools of 300 nontransgenic seeds were spiked, respectively, with 3, 3, 3, and 30 transgenic seeds to yield pools with transgenic seed proportions of 1:5000, 1:1000, 1:100, and 1:10. For these seedlots, the entire pool was ground, and DNA was extracted from the resulting homogenate. A second set of pools was created and assayed via quantitative real-time PCR to assess the impact of the sampling procedure described next. These pools contained nontransgenic seeds spiked with transgenic seeds in various proportions as described in **Table 1**, which had been mixed to randomly distribute transgenic seeds throughout the pool. However, rather than grind the entire sample, multiple samples were collected from each pool, in an attempt to mimic a repeated sampling protocol whereby a number of subsamples are taken from a large bulk lot (e.g., a grain truck) to estimate the frequency of adventitious presence (*14*). A third set of samples was evaluated for the presence of the transgene. These samples were taken from seed harvested from individual plots from the outcrossing trial and were counted mechanically. Pool sizes were selected based on the frequency of transgenic plants observed in the field screen so that only some of the pools were expected to contain transgenic seeds. For example, if the expected transgenic frequency was 0.001, pools of 1000 seeds would be counted.

Seeds were ground in 1 vol of phosphate buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4) until a smooth homogenate was achieved. Grinding was accomplished using either a Quik Stik immersion blender (Waring, Torington, CT), a laboratory blender (Waring, Torington, CT), or a household electric coffee grinder depending on the sample size. Aliquots from this homogenate were taken for ELISA analysis and DNA extraction.

Field Screening for Glufosinate Resistance and Confirmation of PAT Gene Presence in Survivors. For the field bioassay, seeds were planted in 2004 in Edmonton, AB. Seeds were planted in six row plots that were 7 m long with 1 m alleys between plots. Control plots consisting of transgenic or nontransgenic seed were included at random positions within the trial. Emergent plants were counted, and young seedlings were subsequently subjected to two rounds of spraying with glufosinate herbicide (800 g ai ha^{-1} with a volume of 200 L ha^{-1}) with the first spray occurring when plants reached the 4−6 leaf stage and the second spray occurring 1 week after the first. The surviving plants were counted 1 week after the second glufosinate application, and a simple fraction of transgenic plants was determined. Leaf tissue samples from 302 herbicide survivor plants were harvested, frozen,

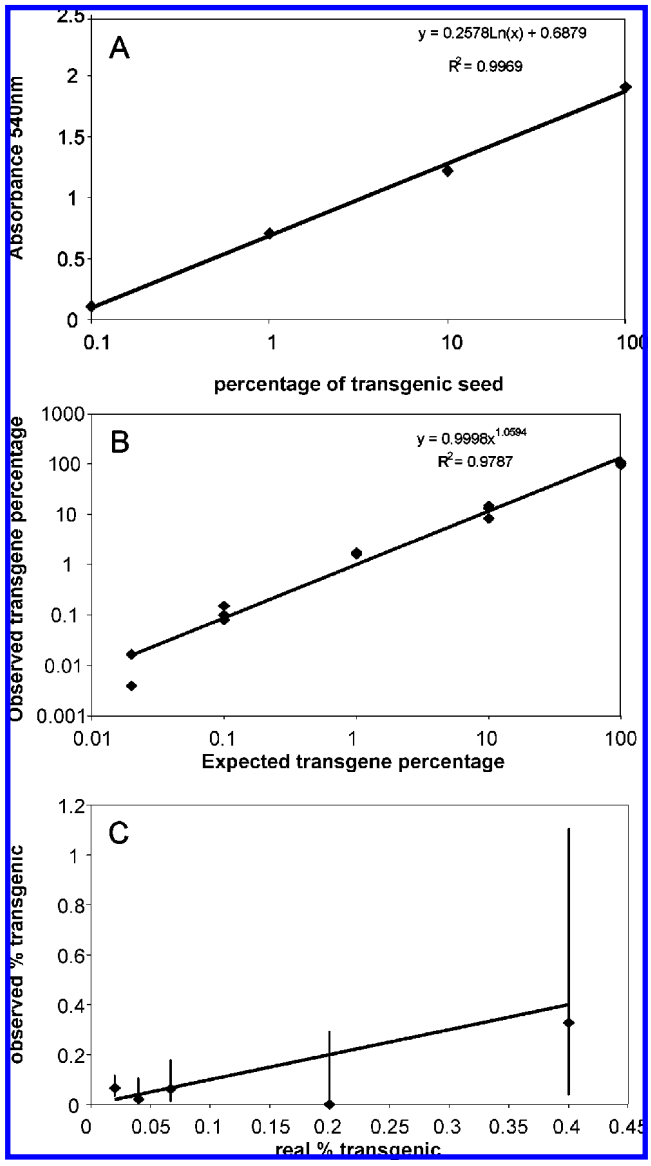

Figure 1. Validation of ELISA and Q-PCR methods for quantifying the presence of transgenic seeds. (**A**) Standard curve for quantification of transgenic content in seed samples with known amounts of transgenic seeds using the ELISA protein assay. Absorbance values were plotted against the percentage of transgenic material in the seed samples (log scale). (**B**) Plot of percentage of transgenic material in seed samples observed via Q-PCR against the known transgenic content of the seed samples assayed. Results are shown for three replicate extractions from each seed sample and were obtained using the PAT2.1 and PAT2.2 primers. Both *x* and *y* axes are plotted on a log scale. (**C**) Plot of percentage of transgenic material in seed pools observed via Q-PCR-based estimation of transgenic content in multiple samples against the known transgenic content of the seed pools assayed. Error bars indicate the 95% confidence interval for each pool estimate (following Zar (*20*)). The solid line is included for reference and shows results expected if the observed results equal the actual transgenic content of the tested pools. Results shown were obtained using PAT2.1 and PAT2.2 primers.

and then tested using either commercially available antibody-based test strips (Strategic Diagnostics Inc., Newark DE) and PCR or PCR alone to confirm the presence of the PAT gene (**Figure 3**). Ninety-nine of these tissue samples were subjected to both a commercially available antibody strip test (Strategic Diagnostics), which detects the phosphinothricin acetyltransferase protein, and a traditional end-point PCR with two sets of primers, one detecting an arbitrarily chosen region of the safflower genome (JCH1 CACACTAAGCCACTCCAACC; JCH4 TTGACAACTCCAATCCCTGC) and one that detected the PAT gene (JCH5 GATCTGGGTAACTGGTCTAACTGG; JCH6 GTTGCAAGATAGATACCCTTGGTT). A further 203 tissue samples were assayed with the end-point PCR alone. The PCR reactions were performed in 20 μL volumes with 40 ng of template DNA, 0.5 mM primers, 2 μL of 10X PCR buffer, 3.0 mM $MgCl_2$, 0.25 mM dNTPs, and 2 units of the *Taq* polymerase. The thermocycling protocol was 95 °C for 10 min; followed by 35 cycles of 95 °C for 20 s, 59 °C for 30 s, and 72 °C for 45 s; and a final elongation step of 72 °C for 5 min. Samples were then separated on 1.0% agarose gels and scored for the presence of the control and PAT gene amplification products.

ELISA Quantification of PAT Protein Levels. Samples of seed homogenates were centrifuged in a benchtop microfuge at 16500*g* for 15 min, which separated the homogenate into a solid pellet, an aqueous layer, and a top layer comprised of seed lipids. The middle aqueous layer was collected, frozen at −20 °C, and later used directly in commercially available ELISA plate assays (Envirologix, Portland, ME). Quantification of the PAT protein was accomplished by comparing the absorbance values of test samples in the ELISA plate assay to a standard curve of absorbance values generated from seed homogenates with a known proportion of transgenic to nontransgenic seeds. In the case of the seed samples obtained from the outcrossing trial conducted in Chile, observed proportions of transgenic material were multiplied by a factor of 2 before comparison with the transgenic plant frequencies observed from field screening. This correction was deemed necessary since the PAT protein in the hemizygous seeds obtained from the outcrossing trial would be present at levels half of those in homozygous seeds, assuming that gene expression is dose-dependent.

DNA Extraction and Purification. Samples of seed homogenate were centrifuged as stated previously, and 0.5 mL of the resulting aqueous layer was collected and mixed with 0.5 mL of 3% CTAB (3% w/v cetyltrimethylammonium bromide, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris (pH 8.0), and 0.4% β-mercaptoethanol). Samples were incubated at 60 °C for 1 h, then extracted with 1 vol of chloroform and precipitated with 0.6 vol of isopropyl alcohol. DNA samples were then resuspended and purified using GENECLEAN Turbo 96 kits (Q-Biogene, Irvine, CA) following the manufacturer's instructions. Purified DNA samples were quantified fluorometrically using picogreen dye (Invitrogen, Carlsbad, CA). All DNA samples were diluted to a concentration of 10 ng/μL.

Q-PCR-Based Quantitation of PAT DNA Proportion. The ΔΔCT (cycle threshold) relative quantification method was used to determine the proportion of transgenic DNA present in each template. The CT value of the sample gene, in this case PAT, was compared to the CT value of a reference gene, in this case glycerol-3-phosphate acyl transferase (GPAT, Genbank accession no. L33841), in a multiplex reaction. The resulting difference between the CT values for these sequences in each reaction was then compared to the difference in CT values observed in a calibrator sample, in this case safflower genomic DNA from transgenic seeds, and expressed as a fold change difference. Primers and Taqman MGB probes specific to each gene are as follows. For the PAT gene, PAT1.1 primer, GTTGAGGGTGTTGTGGCTGG; PAT1.2 primer, TCCAATCGTAAGCGTTCCTAGC; and PAT probe, 6FAM-CTTACGCTGGGCCC or PAT2.1 primer, TGAGGGTGTTGTGGCTGGTA; PAT2.2 primer, GCCTATGTGACACGTAAACAGTACTCT; and PAT probe, 6FAM-CTTACGCTGGGCCC were synthesized. The sequence for the *C. tinctorius* GPAT (Genbank accession no. L33841) gene was obtained from Genbank and a TaqMan MGB probe, and set of primers specific to this sequence (GPAT1 primer, CACCAAGTGTCTCGTTGTCACA; GPAT2 primer, CACCTCCTCTAAACCAAGAGACTAATC; and GPAT probe, 6VIC-CCCTTGCAGTTTCT) was synthesized for use as a reference. The GPAT primers were compared to publicly available plant sequences using the BLAST sequence comparison tool (*19*). No significant matches to other plant sequences were found, implying that these primers are highly specific for safflower.

A total of 100 ng of template DNA was used in Q-PCR reactions with 12.5 μL of 2X TaqMan Fast Universal PCR Master Mix (Applied Biosystems, Foster City, CA) and 150 nM each of PAT and GPAT Taqman MGB probes and 300 nM each of the PAT and GPAT primer

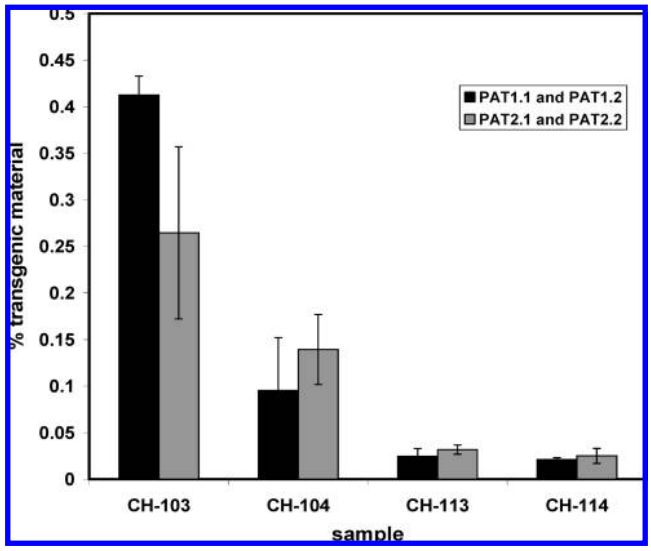

Figure 2. Comparison of two different PAT gene specific primer sets for estimation of transgene content via Q-PCR. Four plots were assayed via Q-PCR with both PAT1.1/PAT1.2 and PAT2.1/PAT2.2 primer sets. Error bars represent 1 SD.

sets. Q-PCR was carried out on an ABI 7500 fast real-time PCR system (Applied Biosystems, Foster City, CA) under the following conditions: one 10 s step at 95 °C and 50 cycles of 95 °C for 10 s and 60 °C for 30 s. Fluorescence data were collected during each 60 °C annealing/extension phase. For each reaction, the difference in the CT value for the PAT gene and the GPAT gene was compared to the difference observed in control samples containing only transgenic safflower genomic DNA.

In the case of seed samples obtained from the outcrossing trial conducted in Chile, observed proportions of transgenic material were multiplied by a factor of 2 before comparison with the transgenic plant frequencies observed from field screening. This correction accounts for the fact that the Q-PCR method will only detect one transgene per hemizygous seed genome, while the field screen counts the entire seed as transgenic.

Statistical Analyses. For each method, 95% confidence intervals were determined for the observed proportions of transgenic material assuming a binomial distribution as described in Zar (*20*). In those cases where the observed proportion was zero, an alternative method for calculating the upper 95% limit was used, as described by Zar (*20*).

RESULTS

Validation of Protein Content Quantification with Commercial ELISA Plates. Commercially available ELISA plate kits (Envirologix, Portland, ME) were used to detect and quantify PAT levels in samples prepared from seed pools varying in size and in their proportion of transgenic to nontransgenic seeds (**Table 1**). Spectrophotometric results of the ELISA showed a logarithmic relationship between absorbance and proportion of transgenic seed with a good linear fit ($R^2 = 0.997$; **Figure 1A**). Samples with a transgenic content below 1:1000 did not give a detectable signal. These experiments were initially performed without quantifying the amount of protein in the assay; however, similar results were obtained where the protein content of the sample was determined using the Bradford method (data not shown (*21*)).

Validation of Quantitative Real-Time PCR. Taqman MGB probes and primers were designed to amplify sequences specific to the PAT transgene and to the GPAT gene present in safflower. Primer efficiencies were determined from a series of DNA dilutions spanning 4 orders of magnitude by graphing the CT values for each dilution against the log of DNA concentration

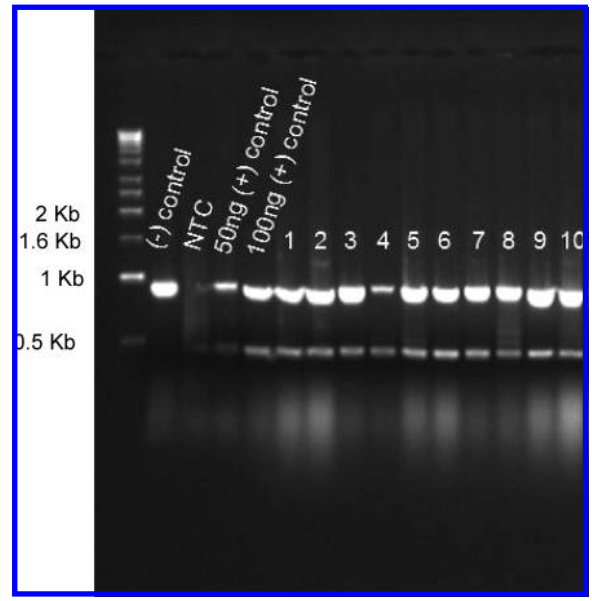

Figure 3. PCR confirmation of the presence of the PAT gene in plants that survived the field bioassay. DNA extracted from leaf tissue of plants surviving the field bioassay was used as a template in a multiplex PCR reaction with primer sets JCH1/JCH4 and JCH5/JCH6, which amplify an arbitrary ~900 bp safflower genomic sequence and a 424 bp sequence of the PAT gene, respectively. DNA from a nontransgenic safflower plant was used as a negative control.

and calculating the slope of the resulting line. PAT1.1 and PAT1.2 primers were found to have an amplification efficiency of 99.2%, PAT2.1 and PAT2.2 primers had an amplification efficiency of 96.4%, and GPAT primers had an amplification efficiency of 98.9%. A ΔΔCT comparison method (*22*) was used to determine the proportion of transgenic templates in a DNA sample. To validate the quantitative real-time PCR methods, a set of nontransgenic seed samples spiked with known quantities of transgenic seed was used as a control. Triplicated DNA extractions from these spiked seed samples were then processed and assayed via Q-PCR to compare the measured transgenic proportions to the known proportions of transgenic to nontransgenic seeds (**Figure 1B**). The correlation between the known transgenic content of the seed samples and the PCR-based quantification of transgenic content was high ($R^2 = 0.979$), indicating that proportionality was maintained throughout the seed homogenization and DNA extraction processes. Seed samples with transgenic contents of 0.02% showed the greatest amount of variability between replicate DNA extractions, with one of the three extractions measuring a transgenic proportion 5-fold lower than expected. The genome size of safflower was estimated in a separate real-time PCR-based experiment, following the procedure outlined by Wilhelm et al. (*23*). Identifying a true single copy gene is one of the challenges in estimating genome size in species such as safflower that are not well-characterized genetically. Preliminary Southern blots using a number of cDNA probes indicated significant levels of gene duplication and complexity in this species (R. Mayerhofer, unpublished results). However, it was determined that the transgenic line used in this study was homozygous for a single copy gene based on Southern blots, PCR amplification of the transgene, and sequencing of the border sequences (S. Zaplachinski, SemBioSys Genetics, personal communication). Amplification profiles generated from known quantities of a PAT derived 392 bp PCR product spanning the transgene amplicon produced by the PAT real-time PCR primers described here were used to construct a standard curve, against which the amplification of

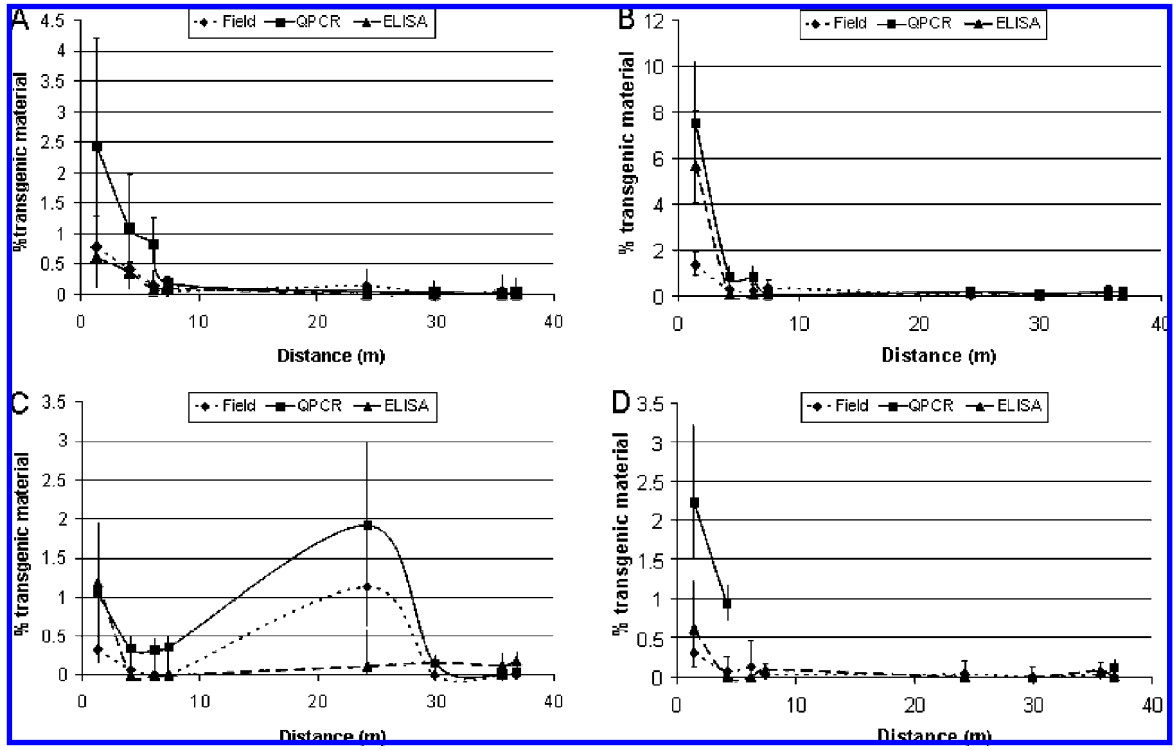

Figure 4. Proportion of transgenic material detected by different methods in seed samples taken at increasing distances from a transgenic pollen source. The proportion of transgenic seeds in samples from an outcrossing field trial designed to measure pollen flow was determined using a field-based bioassay method (dotted line), ELISA-based determination of transgenic protein proportion (dashed line), or Q-PCR-based determination of transgenic DNA content (solid line). Proportions determined from (**A**) plots at various distances north of the transgenic source, (**B**) plots at various distances east of the transgenic source, (**C**) plots at various distances south of the transgenic source, and (**D**) plots at various distances southeast of the transgenic source. Error bars represent the upper and lower 95% confidence intervals for a particular data point, following the method of Zar (*20*).

a known amount of transgenic DNA was compared. Dividing the weight in nanograms of transgenic DNA used in the experiment by the number of transgene copies determined from the standard curve yielded a haploid genome weight of ~2.7 pg. This estimated value is the same value as that determined recently by flow cytometry (*24*). Using this value, the number of genome copies contained in a 100 ng sample as used in real-time PCR amplifications was 37 037. Following Hubner et al. (*25*), the theoretical limit of reliable quantification for safflower using 100 ng of genomic DNA was 0.09% (1:1058). The 0.02% transgenic sample, which is below this limit of quantification, had a high observed variability (**Figure 1B**), but the low measured transgenic proportion in one replicate likely was due to a DNA template sampling error.

We next evaluated a method using measurements of multiple samples derived from a common seed pool to detect and quantify transgenic presence below the limit of quantification for a single sample in safflower. In this experiment, seedlots containing transgenic material in proportions ranging from 1 in 250 to 1 in 5000 were made (**Table 1**) and mixed to distribute transgenic seeds randomly throughout the pool. To mimic situations where the transgenic content of a seed pool may be quite low, and any single seed sample of workable size may have a low probability of containing any transgenic seeds, the number of samples to be drawn was set to five, and the sample size was determined such that each sample had an approximate probability of 60% of containing one transgenic seed. Over five samples, the probability of drawing at least one transgenic seed ranged from 95 to 99%. The number of pools and sample sizes to be used, and hence the number of seeds to be tested overall, and the sensitivity of the Q-PCR estimates was determined before we established the genome size of safflower and thus do not reflect an optimal sampling strategy (see Discussion). Transgenic estimates from each sample were combined to arrive at an estimate for the entire seed pool. Because of sampling error, the estimates of transgenic proportion derived by this method were not as accurate as those derived from the previous experiment where the whole seed pool was used for analysis without sampling. For pools containing 0.4% (1 in 250), 0.066% (1 in 1500), 0.04% (1 in 2500), and 0.02% (1 in 5000), the Q-PCR estimates were 81, 91, 52, and 329%, respectively, of the true transgenic proportions for the seedlots sampled. For each of these estimates, except the 0.02% transgenic seed pool, the true transgenic content value fell within the 95% confidence range, based on the method for determining confidence limits for population proportions of Zar (*20*) (**Figure 1C**). The Q-PCR derived estimate for the 0.2% sample was 0. Even so, the true value of 0.2% fell within the 95% upper limit for this estimate. The estimated transgenic percentage for the 0.02% transgenic seed pool resulted from a very high estimate of transgenic content in one of the five samples used. It is possible that the unexpectedly high estimate of transgenic content for this pool is due to inadequate seed mixing of the pool, prior to seed sampling for analysis.

During the course of our initial investigations, we observed an amplification curve from the PAT gene specific probe very late in the cycling program in no-template control reactions (NTC). Template contamination was ruled out as a source of this signal, and redesigned PAT gene specific primers (PAT2.1 and PAT2.2) did not result in this amplification artifact when used. The late amplification artifact generated by the PAT1.1 and PAT1.2 primers could be differentiated from genuine amplifications that gave higher CT values, and similar results were observed when seed samples with a known transgenic

content were assayed with either the PAT1.1/PAT1.2 primer set or the PAT2.1/PAT2.2 set (**Figure 2**). Consequently, data generated using either primer set are presented here.

Comparison of Methods for Determining Transgene Presence and Quantity. To assess the utility of each method (bioassay, ELISA, and Q-PCR) for detecting and quantifying transgene presence in authentic situations, seed samples harvested from a field trial designed to establish outcrossing distance in safflower were assayed using each method. The samples used represent seed sets at 16 plots between 1.4 and 36.8 m from a transgenic source. These plots were expected to contain a range of transgenic seed proportions, and the transgenic proportion was expected to decrease with distance. Transgenic seed content for the samples was first assessed using a field bioassay. All 302 plants tested were confirmed to possess the PAT gene. Given the limitations of screening large numbers of samples in the field, only 2000−3000 seedlings representing each plot in the outcrossing trial could be screened, decreasing the sensitivity of this test. For the purposes of comparing these methods, five samples of varying numbers of seed were taken from the first four and last four plots from four of the eight blocks in the outcrossing trial and used for analysis of transgenic protein content by ELISA and transgenic DNA content by Q-PCR. The size of these samples was determined based on the results obtained from the field bioassay. Sample sizes were adjusted such that only some of the samples would be expected to contain a transgenic seed, to a maximum sample size of 2000 seeds. The use of multiple samples allowed us to test up to 10000 seeds from a single plot and to detect and quantify transgene presence at levels lower than the limits of quantification either of these methods could achieve on a single sample, although the impact of sampling error is magnified when the expected frequency of transgenic seeds is close to zero. Estimates of transgenic proportions and 95% confidence intervals in each plot, as determined by each method, are presented in **Table 2** and **Figure 4**. Transgenic proportions were consistently higher in observations made by Q-PCR than those made by either the field bioassay or ELISA, especially at higher frequencies. In many cases, the Q-PCR observations were outside of the 95% confidence interval of the bioassay observations, indicating a substantial difference between the estimates of transgenic proportions made by each method. Despite this consistent difference in estimates, plots of estimated transgenic proportion at increasing distances from the transgenic source exhibited a very similar shape for the Q-PCR and bioassay methods, and there was a strong positive correlation between the estimates by these two methods ($R^2 = 0.66$). Transgenic proportions were below the ELISA detection threshold in many of the samples tested. However, the estimates reached by this method showed high correlation with the Q-PCR results ($R^2 = 0.86$; **Table 2**), although this correlation is heavily influenced by the high proportion of transgene content in plot CH-301 estimated with both of these methods. When this point is omitted from the analysis, the correlation between ELISA- and Q-PCR-based estimates falls to $R^2 = 0.24$. Estimates reached by the field screen were linearly correlated with those reached by ELISA ($R^2 = 0.46$).

Table 2. Estimates of Transgenic Frequency in Seed Samples from a Pollen Flow Field Trial As Derived by Three Different Methodologies

		Q-PCR		bioassay		ELISA	
plot	distance from source (m)	estimate[a]	upper limit[b]	estimate[a]	upper limit[b]	estimate[a]	upper limit[b]
CH-101	1.4	2.44	4.22	0.781	1.36	0.603	1.28
CH-102	4.2	1.09	1.97	0.412	0.809	0.361	0.718
CH-103	6.2	0.825	1.26	0.154	0.393	0.076	0.243
CH-104	7.4	0.191	0.298	0.057	0.205	0.105	0.122
CH-113	24.2	0.049	0.118	0.137	0.401	0	0.044
CH-114	29.9	0.042	0.107	0.000	0.227	0	0.044
CH-115	35.6	0.013	0.065	0.056	0.311	0	0.044
CH-116	36.8	0.054	0.124	0.050	0.276	0	0.044
CH-301	1.4	7.50	10.2	1.33	1.89	5.65	4.69
CH-302	4.2	0.826	1.27	0.274	0.637	0.055	0.211
CH-303	6.2	0.817	1.26	0.213	0.464	0.027	0.175
CH-304	7.4	0.067	0.520	0.335	0.690	0.013	0.437
CH-313	24.2	0.175	0.278	0.000	0.147	0.113	0.126
CH-314	29.9	0.075	0.153	0.000	0.202	0	0.044
CH-315	35.6	0.163	0.264	0.000	0.443	0	0.044
CH-316	36.8	0.171	0.274	0.000	0.386	0	0.044
CH-501	1.4	1.05	1.78	0.328	0.603	1.18	1.20
CH-502	4.2	0.341	0.502	0.069	0.247	0	0.058
CH-503	6.2	0.324	0.457	0.000	0.270	0	0.044
CH-504	7.4	0.359	0.498	0.000	0.319	0	0.044
CH-513	24.2	1.92	2.99	1.13	1.89	0.111	0.483
CH-514	29.9	0.152	0.250	0.000	0.126	0.157	0.157
CH-515	35.6	0.006	0.049	0.000	0.267	0.118	0.132
CH-516	36.8	0.035	0.095	0.000	0.172	0.180	0.171
CH-601	1.4	2.23	3.21	0.308	0.634	0.609	0.811
CH-602	4.2	0.925	1.17	0.072	0.259	0	0.058
CH-603	6.2			0.128	0.461	0	0.088
CH-604	7.4			0.031	0.173	0.091	0.111
CH-613	24.2			0.035	0.194	0	0.044
CH-614	29.9			0.000	0.122	0	0.044
CH-615	35.6			0.033	0.185	0.077	0.103
CH-616	36.8	0.123	0.215	0.000	0.143	0	0.044

[a] Frequency of transgenic seed in sample expressed as a percentage. [b] Upper 95% confidence limit for each estimate based on the method for determining confidence limits for population proportions (*20*).

DISCUSSION

General. Three different methods for detecting and quantifying transgenic seeds in bulk seedlots were used on seed samples obtained from a field trial designed to assess pollen flow in transgenic safflower. These methods each showed rapidly decreasing transgenic proportions with increasing distance from the transgenic source plot. By all methods used, levels of observed transgenic content quickly diminished to levels below 0.2% within 6.2 m of the transgenic pollen source. The sole exception was one plot at 24.2 m from the transgenic source (**Figure 4C**). High levels of transgenic seed were detected in this plot by both the bioassay and the Q-PCR methods. It is not clear as to why the levels of transgenic seed in this one plot should be so high, nor is it clear as to why only the Q-PCR and bioassay methods should detect it.

Comparison of Three Methods of Quantification. On the basis of results of antibody test strips and end-point PCR to confirm transgene presence in bioassay survivors, the false positive rate of the bioassay was lower than 1:300. However, we were unable to assess as to what the false negative rate was, as a positive plant that was killed by overtreatment of a herbicide, or some other factor, would be judged as being negative. Another disadvantage of this method included being limited to screening smaller numbers of samples in the field (decreased sensitivity) and being limited by the seasonal and resource requirements for a controlled field trial. The Q-PCR-based method consistently gave estimates 2−3-fold higher than the field bioassay.

The estimates of transgenic proportion determined by measuring the PAT protein with ELISA were in agreement with the Q-PCR estimates and with the bioassay estimates; however, the correlation between these methods was not as strong as anticipated. Because of the low amounts of transgenic seed in all of these samples and the relatively small numbers of seeds that could be screened by any of the methods used here, the sampling error is very large relative to the determined estimates. Confidence levels based on sampling probabilities for a binomial

population can be placed around the estimated values for each distance point. However, because the frequency of transgenic seed decreases very quickly as one moves away from the transgenic pollen source, the data do not allow trend analysis to determine as to whether the decay trends differ between the different sampling protocols. In many cases, the estimates arrived at by Q-PCR are outside the 95% confidence limits for the field-based methods, while the ELISA estimates varied, sometimes being within the Q-PCR 95% confidence intervals and sometimes within the bioassay 95% confidence intervals. It is unclear as to why the transgenic estimates arrived at by Q-PCR-based methods should consistently be higher than those arrived at by field bioassays. It is also puzzling as to why the ELISA-based transgenic proportion estimates do not consistently agree with the estimates arrived at by Q-PCR, given that both methods measure the same sample pools. It could be that there is either a bias in the Q-PCR method or a bias in the field-based method. Results from the Q-PCR validation experiments estimating transgenic proportion in spiked seedlots with or without subsampling indicated that Q-PCR measurements were quite accurate when measuring entire seedlots without sampling and accurate to within the 95% confidence interval expected for a sampling error in a binomial population.

Costs, detection sensitivities, execution times, and skill requirements differed for each method. While the ELISA-based methods were quickest to perform, were the least costly, and required no specialized skill, they were less sensitive and had the highest amount of variability as seen in the validation experiments and were least in agreement with the other two methods when applied to field samples. Sensitivity of the ELISA method varies depending on the protein to be detected and the specific antibodies used for detection, the manner in which the protein extract is prepared, and the level of expression in the particular tissue analyzed. We were able to reliably detect the PAT protein in ground seed samples at a proportion of 0.1%. Previous validations of ELISA methods for the detection of the CP4 EPSPS protein in Roundup Ready soybean could reliably detect this protein in homogenized samples at a proportion of 0.3% (*26*). Quantification of transgenic content in unknown samples using the ELISA method requires appropriate reference material to generate a standard curve. This requirement will affect the cost of using ELISA-based methods as preparation, validation, and maintenance of reference materials will be an ongoing process.

The Q-PCR method used here can be quite accurate when it is used to measure samples with known proportions of transgenic seed but is technically demanding and expensive. The method is quite sensitive and was able to detect the PAT transgene at proportions of 0.02% (~7.4 transgenic genome copies in a 100 ng sample of template). However, detection at this level is very difficult; aliquots of the PCR template may fail to contain any transgenic sequence, and PCR becomes more error prone at very low levels of the target sequence. Following Hubner et al. (*25*), the effective limit of quantification is only 1:1058 for a 100 ng genomic DNA sample, meaning that accurate quantification of transgene proportions lower than 0.1% require multiple samples that may or may not contain any transgenic material, introducing an added layer of sampling error. We attempted to validate this multiple sampling strategy and found that our real percentages of transgenic material were within 95% confidence limits of our estimates for all transgenic proportions from 0.04% and up. Our estimate for a seed pool with a known transgenic content of 0.02% was much higher than expected, and we suspect that this was the result of inadequate homogenization of the seed sample. Notwithstanding this error, limiting ourselves to using five seed samples of 3500 seeds would likely have yielded inaccurate results anyway, as the expected frequency of transgenic seeds in this sample size is below the limit of quantification for our Q-PCR method. To achieve a reliable estimate of transgenic proportion for this sample pool, a better strategy would have been to have used 16 samples of 1100 seeds.

Many other published quantitative PCR-based methodologies for quantification of transgene content utilized series of reference standards containing varying proportions of transgenic material (*3, 5, 7, 9*). Here, we used a Taqman Q-PCR assay where primers and probes for the PAT gene and reference GPAT gene were multiplexed in a single reaction, and the relative quantifications were determined using the ΔΔCT method described by Pfaffl (*22*). This methodology has the advantages of carrying out target and reference gene reactions in a single tube, increasing throughput, and not requiring a set of standards, relying instead on one reference sample (in this case 100% transgenic DNA). This eliminates the need for preparing and validating reference materials of known transgene content.

Two sets of Taqman primers were designed to amplify the PAT gene, as the first set of Taqman primers consistently showed late amplification in the no-template control reactions. Although the Taqman system is specifically designed to eliminate nonspecific amplification, exhaustive testing could not locate a source of template contamination in these no-template control reactions, and the problem persisted even after replacement of all reagents and plasticware used. The use of redesigned PAT gene primers eliminated the observed late amplification in the no-template controls. The presence of this apparent product in no-template controls highlights the need to ensure that the primers and conditions for amplification are carefully determined, to prevent false positives. In contrast, the detection threshold of a field assay is limited only by the space available. Field bioassays are costly, require a large amount of skilled field labor, and require the most time. Field-based assays also require an easily screenable trait such as herbicide resistance. In addition, any field assays involving plant varieties with novel traits, such as plant-made pharmaceutical-bearing safflower, will be subject to regulatory approval and monitoring. On the basis of these data, a logical monitoring protocol for the adventitious presence of transgenic safflower first would involve screening seedlots with an antibody-specific test (providing one is available), due to the simplicity of the test and the low probability of a false positive. Second, a DNA-based Q-PCR test could be envisioned since this would permit determination of the specific transgenic construct. Finally, a field-based bioassay would be the least useful, due to cost, the potential requirement for additional monitoring of the site after initial testing, and the seasonality of the testing program.

ACKNOWLEDGMENT

We thank SemBioSys Genetics Inc. for all of our specific seed stocks.

LITERATURE CITED

(1) Hills, M. J.; Hall, L.; Arnison, P.; Good, A. G. Genetic use restriction technologies (GURTs): Strategies to impede transgene movement. *Trends Plant Sci.* **2007**, *12*, 177–183.

(2) Council for Agricultural Science and Technology (CAST). *Adventitious Presence: Inadvertent Commingling and Coexistence Among Farming Methods; CAST Commentary QTA 2005-1*; CAST: Ames, IA, 2006.

(3) Berdal, K. G.; Holst-Jensen, A. Roundup Ready soybean event-specific real-time quantitative PCR assay and estimation of the

practical detection and quantification limits in GMO analyses. *Eur. Food Res. Technol.* **2001**, *213*, 432–438.
(4) Birch, L.; Archard, C. L.; Parkes, H. C.; McDowell, D. G. Evaluation of Labchip technology for GMO analysis in food. *Food Control* **2001**, *12*, 535–540.
(5) Vaïtilingom, M.; Pijnenburg, H.; Gendre, F.; Brignon, P. Real-time quantitative PCR detection of genetically modified maximizer maize and Roundup Ready soybean in some representative foods. *J. Agric. Food Chem.* **1999**, *47*, 5261–5266.
(6) Forte, V. T.; Di Pinto, A.; Martino, C.; Tantillo, G. M.; Grasso, G.; Schena, F. P. A general multiplex-PCR assay for the general detection of genetically modified soya and maize. *Food Control* **2005**, *16*, 535–539.
(7) Rott, M. E.; Lawrence, T. S.; Wall, E. M.; Green, M. J. Detection and quantification of Roundup Ready soy in foods by conventional and real-time polymerase chain reaction. *J. Agric. Food Chem.* **2004**, *52*, 5223–5232.
(8) Hernández, M.; Río, A.; Esteve, T.; Prat, S.; Pla, M. A rapeseed-specific gene, acetyl-CoA carboxylase, can be used as a reference for qualitative and real-time quantitative PCR detection of transgenes from mixed food samples. *J. Agric. Food Chem.* **2001**, *49*, 3622–3627.
(9) Alary, R.; Serin, A.; Maury, D.; Jouira, H. B.; Sirven, J.; Gautier, M.; Joudrier, P. Comparison of simplex and duplex real-time PCR for the quantification of GMO in maize and soybean. *Food Control* **2002**, *13*, 235–244.
(10) Permingeat, H. R.; Reggiardo, M. I.; Vallejos, R. H. Detection and quantification of transgenes in grains by multiplex and real-time PCR. *J. Agric. Food Chem.* **2002**, *50*, 4431–4436.
(11) Alexander, T. W.; Reuter, T.; McAllister, T. A. Qualitative and quantitative polymerase chain reaction assays for an alfalfa (*Medicago sativa*) specific reference gene to use in monitoring transgenic cultivars. *J. Agric. Food Chem.* **2007**, *55*, 2918–2922.
(12) Middleton, A. T.; Stone, E. M. DNA extraction from bulked samples of canola seed and the use of multiplex PCR for detection of adventitious contamination with genetically modified seed. *Seed Sci. Technol.* **2003**, *31*, 487–495.
(13) Remund, K. M.; Dixon, D. A.; Wright, D. L.; Holden, L. R. Statistical considerations in seed purity testing for transgenic traits. *Seed Sci. Res.* **2001**, *11*, 101–119.
(14) Laffont, J.; Remund, K. M.; Wright, D.; Simpson, R. D.; Grégoire, S. Testing for adventitious presence of transgenic material in conventional seed or grain lots using quantitative laboratory methods: Statistical procedures and their implementation. *Seed Sci. Res.* **2005**, *15*, 197–204.
(15) Claassen, C. E. Natural and controlled crossing in safflower, *Carthamus tinctorius* L. *Agron. J.* **1950**, *42*, 301–304.
(16) Kadam, B. S.; Patankar, V. K. Natural cross-pollination in safflower. *Indian J. Genet. Pl. Breed.* **1942**, *2*, 69–70.
(17) Deokar, A. B.; Patil, F. B. Vicinism in safflower. *J. Maharashtra Agric. Univ.* **1976**, *1*, 232–234.
(18) Howard, A.; Howard, G. L. C.; Khan, A. R. Studies in Indian oil seeds. *Memoirs Dept. Agric. India* **1915**, *7*, 237–272.
(19) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.
(20) Zar, J. H. *Biostatistical Analysis*, 4th ed.; Prentice-Hall, Inc.: Upper Saddle River, NJ, 1999; pp 527−528.
(21) Bradford, M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein−dye binding. *Anal. Biochem.* **1976**, *72*, 248–254.
(22) Pfaffl, M. W. A new mathematical model for relative quantification in real-time PCR. *Nucleic Acids Res.* **2001**, *29*, 2002–2007.
(23) Wilhelm, J.; Pingoud, A.; Hahn, M. Real-time PCR-based method for the estimation of genome sizes. *Nucleic Acids Res.* **2003**, *31*, 56.
(24) Garnatje, T.; Garcia, S.; Vilatersana, R.; Valles, J. Genome size variation in the genus *Carthamus* (Asteraceae, Cardueae): Systematic implications and additive changes during allopolyploidization. *Ann. Bot.* **2006**, *97*, 461–467.
(25) Hubner, P.; Waiblinger, H. U.; Pietsch, K.; Brodmann, P. Validation of PCR methods for quantitation of genetically modified plants in foods. *J. AOAC Int.* **2001**, *84*, 1855–1864.
(26) Lipp, M.; Anklam, E.; Stave, J. W. Validation of an immunoassay for detection and quantitation of a genetically modified soybean in food and food fractions using reference materials: Interlaboratory study. *J. AOAC Int.* **2000**, *83*, 919–927.

Received for review March 5, 2008. Revised manuscript received April 22, 2008. Accepted April 30, 2008. We thank SemBioSys Genetics Inc. for continued financial support. This research was funded by a Natural Sciences and Engineering Research Council of Canada, Industrial CRD grant to A.G.G.

JF800683G