Tree Physiology 29, 299–312
doi:10.1093/treephys/tpn028

Stability of transgenes in trees: expression of two reporter genes in poplar over three field seasons

JINGYI LI,[1,*] AMY M. BRUNNER,[3] RICHARD MEILAN[4] and STEVEN H. STRAUSS[1,2]

[1] *Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331-5752, USA*

[2] *Corresponding author (Steve.Strauss@oregonstate.edu)*

[3] *Department of Forestry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0324, USA*

[4] *Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907-2061, USA*

Received August 12, 2008; accepted October 22, 2008; published online December 23, 2008

Summary High stability of transgene expression is essential for functional genomics studies using transformation approaches and for application of genetic engineering to commercial forestry. We quantified expression of two reporter genes, green fluorescent protein (*GFP*) and the herbicide bialaphos resistance gene (*BAR*), in 2256 transgenic poplar trees derived from 404 primary events, and in 106 in vitro-redifferentiated subevents, over 3 years in the greenhouse and in the field. No gene silencing (complete breakdown of expression) was observed for *GFP* or *BAR* expression in any of the primary transgenic events during the course of the study. Transgenic cassettes were physically eliminated in four subevents (2.5%) derived from three different primary events during re-organogenesis. Transgene copy number was positively correlated with transgene expression level; however, a majority of transformants (85%) carried single-copy transgenes. About one-third of the events containing two-copy inserts had repeats formed at the same chromosomal position, with direct repeats being the main type observed (87%). All events containing more than two transgene copies showed repeat formation at least at one locus, with direct repeats again dominant (77%). Loci with two direct repeats had substantially greater transgene expression level than other types of two-copy T-DNA configurations, but insert organization was not associated with stability of transgene expression. Use of the poplar *rbc*S promoter, which drove *BAR* in the transgenic constructs, had no adverse effect on transgene expression levels or stability compared with the heterologous CaMV *35S* promoter, which directed *GFP* expression.

Keywords: BAR gene, functional genomics, GFP gene, Populus tremula × *Populus tremuloides, Populus tremula* × *Populus alba, rbcS promoter, transgene expression stability.*

Introduction

Transgenes are widely known to be susceptible to loss of expression as a result of gene loss and transcriptional or post-transcriptional gene silencing (TGS or PTGS, respectively) (Stam et al. 1997, Kooter et al. 1999, Fagard and Vaucheret 2000). Although the rates and the causes of instability vary widely among species, environments and transformation systems, the detection and removal of unstable transgenic events are the essential steps for commercial development (Meza et al. 2001, Kohli et al. 2003).

Unstable expression of transgenes is of particular concern for trees as a result of their long life cycles, difficult transformation and general absence of sexual propagation before field and commercial deployment (Han et al. 1997, Bradshaw et al. 2000). Meiosis is widely known to promote gene silencing, allowing unstable transgenic lines to be rapidly recognized and avoided in further development (Scheid et al. 1991, Iglesias et al. 1997, Metz et al. 1997). For some applications of transgenic trees, biosafety traits, such as reproductive sterility, may be required by regulatory agencies, marketplace forces, or stewardship programs (Strauss et al. 1995, Brunner et al. 2007). High levels of stability may, therefore, be dictated by social and commercial needs, as well as by regulatory requirements.

Expression of a transferred gene can be initially silenced or inactivated over time or generations by either TGS or PTGS triggered by diverse host–defense responses (Matzke et al. 2000). Several factors that might serve as a trigger for silencing mechanisms include transgene copy number, T-DNA structure and integration sites. The presence of multiple copies of

* Present address: Department of Plant and Microbial Biology, The University of California, Berkeley, CA 94720-3102, USA

© The Author 2008. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oxfordjournals.org

transcripts that are coded by transgenes might cause homology-dependent transgene silencing due to TGS or PTGS. Studies on the correlations between transgene silencing and copy number have shown conflicting results; in most cases no correlation was observed (Scheid et al. 1991, Gallo-Meagher and Irvine 1996, Iglesias et al. 1997, Kohli et al. 1999, Meza et al. 2001, Hawkins et al. 2003, Leibbrandt and Snyman 2003). Formation of transgene duplications and its association with stability have been reported in both annual plants and trees (Stam et al. 1997, Kumar and Fladung 2000, Lechtenberg et al. 2003). Transgenes that are organized as inverted repeats can show low expression or complete silencing (Morino et al. 1999), but this is not always the case (Meza et al. 2002, Lechtenberg et al. 2003).

Stress-induced transgene inactivation has been reported in plants (reviewed in Broer 1996). For instance, a heat treatment (37 °C) lasting for 10 days resulted in an almost complete (95%) loss of the phosphinothricin resistance in suspension culture cells derived from a single *Medicago sativa* transgenic line (Walter et al. 1992). Studies on transgenic maize (*Zea mays* L.) resistant to European corn borer (*Ostrinia nibilalis* L.) showed that water stress affects the level of expressed *Bt* proteins (Traore et al. 2000). The various stresses that occur during plant transformation and tissue culture are also known to cause genomic instability (reviewed in Phillips et al. 1994). Choi et al. (2000) reported increased chromosomal variation in transgenic barley (*Hordeum vulgare* L.) plants. Of 59 independent transgenic lines, 27 (46%) were tetraploid or aneuploid around the tetraploid level, whereas non-transgenic (NT) plants regenerated after in vitro culture alone had a much lower percentage of tetraploids (0–4.3%). However, an amplified fragment length polymorphism study of four aspen transgenic lines and wild-type plants showed high genomic stability (reviewed in Hoenicka and Fladung 2006).

Transgenic instability has been frequently observed in annual plants undergoing sexual propagation (reviewed in Brunner et al. 2007). Several studies have been conducted in perennial plants, including poplar trees, over multiple seasons and after vegetative propagation (Gallo-Meagher and Irvine 1996, Bettany et al. 1998, Cervera et al. 2000, Kumar and Fladung 2001, Meilan et al. 2002*a*, Hawkins et al. 2003, Leibbrandt and Snyman 2003). Most of these studies reported that transgene expression following vegetative propagation was highly stable and predictable. For instance, Meilan et al. (2002*a*) reported high stability of herbicide resistance in 40 poplar transgenic lines over 4 years in the field. Similarly, Li et al. (2008*a*) reported long-term stability of herbicide resistance in poplar over 8 years in the field. Long-term expression stability of the *GUS* gene was also demonstrated in another study in poplar (Hawkins et al. 2003). However, phenotypic instability of a *rolC* gene was observed in some transgenic aspen lines grown in vitro, in greenhouse, or in field conditions (Kumar and Fladung 2001). Instability of *rolC* expression was most obvious after transfer from in vitro to greenhouse conditions.

Here we report a study of stability of transgene expression in trees using a large number of transgenic events. This study was designed to be comparable to a commercial program where a large number of events are typically examined. Expression of 404 poplar transgenic events containing two reporter genes, green fluorescent protein (*GFP*) and bialaphos resistance gene (*BAR*), were studied under vegetative propagation in the greenhouse and in the field over 3 years. We investigated effects of several factors on transgene expression level, variability and stability, including the effects of secondary organogenesis, transgene copy number and T-DNA structure. We report extremely high levels of stability with respect to all these factors.

Materials and methods

Transgenic vectors, poplar transformation and organogenesis treatment

We used two reporter genes, *GFP* and *BAR*, to study transgene expression stability. The construction of transgenic vectors has been described in detail by Li et al. (2008*b*). The *GFP* gene was driven by the cauliflower mosaic virus *35S* promoter, and the *BAR* gene was driven by the promoter from the poplar gene for the small subunit of ribulose bisphosphate carboxylase (*rbc*S). The *NPTII* gene driven by the *NOS* promoter was used as the selectable marker. These three transgenic cassettes were assembled in the pGreenII binary vector to produce two versions of transgenic vectors: one designated pG3MKGB, which contained flanking matrix attachment region (MAR) elements from the tobacco *RB7* gene, and the other, pG3KGB, lacked MARs (Li et al. 2008*b*).

Two hybrid poplar genotypes INRA 353-38 (*Populus tremula* × *Populus tremuloides*) and 717-1B4 (*Populus tremula* × *Populus alba*) were used to transform the genetic constructs with *Agrobacterium tumefaciens* strain C58/pMP90 (GV3101) as described previously (Li et al. 2008*b*). A total of 404 independent transgenic events were produced. For organogenesis treatments, a total of 80 primary transformants, 20 per clone–construct combination, were randomly selected for secondary organogenesis. Leaf disks from those selected events were reintroduced into in vitro culture following protocols similar to the one used to produce the transgenic events, except that the leaf explants were neither co-cultivated with *Agrobacterium* nor was kanamycin added to the culture media. Two subevents (i.e., regenerated from two different explants) were derived from each of 80 reintroduced events. As a result, a total of 160 subevents [20 events × 4 (clone–construction) × 2 subevents] were regenerated from the 80 primary events.

Six to eight ramets for all transgenic events and subevents, and 32 NT plants for each of the two genotypes were

propagated in vitro, potted in soil and grown in greenhouse for 3–4 months. Together with NT plants, four ramets from the individual transgenic events and subevents were planted at a field site near Corvallis, OR, in October 2003. The plants were distributed in four blocks with a random split-plot design, where the two genotypes were separated and randomly assigned to subplots of each block, and transgenic events from one of two genotypes were randomly distributed within corresponding subplots. The plants were coppiced in spring 2004 to stimulate growth of multiple shoots, a common management and propagation practice for poplar. For an image of the field trial, see the cover of Plant Biotechnol. J. 6(9) 2008.

Measurement of transgene expression levels

Expression levels of the *GFP* and *BAR* transgenes were measured in three different years as previously described (Li et al. 2008*b*). Briefly, *GFP* expression was quantified with a hand-held *GFP*-meter (Opti-Sciences, Hudson, NH), and the relative concentration of the phosphinothricin acetyl transferase (PAT) encoded by *BAR* transgene was quantified by an enzyme-linked immunosorbent assay (ELISA) using the commercial LibertyLink PAT/*BAR* ELISA Kit (Envirologix Portland, Maine). For greenhouse plants, *GFP* measurements were made in September 2003 on the day before planting in the field. For the field plants, measurements were made in August during the day time, mainly from 0800 to 1300 h.

Comparative real-time PCR for transgene copy number and inverse PCR for T-DNA structure

To estimate the transgene copy number in transgenic plants, quantitative real-time polymerase chain reaction (qPCR) was performed on 396 transgenic events as previously described (Li et al. in press). Briefly, the single-copy endogene *PTLF* (poplar *LEAFY*/*FLORICAULA*) was used as a reference. The transgene *GFP* was chosen for the copy-number analysis because it was closer to the left-hand T-DNA border than the *BAR* gene. Real-time PCR was performed using dual-labeled TaqMan® probes (Biosearch Technologies, Novato, CA). For *GFP*, the primers used for the amplification were: 5′-ttaagggaatcgatttcaag-3′ and 5′-acgttgtgggagttgtagtt-3′. The hybridization probe was dual labeled with 6-hexachlorofluorescein (HEX) and Black Hole Quencher 1 (BHQ-1): 5′-HEX d(cctcggccacaagttggaatac) BHQ-1–3′. For the *PTLF* gene, PCR primer sequences were: 5′-ggtttctctgaggagccagtacag-3′ and 5′-gcctcccatgtccctcttc-3′. The hybridization probe was labeled with carboxyfluorescein (FAM) and carboxy tetraethyl-rhodamine (TAMARA): 5′-FAM d(caaggaggcagcagggagcggt) TAMARA3′. All PCR analyses were carried out using QuantiTect Multiplex PCR buffer (Qiagen, Valencia, CA) in a volume of 25 µl. For each sample, 100 ng genomic DNA purified with the Qiagen DNeasy Plant Mini Kit was used in a duplexed reaction to amplify both *PTLF* and *GFP* targets. Each sample was duplicated in two wells. The amplification was performed in an Mx3000P real-time PCR machine (Stratagene, La Jolla, CA) using the following program: 95 °C for 15 min, and 45 cycles of: 94 °C for 30 s, and 60 °C for 1 min. Transgene copy number of *GFP* transgene was determined by the formula: $2^{(1-\Delta C_t(GFP-\mathrm{PTLF}))}$.

To investigate T-DNA structure for inserts with multiple copies, five different primers located along the T-DNA sequence were used for PCR amplification as described by Li et al. (2008*b*). The presence and the size of amplified bands were used for inferring T-DNA structure.

Southern analysis to estimate transgene copy number

Southern blotting was performed to validate the comparative PCR method. Probes were labeled using PCR amplification with DIG-11-dUTP (Digoxigenin-11, 2′-deoxy-uridine-5′-triphosphate, alkali-labile; Roche, Indianapolis, IN, 25 µl). A 494-bp *GFP* probe fragment was amplified and labeled using the primers: 5′-tggccaacacttgtcactac-3′ and 5′-agaaggaccatgtggtctct-3′ in a 50-µl reaction with the following conditions: 200 µM dNTPs, with a ratio of 1:2 for DIG-dTTP and dTTP, 1.5 nM $MgCl_2$, 0.4 µM primers, 3 U *Taq* polymerase and 10 pg DNA for pG3KGB. A 570-bp *PTLF* probe fragment was amplified from genomic DNA extracted from one transgenic event. Forward and reverse primers used for amplification and labeling were: 5′-gccagtacagcaagacaagg-3′ and 5′-tgtgggtccaagacaagaac-3′. The labeling reaction was the same as for labeling of the *GFP* probe except that a ratio of 1:3 was used for DIG-dTTP and dTTP, and 10 ng of genomic template DNA was used. The PCR program for labeling was: 95 °C for 2 min, followed by 30 cycles of: 95 °C for 30 s, 56 °C for 30 s and 72 °C for 40 s and 72 °C for 7 min.

A total of 16 µg genomic DNA extracted from each transgenic sample was digested separately with *Hind*III and *Sca*I. Both enzymes cut within the T-DNA but outside of the *GFP* probe. The digested genomic DNA was electrophoretically separated along with 25 ng Digoxigenin-labeled DNA molecular mass marker II (Roche, Indianapolis, IN), on a 0.8% agarose gel (16 cm × 14 cm) for 20 h at 22 V. The DNAs were transferred to a Nytran SuperCharge membrane using TURBOBLOTTER™ Rapid Downward Transfer System (Schleicher & Schuell Bioscience, Keene, NH). The membrane was baked at 80 °C for 2 h to fix the DNA.

The probe hybridization and DIG luminescent detection were carried out using the DIG Luminescent Detection Kit (Roche, Indianapolis, IN). The prehybridization and hybridization of the *GFP* probe to the target was performed at 45 °C using 4 µl of PCR labeling reaction per milliliter of DIG Easy Hyb buffer (Roche, Indianapolis, IN) in a hybridization incubator with rotation. All steps were carried out according to the manufacturer's

instructions except that 1:3000 Anti-Digoxigenin-AP (23 mU ml^{-1}) was used. Membranes were exposed to X-ray film for 30 min.

After exposure, membranes were stripped twice for 15 min at 37 °C in 0.2 M NaOH containing 0.1% sodium dodecyl sulfate to remove the DIG-labeled *GFP* probe, rinsed for 5 min in 2 × SSC, and stored in 2 × SSC buffer. Stripped membrane was then prehybridized and hybridized with the *PTLF* probe at 42.5 °C using the same procedure for the *GFP* probe hybridization and detection. Developed films were scanned with a Personal Densitometer (Molecular Dynamics, Sunnyvale, CA) at the Center for Genome Research and Biocomputing at Oregon State University, and signal intensities were analyzed with ImageQuant 5.0 (Molecular Dynamics, Sunnyvale, CA).

Statistical analysis

The GFP fluorescence and PAT protein concentrations measured in three different years were normalized and transformed for statistical test as described previously (Li et al. 2008*b*). Except for testing the effect of transgene copy number on transgene variation among ramets and analysis of variance (ANOVA) of expression levels, mean expression values averaged over four ramets for each event were used for statistical analysis. Overall ANOVA of *GFP* expression was performed based on a random split-block design with field expression data (years 2004 and 2005). Event was considered a random effect, whereas block, clone and construct were treated as fixed effects. The ANOVA of *BAR* expression was performed on quantified PAT concentrations from all 3 years using a mixed model, with event as the only random effect.

For organogenesis treatment effects, linear regression and 99% prediction intervals were used to identify unstable subevents. For year-to-year stability, the median expression value of each of 3 years was computed and used to order the years from small to large. The expression value of each event was then plotted against the median expression value of the ordered years. The slopes and the pooled residual variances of the regression line for each event were used to identify unstable events over time. Normality of regression slope distribution was tested with a Quartile–Quartile (QQ) normal plot and Kolmogorov–Smirnov test. Unstable events with high sensitivity to the general environment (year of measurement) were then identified according to the normal distribution at significance level of 0.01. Events with extreme regression residual variances distributed in the tail of calculated variances were considered unstable with large random variation over years. Two-sample *t* tests (unequal variance) and Wilcoxon rank tests were used to test effects of transgene copy number and T-DNA structure on regression slope and residual variance, respectively. Levene's test on equality of variances was used to examine the effect of transgene copy number on variance of expression levels among events.

Results

Variation of transgenic expression among transgenic events

Two transgenic constructs, one of which contained *GFP* and *BAR* transgenes flanked by the MAR element from tobacco *RB7* (designated pG3MKGB), and another which contained *GFP* and *BAR* transgenes only (designated pG3KGB), were transformed into two poplar hybrid clones 353-38 and 717-1B4 (Li et al. 2008*b*). A total of 404 independent transgenic events were produced. Among them, 99 events were for the pG3KGB in 353-38, 98 events for the pG3KGB in 717-1B4, 108 events for the pG3MKGB in 353-38 and 99 events for the pG3MKGB in 717-1B4.

Expression levels of *GFP* and *BAR* transgenes were studied for three consecutive years (2003–2005). We measured *GFP* expression on 1616 transgenic poplars that originated from the 404 independent transgenic events, whereas the expression of *BAR* gene was measured on a single ramet for all 198 events in the poplar genotype 717-1B4 background. For the final *BAR* stability analysis, we used only the expression levels of 168 events because some plants did not survive through to the second or the third year. A wide range of transgene expression levels were observed for both *GFP* and *BAR* transgenes (Figure 1). Of 404 transgenic events, four were initially silenced for *GFP* expression and remained so in the field (*GFP* expression same as for NT controls). Three were transformants of 717-1B4, and one was a transformant of 353-38. None of the three 717-1B4 transgenic events showed initial silencing of *BAR* gene based on ELISA results, but initial silencing of *BAR* occurred for the 353-38 event (data not shown).

No significant difference in mean *GFP* expression was observed between clones 353-38 and 717-1B4 ($P > 0.05$). There was also no significant difference in *GFP* expression between the two constructs (Li et al. 2008*b*), though the mean *BAR* expression was somewhat higher for non-MAR pG3KGB transformed events. The distribution of *GFP* and *BAR* expression followed a similar pattern in all the three different years (Figure 1), and there was no significant difference in distributions between any 2 years (Kolmogorov–Smirnov test, $P > 0.05$).

Stability of transgene expression over time

A strong correlation in expression levels was observed between years for both transgenes (Figure 2). The correlation in *GFP* fluorescence values between 2 years in the field appeared to be stronger than the correlation between the greenhouse and the field, whereas no clear correlation of this kind was seen for the measured relative PAT protein concentrations.

No instances of complete breakdown (gene silencing) over time were observed for either transgene. We used regression approaches to detect events with significant changes in expression over time. Large absolute slope values indicate high sensitivity of transgene expression to the

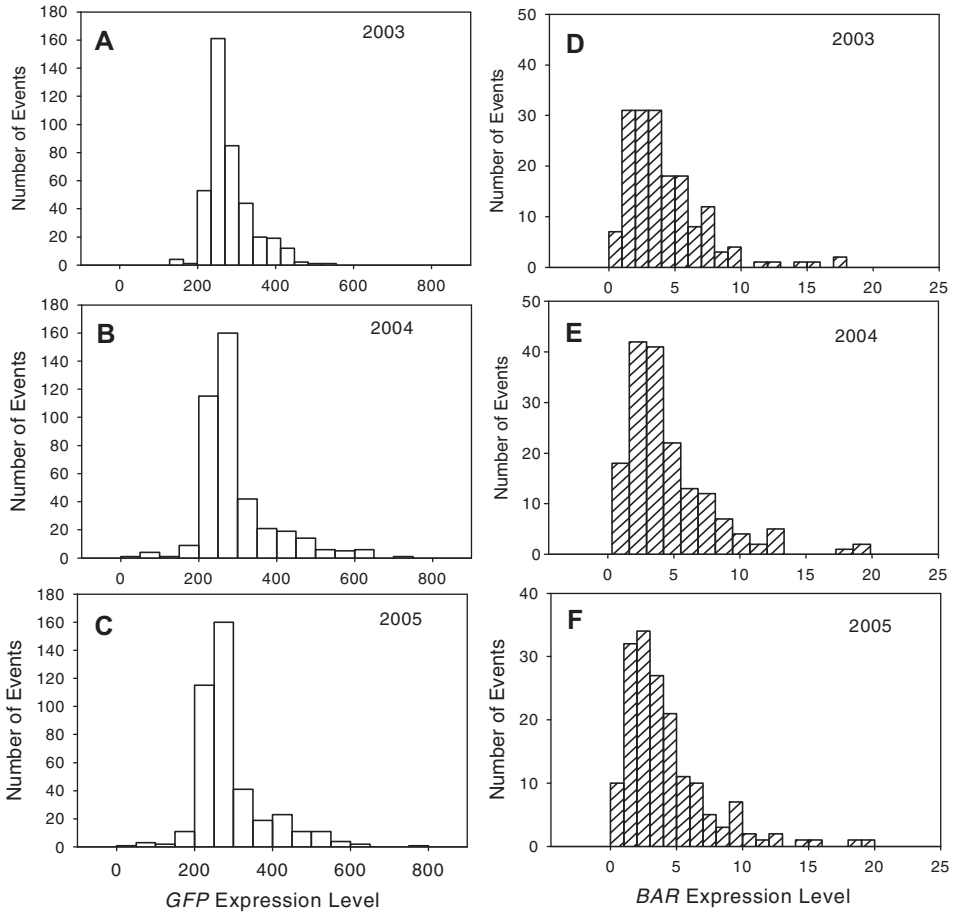

Figure 1. Histograms of *GFP* and *BAR* expression levels in three different years. Quantified expression values in the different years were scaled to their corresponding median values. (A–C) *GFP* expression levels of 404 transgenic events produced from poplar clones 717-1B4 and 353-38 in year 2003 (greenhouse), 2004 (field) and 2005 (field), respectively. Expression of *GFP* was quantified with a *GFP* meter (units in counts per seconds, 530 nm). (D–F) Quantified *BAR* expression levels of 168 transgenic events produced from 717-1B4 clone in years 2003–2005, respectively. Expression of *BAR* was quantified with an ELISA (normalized OD values).

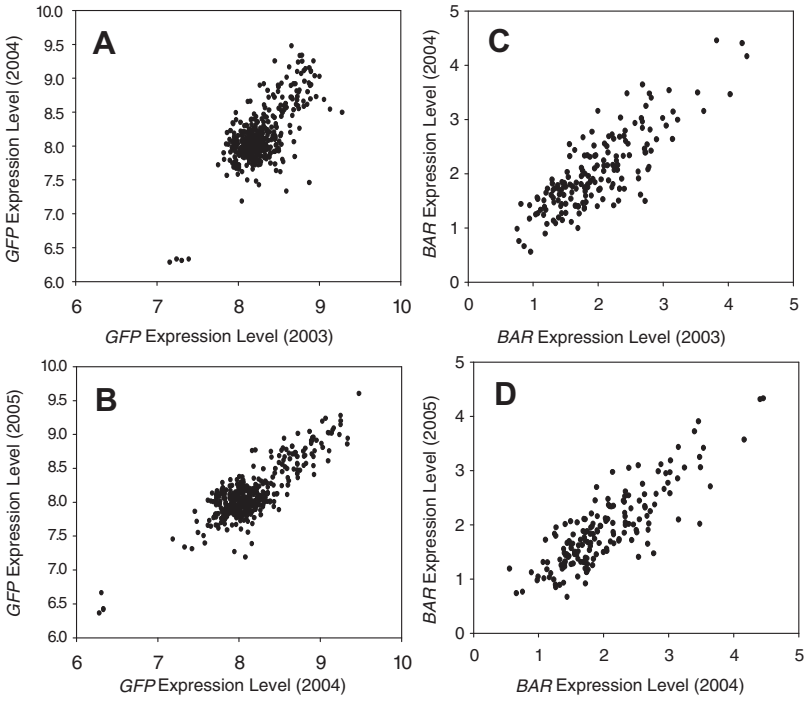

Figure 2. Correlation of transgene expression levels between different years. (A) Correlation of *GFP* expression levels between year 1 (2003) and year 2 (2004) ($n = 404$, $r = 0.70$). (B) Correlation of *GFP* expression levels between year 2 (2004) and year 3 (2005) ($n = 404$, $r = 0.86$). (C) Correlation of *BAR* expression levels between year 1 (2003) and year 2 (2004) ($n = 168$, $r = 0.81$). (D) Correlation of *BAR* expression levels between year 2 (2004) and year 3 (2005) ($n = 168$, $r = 0.84$). All correlations were statistically significant at $P < 0.01$.

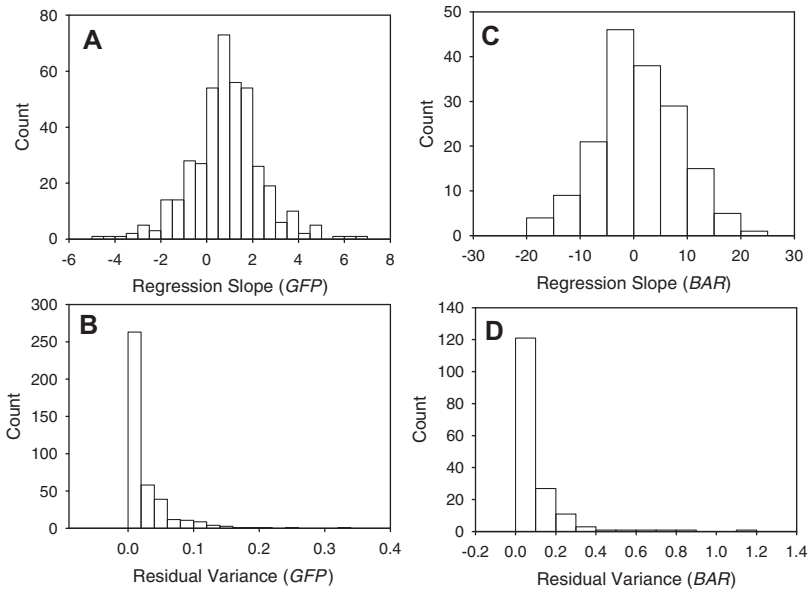

Figure 3. Regression slopes and residual variance for stability evaluation of transgene expression over 3 years in the greenhouse and in the field. Mean expression level of each event was plotted against median expression levels of each of 3 years. (A and B) Distributions of regression slopes (A) and residual variances (B) of *GFP* expression level of 404 transgenic events. (C and D) Distributions of regression slopes (C) and residual variances (D) of *BAR* expression levels of 168 transgenic events.

general environment (year of measurement), and large residual variances indicate high random variation not explained by the general environment in which the transgene was expressed. For *GFP* expression, the regression slopes showed a normal distribution (Figure 3), with a mean slope value of 0.81 and a standard deviation (SD) of 1.5. Four events with a slope value beyond three SDs of the mean were excluded from the estimate of the mean and the SD. Statistics based on the normality were then used to identify unstable events at a significance level of 0.01. In addition to the four excluded events, another four events were identified as highly sensitive to the year of measurement. Thus, a total of eight events (2%) showed significant changes in expression over time. Five of these showed a consistent increase in *GFP* expression, whereas the remaining three (0.7%) showed a consistent decrease. None of the three reduced expression events showed signs of complete gene silencing; all had an expression level at least two times greater than background values (NT controls). Based on chance alone, we expected a significant change in the expression in four of 404 total events.

For *BAR* expression, regression slopes also followed a normal distribution (Kolmogorov–Smirnov test, $P > 0.05$; Figure 3). The calculated means and SDs of the distribution were then used to identify events that were beyond the 99% distribution area ($P = 0.01$). One of the 168 (0.6%) events was identified as significantly sensitive to the environment with increased expression over time, also well within what is expected due to chance alone.

Pooled residual variances of each regression line are an indication of changes in transgene expression caused by random variation that was not explained by the average effect of year. A majority of events had a residual variance around zero, as expected (Figure 3), indicating that the majority of variation in transgene expression was explained by regression or by effect of year of measurement. We identified those with residual variances at the tails of the distributions as potentially unstable events. With this approach, four of the 404 (1%) events were identified for *GFP* expression and five of the 168 (3%) events were identified for *BAR* expression. Only the latter result exceeds what would be expected due to chance alone.

In summary, about 3% of the transgenic events showed variable *GFP* expression over time. If only reduced expression over time or in a certain year is considered, the frequency is about 1% at the event level, similar to the employed *P* value. About 4% of the events showed significantly variable *BAR* expression over time, and all of those variable events showed increased *BAR* expression in the second or the third year, when compared with the first year. Neither there were events that had statistically significant expression changes in both genes, nor were any strongly silenced events identified during the study period.

Effect of organogenesis on the stability of transgene expression

A total of 80 randomly selected transgenic events, 20 from each clone–construction combination, were subjected to a further round of organogenesis to produce subevents. Two independent subevents from each of 80 reintroduced events were produced; therefore a total of 160 subevents were studied. Quantification of *GFP* was done on all the 160 subevents, whereas PAT ELISA was performed on 76 subevents in the 717-1B4 background. Expression levels of regenerated subevents generally correlated well with their corresponding events (Figures 4 and 5), but there were several subevents that showed large changes in transgene expression levels compared with their progenitor events based on visual inspection of scatter plots. Four subevents

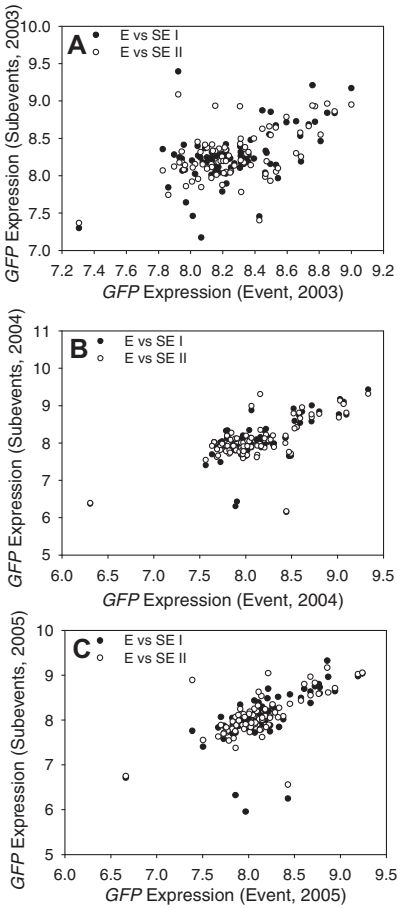

Figure 4. Scatter plots of *GFP* expression between parent events and their corresponding subevents in three different years. The 160 subevents were plotted against their corresponding events. (Abbreviations: E, event; S I, subevent I; and S II, subevent II.)

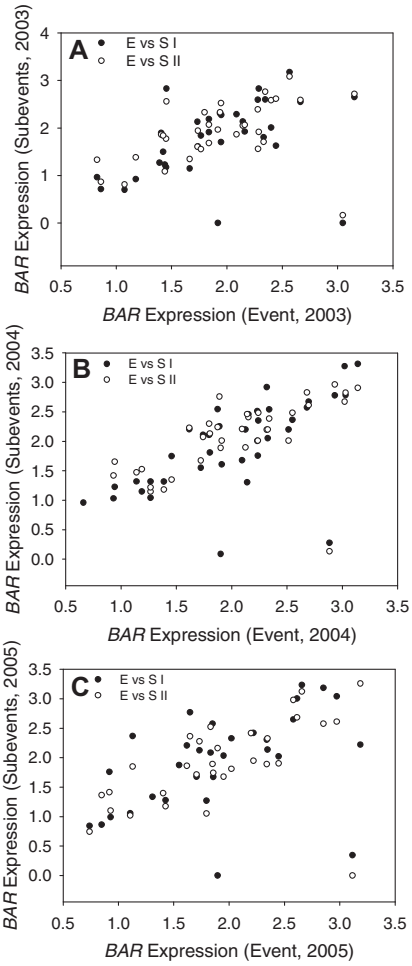

Figure 5. Scatter plots of *BAR* expression between parental events and their corresponding subevents in three different years. The 76 subevents were plotted against their corresponding events. (Abbreviations: E, event; S I, subevent I; and S II, subevent II.)

from three different events failed to express *GFP* and *BAR* transgenes above background values characteristic of NT controls, whereas expression levels of their corresponding parental events were in the middle or in the top half of the expression ranges. Amplifications by PCR with transgene-specific primers were performed on these four subevents and their corresponding events, and showed that all genes in the T-DNA, including the kanamycin selectable marker, were absent from these four subevents, whereas all their corresponding parental events had the present of transgenes (data not shown).

To further identify the subevents with statistically significant changes in transgene expression levels, regression analysis of expression levels of subevents on their corresponding events was performed. Regression prediction intervals (99%) were used to identify points falling outside the bounds as potential unstable subevents. The four subevents that lost the transgenic cassettes were excluded from the analysis. Significant outliers that affected normality of the regression residuals were excluded from the construction of prediction intervals, but considered as potential unstable subevents. Three subevents, 3mgb161S1, 3mgb161S2 and 3mgb281S2, always showed substantially elevated *GFP* expression in subevents in all 3 years, and were considered as outliers in regression analysis. The normality of regression residuals was checked with a QQ normal plot and tested with a Kolmogorov–Smirnov normality test ($P > 0.05$). Five additional subevents (3%) were identified as having significant changes in transgene expression levels, using the 99% prediction intervals, but instability was not observed in all 3 years. Among them, two subevents (1%) showed reduced *GFP* expression compared with the event from which they were derived, whereas the other three (2%) showed increased *GFP* expression. Based on the number of regression points assessed, we expected two of the 160 total points to be statistical outliers.

For *BAR* transgene expression, three subevents (4%) were identified as unstable by regression analysis, and two of them (2.5%) showed reduced *BAR* expression in subevents. One of 76 regression points would be expected due to chance alone. However, the instability was not maintained in all

3 years, and only one subevent, 7gb247S1, had significant expression changes for both genes in one of 3 years. None of the other identified subevents showed simultaneous expression changes in both transgenes.

Transgene copy number and its correlation with transgene expression and stability

Quantitative PCR has been previously used to estimate transgene copy number (Ingham et al. 2001). In this study, we used the endogenous gene *PTLF*, the poplar homolog of *LEAFY*, as a reference gene. *PTLF* exists as a single-copy gene in the poplar genome, which we confirmed by Southern blot analysis (data not shown; Rottmann et al. 2000). Transgene copy number was estimated using the *GFP* transgene, which is closer to the left border of the T-DNA than *BAR,* and thus more reflective of complete T-DNA transfers. The TaqMan® probe and the primer sets were optimized to achieve high and comparable PCR efficiency for both genes (data not shown). The DNA extracted from each of the 396 transgenic events was run, in duplicate, in a duplex TaqMan® assay. Transgene copy number was calculated by comparing the C_t (cycle-threshold, the number of PCR cycles in which all the PCR amplification graphs being compared are in the linear range) of the *GFP* transgene to that of the endogenous *PTLF*, as described in the section Materials and methods. We used duplicate instead of triplicate reactions, because the C_t values were generally very similar ($\Delta Ct < 0.2$, data not shown). There was also a good correlation in calculated copy number between runs (i.e., with different reaction mixes and plates run on different days) (data not shown).

We estimated copy number for 396 events, and one of these events had multiple copies but the exact number was not determined. We estimated copy number using the following ranges: QV < 1.5: one copy and 1.5 < QV < 2.5: two copies. We found that our *Agrobacterium*-mediated transformation system produced a majority of single-copy transformants (Table 1), with 85% of the transformants containing a single copy, 11% having two copies, 3% having three copies and the remaining 1% containing four copies. None of studied events contained > 4 copies of the insert. These results are generally consistent with previous observations of *Agrobacterium*-transformed dicotyledonous and monocotyledonous plants (Hansen and Chilton 1999, Ingham et al. 2001). The frequency distributions of copy number were also similar for the two constructs. For events transformed with the construct without MARs, 82.7% contained a single copy of the insert compared with 83.6% for events transformed with the construct containing MARs.

To validate this qPCR method to estimate copy number, we performed Southern blot analyses on a subset of transgenic events. Two restriction enzymes, *Sca*I and *Hind*III, were used separately to digest the genomic DNA. Both enzymes cut outside of the *GFP* probe sequences (Appendix – Figure A1). The number of hybridization bands from the *Sca*I digestion indicated the number of transgene loci, whereas signal intensities of hybridization bands from the *Hind*III digestion indicated the total number of transgene copies present in the genome.

We performed Southern blots on 15 events with a QV between 1.4 and 1.6, and on two events with a QV of 3.4 (Table 1). Of the 17 events studied, 14 events (82.3%) were consistent between Southern blot and qPCR (Table 1). In addition to those 17 events, another 18 events with a more decisive QV were randomly chosen for Southern analysis. For all the 35 events, we found a 95% correlation (r^2) between qPCR and Southern blot results when *Sca*I was used for restriction digestion. Nine of those events, and an additional event, were further confirmed with *Hind*III digestion. We found 100% correlation when qPCR-determined copy number was < 4 copies, as a result of hybridization signal reaching saturation at three copies (we found two samples with four copies as determined by qPCR had the same intensity as one with three copies). One of these two events was also digested by *Sca*I, and was shown to have four copies (Appendix – Figure A1, Lane 7). Thus,

Table 1. Estimated transgene copy number of the 396 transgenic samples and the correlation with Southern blot results.

Copy no.	No. of samples determined by qPCR assay (% of total)	Correlation with Southern blot			
		(*Sca*I)		(*Hind*III)	
		No. assayed	No. consistent (%)	No. assayed	No. consistent (%)
1	338 (85%)	17	16 (94.1%)[1]	2	2 (100%)
2	44 (11%)	11	9 (82%)[2]	2	2 (100%)
3	10 (3%)	4	4 (100%)	4	4 (100%)
4	4 (1%)	3	2 (67%)[3]	2	0 (0%)[4]
Total	396	35		10	

[1] The only inconsistent event had a QV of 1.4 and displayed two bands on the Southern blot.
[2] Two inconsistent events had a QV of 1.6 and displayed a single band on the Southern blot.
[3] One sample determined to have four copies by qPCR displayed two bands on the Southern blot. Copy number of this sample was set as multiple copies (not determined for its exact number due to unclear results based on Southern).
[4] Two samples determined to have four copies by qPCR assay displayed a signal intensity of three copies by Southern.

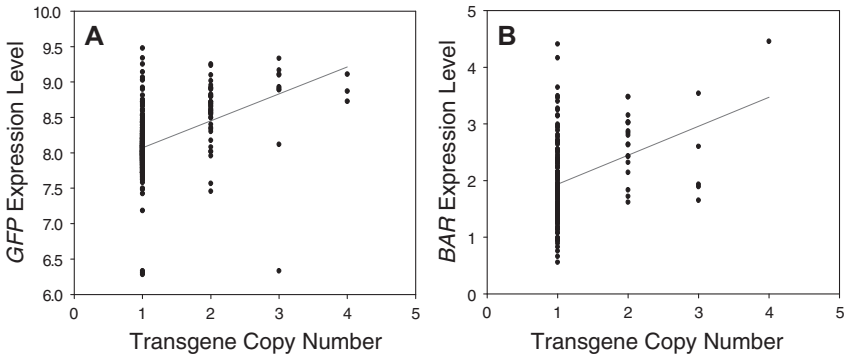

Figure 6. Correlation of transgene copy number with transgene expression level. (A) Correlation of *GFP* expression with transgene copy number ($n = 394$, $r = 0.46$, $P < 0.001$). (B) Correlation of *BAR* expression with transgene copy number ($n = 187$, $r = 0.35$, $P < 0.001$).

our qPCR method was highly effective for events with three or fewer copies, which represented the large majority of the population studied (99%).

Correlation of the transgene copy number with transgene expression level

The positive correlations between copy number and expression level were observed for both *GFP* and *BAR* transgenes (Figure 6). The events with multiple transgene inserts had a higher mean expression level ($\sim$ 1.5-fold for *GFP* and $\sim$ 2-fold for *BAR*); the difference in means was significant for both transgenes ($P < 0.05$). The regression of expression versus copy number explained 12% and 21% of its variance with *GFP* and *BAR*, respectively.

We performed Levene's test on equality of variance in transgene expression levels between single-copy and multi-copy groups. The presence of multiple copies increased variance in the *GFP* expression among events by 139% ($P < 0.05$). Although a similar effect was observed for the *BAR* expression (79% greater for high-copy events), it was not statistically significant ($P > 0.05$). A large part of the high variance for high-copy-number events was a result of one three-copy event that had silenced *GFP* expression, whereas the rest of the multi-copy events had high expression (Figure 6). With the silenced event excluded, however, the high-copy-number events still showed higher variance on both linear ($P < 0.0001$) and log scales ($P = 0.0178$). Multiple copies also tended to increase variance in the *GFP* expression among ramets ($P < 0.05$, data not shown).

T-DNA repeat formation

Multiple copies of transgenes can be arranged at one locus as direct, inverted, or combinations of both types of repeat (summarized in Krizkova and Hrouda 1998). We investigated repeat formation for the 42 events containing two copies of the transgenes, and for the 13 events containing three or four copies of the transgenes (Li et al. 2008*b*). The overall frequency for the integration of multiple copies at the same locus was 0.55; direct repeat was the major type of repeat formation (77%). We plotted *GFP* expression level of 40 two-copy events against repeat types to see if repeat formation and types of repeats affected transgene expression (Figure 7). Events with direct repeats had a 24% higher *GFP* expression than events without direct repeats ($P < 0.05$). Because of the small number of events with inverted repeats (2), no strong statistical inferences can be made regarding their effect on transgene expression. Both of the inverted repeat events showed *GFP* expression at the lower end of the range, and were significantly lower in expression than events with direct repeats ($P = 0.004$). However, the mean expression level of these two events did not differ significant from that of the events with no repeats ($P = 0.14$). There was no significant difference in the stability between the direct repeats and the other repeat types based on regression slopes (two sample t test, $P > 0.05$), regression residual variances (Wilcoxon rank test, $P > 0.05$), or variance among event means.

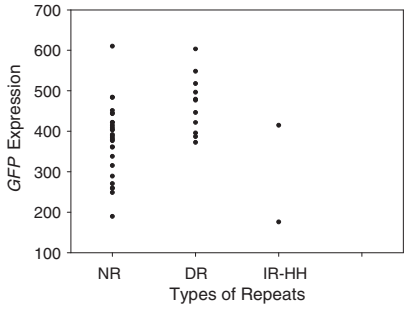

Figure 7. Correlation of *GFP* expression levels with repeat formation of T-DNAs. The scatter plot was based on 40 two-copy transgenic events whose T-DNA structure was determined by PCR amplification. (Abbreviations: NR, no repeats; DR, direct repeats; and IR-HH, head-to-head inverted repeats.)

Discussion

A major biological concern about transgenic trees is whether transgene expression will be stable over a full harvest rotation, which can span decades. Several previous studies with transgenic poplar trees have suggested that the expression of transgenes after vegetative propagation

can be highly stable over many years in the field (Meilan et al. 2002*b*, Hawkins et al. 2003, Li et al. 2008*b*). We monitored expression of 2256 transgenic poplar trees derived from 404 primary events and 160 subevents over 3 years in the greenhouse and in the field. Although some variation in the transgene expression level was observed among years, for the large majority of events there was a strong correlation in expression over time, and no cases of gene silencing were observed during the study period for two different transgenes. Although transgene expression seemed to change in both directions (i.e., reduced and elevated), of most concern for the application of transgenic trees in a commercial plantation is whether transgenes can be expressed sufficiently to impart targeted traits for the duration of the rotation. In this study, 3% of the transgenic events showed significant changes in *GFP* expression, which is just 1% above the employed *P* value, and 1% of the transgenic events showed reduced *GFP* expression in one or in both of 2 years in the field, which is expected to occur by chance alone. None of those events showed signs of complete gene silencing: their expression levels were more than twice that of the background value characteristic of NT plants. In contrast, no events showed significant reduction in *BAR* expression over time. The difference in observed instability in *GFP* and *BAR* may be associated with the greater accuracy of the ELISA used to determine relative protein concentrations for *BAR* compared with the fluorescence measurements used for *GFP*. It may also result from strong influences of leaf age and environment in which leaves develop a background fluorescence that obscures *GFP* detection (Halfhill et al. 2003).

The stresses that occur during tissue culture have been associated with genomic alterations (Choi et al. 2000). These include changes in chromosome structure, DNA methylation and DNA sequence (reviewed in Phillips et al. 1994). The stress imposed by in vitro conditions tends to induce genomic instability at particular loci, and these sites may also be preferential targets for the integration of exogenous DNA (Gould 1986, Romano et al. 2005). Transgene expression could, therefore, change as a consequence of plant tissue culture and regeneration. We reintroduced 80 primary transformants into tissue culture using our system, which involved callogenesis followed by organogenesis, to evaluate transgene stability during regeneration. Because selection and co-cultivation with *Agrobacterium* were not performed, this procedure is likely to provide a lower estimate of transformation-induced instability. A majority of subevents showed expression levels that were highly correlated with their corresponding parental events. Four subevents (2.5%) derived from three different events showed a complete loss of expression of both transgenes. Amplification by PCR with different pairs of primers confirmed that both *GFP* and *BAR* genes had been lost from those subevents, but present in their corresponding parental events and expressing subevents. Loss of transgenes during sexual propagation has been reported for annual plants (Srivastava et al. 1996, Feldmann et al. 1997, Joersbo et al. 1999). All the three events that gave rise to these four subevents contained a single copy of the insert; it is, therefore, unlikely that the loss of transgenes was caused by intrachromosomal recombination between transgenic copies integrated at different loci, as hypothesized by Fladung (1999) to explain transgene loss in poplar. Stress can increase the frequency of homologous recombination (Molinier et al. 2006), and if a transgene is inserted into duplicated region of the genome, this still may provide an explanation for transgene loss. Although transgene loss occurred somatically, it could also be the result of a yet unknown mechanism responsible for the non-Mendelian inheritance observed in the *Arabidopsis hothead* mutant (Lolle et al. 2005) and in varieties of flax (*Linum narbonense* L.) under certain environmental conditions (Chen et al. 2005). Given this, when reintroducing primary transformants into tissue culture, it is advisable to keep explants under selection pressure to minimize the possibility of transgene loss.

About 3% of subevents had elevated *GFP* expression. In particular, three MAR-containing subevents derived from two events appeared to develop elevated expression over time. It was not determined whether a similar expression change also happened to *BAR* transgenes in the three subevents, because *BAR* expression was not determined for genotype 353-38. About 1% of events showed elevated *BAR* expression in the first year, but this was not maintained in subsequent years in the field.

Transgene copy number and its effect on transgene expression and stability

Estimation of transgene copy number for a large number of samples by traditional Southern blots is time-consuming and labor-intensive. We found high levels of accuracy of TaqMan®-based assays, a result of consistency among different replicates and runs, and high correlations with Southern blot results. However, the limitations of the TaqMan® assay were also evident. There were cases where calculated copy number was between integers and, therefore, difficult to assign to a copy-number class. For these cases, additional qPCR runs or other assays, such as Southern blots, are needed for confirmation.

Truncations of T-DNA may have caused some of the problems in copy-number determination. When one of the multiple copies is truncated, qPCR might fail to detect the truncated copy. This is also true for Southern blots, which will either fail to detect a signal, or produce a weaker hybridization signal from a partial truncation. T-DNA truncations can often happen when multiple insert copies are present (discussed below). When there are no truncations, all samples should produce a single hybridization band from the *Hind*III digestion. The presence of additional bands may indicate that one of the *Hind*III sites on the T-DNA was lost as a result of truncation. This was

the case for three events with multiple copies. It is, therefore, advisable not to choose PCR primers or hybridization probes from sequences close to the T-DNA borders. Quantitative PCR is a better choice than Southern blots when > 2 copies of T-DNA are arranged as tandem repeats, or when inverted repeats are present. Under these circumstances, the number of hybridization bands will not represent the actual number of inserts, and band intensity must, therefore, be carefully quantified.

Previous studies on the effect of transgene copy number on transgene expression and stability have given conflicting results. Copy number and transgene expression levels can be positively correlated (Voelker et al. 1992, Hobbs et al. 1993, van der Hoeven et al. 1994, McCabe et al. 1999, Tang et al. 2003), negatively correlated (Mannerlöf et al. 1997, Cervera et al. 2000) or uncorrelated (Hobbs et al. 1993, Bauer et al. 1998, McCabe et al. 1999). This study showed a positive and significant, but weak, correlation between transgene expression and copy number for both *GFP* and *BAR*. Contrary to the expectation based on the literature on homology-dependent gene silencing (reviewed in Stam et al. 1997), copy number was not correlated with initial transgene silencing nor with stability following organogenesis or over time. Of four events that were initially silenced for *GFP* expression, three contained one copy, the other one had three copies. All four subevents that lost their transgenes during organogenesis were derived from single-copy transgenic events. It is possible that our transgenic tissues had some level of chimerism; this study did not examine instability in expression at the cellular level, but instead focused on the tissue and organism levels. However, if chimerism were common we would have expected to see it reflected as quantitative instability among tissues/years, or as multiple PCR products in our silenced events after organogenesis. Nonetheless, the conventional wisdom that simple, single-copy insertion events should be the ones chosen for research or commercial purposes was not supported by this study of vegetatively propagated poplars. However, the larger numbers of copies and partial copies that can be produced from biolistic transformation could behave differently from the one to four *Agrobacterium*-delivered copies that we studied.

The relationship between copy number and expression was unaffected by the presence of MARs. Therefore, the hypothesis that MARs may prevent homology-dependent gene silencing when multiple copies are present (Allen et al. 1993) was not supported by our observations. The presence of multiple copies substantially increased variance of expression of both transgenes at the event or ramet level, or both, although the effect was not statistically significant for the *BAR* transgene. There was also a significant interaction effect of MARs with copy number with respect to expression. The highest variance at the event level was seen in transgenic events containing multiple copies without flanking MARs. However, this effect was not seen in the *BAR* transgenic population. Thus, MARs had a very weak effect in buffering the variance caused by multiple transgenes (Li et al. 2008*b*)

It is generally believed that homologous promoters should be avoided, because they can lead to transgene and native gene silencing (reviewed in Rathore and Sunilkumar 2005). We tested this hypothesis by using a native poplar promoter from the *rbc*S gene to drive *BAR*. Contrary to expectation, there was no obvious increase in gene silencing with this promoter. Although a small difference in silencing could have been missed as a result of using a different coding gene (*GFP*) and strongly expressing promoter (*35S*) for comparison, it appears that at least some native promoters can be highly reliable in transgenic plants. There are other examples of successful application of native promoters in driving transgene expression, including promoters from the following genes: maize ubiquitin, rice glutelin, rice actin, rice LHCP (light harvesting chlorophyll a/b-binding protein of photosystem II) and rice cytochrome C (summarized in Sunilkumar et al. 2005). The latter study also tested whether the application of a homologous α-globulin B promoter in transgenic constructs can negatively affect the expression of the native gene, and found that the expression of α-globulin B storage protein in cottonseed was not adversely affected.

Repeat formation in the multi-copy events and their effect on expression and stability

The presence of repeated structures can be accompanied by transgene silencing at both transcriptional and post-transcriptional levels (reviewed in Vaucheret and Fagard 2001). Of 40 two-copy transgenic events, 32.5% carried direct repeats in the same locus, and 5% had inverted repeats arranged head-to-head. Unexpectedly, transformants carrying direct repeats at one locus had significantly elevated transgene expression compared with those with two copies inserted at different loci. Inverted DNA repeats have been demonstrated to be associated with TGS or PTGS, or both (Muskens et al. 2000). Inverted repeat structure was consistently associated with T-DNA truncations in this study, but the two events carrying inverted repeats (head-to-head) had stable expression. Inverted and direct repeat arrangements, or T-DNA truncations, were also not sufficient to induce transgene silencing in *Arabidopsis* (Meza et al. 2002, Lechtenberg et al. 2003).

In all transformants containing three or four copies, at least two of the copies were inserted at the same locus. A majority of these (67%) carried direct repeats, 6.7% had inverted repeats and the remaining 26.7% had both types of repeats. Inverted repeats were accompanied by T-DNA truncations. As shown by the positive correlation between copy number and transgene expression, repeat formation did not reduce overall expression level. It also did not increase instability over time.

Transgene silencing, both initial and subsequent silencing over time, appears to be a minor concern in

Agrobacterium-transformed poplar. Because of the wide range but consistency of transgene expression, it is advisable to perform early screening to select desired expression levels. Once selected, our results suggest that expression will remain consistent under field conditions. We found a high percentage of single-copy integration of T-DNAs, a positive correlation between transgene copy number and expression, and a high percentage of strongly expressing events that had direct repeats at single loci. Selection of transformants on the basis of expression level, without molecular characterization, therefore, appears to provide a reliable prediction about the level of expression that will occur in future years.

Although we observed very high stability of transgene expression, this study spanned only 3 years, two types of transgenes, and two genotypes. Studies that span full propagation and commercial harvesting cycles – and have a much wider range of genes, germplasm and environments – are highly desirable. In addition, studies are needed that evaluate stability after sexual propagation, because instability might be reversed or amplified in association with meiosis. Current regulations effectively assume that even small releases of reporter transgenes are hazardous until proven otherwise on an event-by-event basis (USDA 2007). Moreover, large and long-term studies are difficult and expensive to conduct. It is likely that commercial uses of transgenic trees will be needed before large-scale studies of stability are feasible. Nonetheless, our results strongly suggest that instability of transgene expression, including in multi-copy transformants, is low, and stable events can be readily identified.

Acknowledgments

We thank Dr. C. Neal Stewart for the gift of *mgfp5er*, Dr. Steven Spiker for providing the MAR element and Dr. John Davis for the use of his poplar *rbc*S cDNA. We also thank Dr. Lisa Ganio for statistical advice and Dr. Christopher J. Bayne for sharing his microtiter plate reader. The research project was supported by USDA Initiative for Future Agriculture and Food Systems Grant No. 00-52100-9623, industrial members of the TBGRC/TGERC Research Cooperative at Oregon State University and the OSU Foundation Forest Biotechnology Fund.

References

Allen, G.C., G.E. Hall Jr., L.C. Childs, A.K. Weissinger, S. Spiker and W.F. Thompson. 1993. Scaffold attachment regions increase reporter gene expression in stably transformed plant cells. Plant Cell 5:603–613.

Bauer, M., J. Libantová, J. Moravíková and I. Békésiová. 1998. Transgenic tobacco plants constitutively expressing acidic chitinase from cucumber. Biologia 53:749–758.

Bettany, A., S. Dalton, E. Timms and P. Morris. 1998. Stability of transgene expression during vegetative propagation of protoplast-derived tall fescue (*Festuca arundinacea* Schreb.) plants. J. Exp. Bot. 49:1797–1804.

Bradshaw, H.D., R. Ceulemans, J. Davis and R. Stettler. 2000. Emerging model systems in plant biology: poplar (*Populus*) as a model forest tree. J. Plant Growth Regul. 19:306–313.

Broer, I. 1996. Stress inactivation of foreign genes in transgenic plants. Field Crops Res. 45:19–25.

Brunner, A.M., J. Li, S.P. DiFazio et al. 2007. Genetic containment of forest plantations. Tree Genet. Genomes 3:75–100.

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

Chen, Y., R.G. Schneeberger and C.A. Cullis. 2005. A site-specific insertion sequence in flax genotrophs induced by environment. New Phytol. 167:171–180.

Choi, H.W., P.G. Lemaux and M.J. Cho. 2000. Increased chromosomal variation in transgenic versus nontransgenic barley (*Hordeum vulgare* L.) plants. Crop Sci. 40:524–533.

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

Feldmann, K.A., D.A. Coury and M.L. Christianson. 1997. Exceptional segregation of a selectable marker (KanR) in *Arabidopsis* identifies genes important for gametophytic growth and development. Genetics 147:1411–1422.

Fladung, M. 1999. Gene stability in transgenic aspen (*Populus*). I. Flanking DNA sequences and T-DNA structure. Mol. Gen. Genet. 260:574–581.

Gallo-Meagher, M. and J.E. Irvine. 1996. Herbicide resistant transgenic sugarcane plants containing the *BAR* gene. Crop Sci. 36:1367–1374.

Gould, A.R. 1986. Factors controlling generation of variability in vitro. *In* Cell Culture and Somatic Cell Genetics of Plants. Ed. I.K. Vasil. Academic Press, Orlando, FL, pp 549–567.

Halfhill, M.D., R.J. Millwood, T.W. Rufty, A.K. Weissinger and C.N. Stewart. 2003. Spatial and temporal patterns of green fluorescent protein (GFP) fluorescence during leaf canopy development in transgenic oilseed rape, *Brassica napus* L. Plant Cell Rep. 22:238–343.

Han, K.-H., C. Ma and S.H. Strauss. 1997. Matrix attachment regions (MARs) enhance transformation frequency and transgene expression in poplar. Transgenic Res. 6: 415–420.

Hansen, G. and M. Chilton. 1999. Lessons in gene transfer to plants by a gifted microbe. Curr. Top. Microbiol. Immunol. 240:22–58.

Hawkins, S., J.-C. Leplé, D. Cornu, L. Jouanin and G. Pilate. 2003. Stability of transgene expression in poplar: a model forest tree species. Ann. For. Sci. 60:427–438.

Hobbs, S.L.A., T.D. Warkentin and C.M.O. DeLong. 1993. Transgene copy number can be positively or negatively associated with transgene expression. Plant Mol. Biol. 21:17–26.

Hoenicka, H. and M. Fladung. 2006. Genome instability in woody plants derived from genetic engineering. *In* Tree Transgenesis – Recent Developments. Eds. M. Fladung and D. Ewald. Springer-Verlag, Berlin, pp 301–321.

Iglesias, V.A., E.A. Moscone, I. Papp, F. Neuhuber, S. Michalowski, T. Phelan, S. Spiker, M. Matzke and A.J.M. Matzke. 1997. Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco. Plant Cell 9:1251–1264.

Ingham, D.J., S. Beer, S. Money and G. Hansen. 2001. Quantitative real-time PCR assay for determining transgene copy number in transformed plants. BioTechniques 31:132–140.

Joersbo, M., J. Brunstedt, J. Marcussen and F.T. Okkels. 1999. Transformation of the endospermous legume guar (*Cyamopsis tetragonoloba* L.) and analysis of transgene transmission. Mol. Breed. 5:521–529.

Kohli, A., D. Gahakwa, P. Vain, D.A. Laurie and P. Christou. 1999. Transgene expression in rice engineered through particle bombardment: molecular factors controlling stable expression and transgene silencing. Planta 208:88–97.

Kohli, A., R.M. Twyman, R. Abranches, E. Wegel, E. Stoger and P. Christou. 2003. Transgene integration, organization and interaction in plants. Plant Mol. Biol. 52:247–258.

Kooter, J.M., M.A. Matzke and P. Meyer. 1999. Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends Plant Sci. 4:340–347.

Krizkova, L. and M. Hrouda. 1998. Direct repeats of T-DNA integrated in tobacco chromosome: characterization of junction regions. Plant J. 16:673–680.

Kumar, S. and M. Fladung. 2000. Transgene repeats in aspen: molecular characterisation suggests simultaneous integration of independent T-DNAs into receptive hotspots in the host genome. Mol. Gen. Genet. 264:20–28.

Kumar, S. and M. Fladung. 2001. Gene stability in transgenic aspen (*Populus*). II. Molecular characterization of variable expression of transgene in wild and hybrid aspen. Planta 213:731–740.

Lechtenberg, B., D. Schubert, A. Forsbach, M. Gils and R. Schmidt. 2003. Neither inverted repeat T-DNA configurations nor arrangements of tandemly repeated transgenes are sufficient to trigger transgene silencing. Plant J. 34:507–517.

Leibbrandt, N.B. and S.J. Snyman. 2003. Stability of gene expression and agronomic performance of a transgenic herbicide-resistant sugarcane line in South Africa. Crop Sci. 43:671–677.

Li, J., R. Meilan, C. Ma, M. Barish and S.H. Strauss. 2008*a*. Stability of herbicide resistance over eight years of coppice in field-grown, genetically engineered poplars. West. J. Appl. For. 23:89–93.

Li, J., A.M. Brunner, R. Meilan and S.H. Strauss. 2008*b*. MAR elements have small and variable effects on transgene expression and stability in field-grown *Populus*. Plant Biotech. J. 6:887–896.

Lolle, S.J., J.L. Victor, J.M. Young and R.E. Pruitt. 2005. Genome-wide non-mendelian inheritance of extra-genomic information in *Arabidopsis*. Nature 434:505–509.

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

Matzke, M.A., M.F. Mette and A.J.M. Matzke. 2000. Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Mol. Biol. 43:401–415.

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

Meilan, R., D.J. Auerbach, C. Ma, S.P. DiFazio and S.H. Strauss. 2002*a*. Stability of herbicide resistance and GUS expression in transgenic hybrid poplars (*Populus* sp.) during four years of field trials and vegetative propagation. Hortscience 37:277–280.

Meilan, R., K. Han, C. Ma et al. 2002*b*. The *CP4* transgene provides high levels of tolerance to Roundup® herbicide in field-grown hybrid poplars. Can. J. For. Res. 32:967–976.

Metz, P., E. Jacobsen and W. Stiekema. 1997. Occasional loss of expression of phosphinothricin tolerance in sexual offspring of transgenic oilseed rape (*Brassica napus* L.). Euphytica 98:189–196.

Meza, T.J., D. Kamfjord, A.-M. Håkelien, I. Evans, L.H. Godager, A. Mandal, K.S. Jakobsen and R.B. Aalen. 2001. The frequency of silencing in *Arabidopsis thaliana* varies highly between progeny of siblings and can be influenced by environmental factors. Transgenic Res. 10:53–67.

Meza, T.J., B. Stangeland, I.S. Mercy et al. 2002. Analyses of single-copy *Arabidopsis* T-DNA-transformed lines show that the presence of vector backbone sequences, short inverted repeats and DNA methylation is not sufficient or necessary for the induction of transgene silencing. Nucleic Acids Res. 30:4556–4566.

Molinier, J., G. Ries, C. Zipfel and B. Hohn. 2006. Transgeneration memory of stress in plants. Nature 442:1046–1049.

Morino, K., O.-A. Olsen and K. Shimamoto. 1999. Silencing of an aleurone-specific gene in transgenic rice is caused by a rearranged transgene. Plant J. 17:275–285.

Muskens, M.W.M., A.P.A. Vissers, J.N.M. Mol and J.M. Kooter. 2000. Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. Plant Mol. Biol. 43:243–260.

Phillips, R.L., S.M. Kaeppler and P. Olhoft. 1994. Genetic instability of plant tissue cultures: breakdown of normal controls. Proc. Natl. Acad. Sci. USA, 91:5222–5226.

Rathore, K.S. and G. Sunilkumar. 2005. Homologous promoter use in genetic modification. ISB News Report, September 6–7.

Romano, E., A. Soares, K. Proite, S. Neiva, M. Grossi, J.C. Faria, E.L. Rech and F.J.L. Aragão. 2005. Transgene elimination in genetically modified dry bean and soybean lines. Genet. Mol. Res. 4:177–184.

Rottmann, W.H., R. Meilan, L.A. Sheppard, A.M. Brunner, J.S. Skinner, C. Ma, S. Cheng, L. Jouanin, G. Pilate and S.H. Strauss. 2000. Diverse effects of overexpression of *LEAFY* and *PTLF*, a poplar (*Populus*) homolog of *LEAFY*/*FLORICAULA*, in transgenic poplar and Arabidopsis. Plant J. 22:235–246.

Scheid, O.M., J. Paszkowski and I. Potrykus. 1991. Reversible inactivation of a transgene in *Arabidopsis thaliana*. Mol. Gen. Genet. 228:104–112.

Srivastava, V., V. Vasil and I.K. Vasil. 1996. Molecular characterization of the fate of transgenes in transformed wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 92: 1031–1037.

Stam, M., J.N.M. Mol and J.M. Kooter. 1997. The silence of genes in transgenic plants. Ann. Bot. 79:3–12.

Strauss, S.H., W.H. Rottmann, A.M. Brunner and L.A. Sheppard. 1995. Genetic engineering of reproductive sterility in forest trees. Mol. Breed. 1:5–26.

Sunilkumar, G., L.M. Campbell, M. Hossen, J.P. Connell, E. Hernandez, A.S. Reddy, C.W. Smith and K.S. Rathore. 2005. A comprehensive study of the use of a homologous promoter in antisense cotton lines exhibiting a high seed oleic acid phenotype. Plant Biotechnol. J. 3:319–330.

Tang, J., R. Scarth and B. Fristensky. 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. Mol. Breed. 12:71–81.

Traore, S.B., R.E. Carlson, C.D. Pilcher and M.E. Rice. 2000. Bt and non-Bt maize growth and development as affected by temperature and drought stress. Agron. J. 92:1027–1035.

USDA. 2007. Introduction of genetically engineered organisms. USDA APHIS Draft Programmatic Environmental Impact Statement. Available at: http://www.aphis.usda.gov/newsroom/content/2007/07/content/printable/complete_eis.pdf; last accessed May 12, 2008.

van der Hoeven, C., A. Dietz and J. Landsmann. 1994. Variability of organ-specific gene expression in transgenic tobacco plants. Transgenic Res. 3:159–166.

Vaucheret, H. and M. Fagard. 2001. Transcriptional gene silencing in plants: targets, inducers and regulators. Trends Genet. 17:29–35.

Voelker, T.A., A.C. Worrell, L. Anderson, J. Bleibaum, C. Fan, D.J. Hawkins, S.E. Radke and H.M. Davies. 1992. Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257:72–74.

Walter, C., I. Broer, D. Hillemann and A. Pühler. 1992. High frequency, heat treatment-induced inactivation of the phosphinothricin resistance gene in transgenic single cell suspension cultures of *Medicago sativa*. Mol. Gen. Genet. 235:189–196.

Appendix

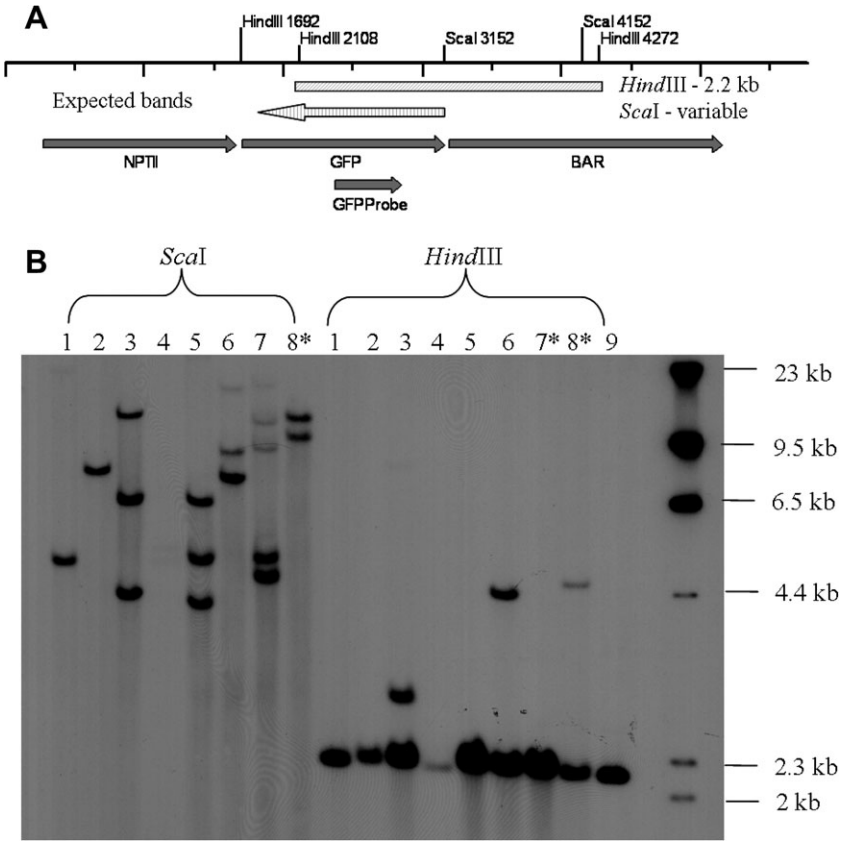

Figure A1. Southern analysis of transgene copy number. (A) Graphic representation of the T-DNA of pG3KGB. The relative positions of two restriction enzymes *Sca*I and *Hind*III, three transgenes (*NPTII*, *GFP* and *BAR*), and *GFP* probe region are shown. Expected hybridization bands are indicated by blue arrows. (B) Southern blot of genomic DNA extracted from eight transgenic events (Lanes 1–8) plus plasmid DNA of transgene construct pG3KGB (Lane 9). Aliquots of genomic DNA were digested with *Sca*I or *Hind*III, and probed with a DIG-labeled *GFP*-specific probe. Inconsistent events between Southern blot and qPCR are indicated by an asterisk (*). The DIG-labeled DNA molecular mass marker is shown at the right.