AN ABSTRACT OF THE DISSERATION OF

Jingyi Li for the degree of Doctor of Philosophy in Forest Science presented on September 20, 2006.

Title: Stability of Reporter Gene Expression and RNAi in Transgenic Poplars over Multiple Years in the Field under Vegetative Propagation.

Abstract approved:

_____

Steve H. Strauss Amy M. Brunner

Stable expression of transgenes is required for commercial uses of genetically engineered trees. To better understand the stability of transgene expression under field conditions, we studied transgene expression and RNA interference (RNAi)-induced transgene suppression in 2,480 transgenic poplars (460 transgene insertion events) over three years. Stability of expression was assessed based on two reporter genes, green fluorescent protein (*GFP*) driven by a 35S promoter and the herbicide resistance gene *BAR* driven by a poplar *rbc*S promoter. No cases of gene silencing (complete breakdown of expression) were observed for either gene, although physical loss of the transgenes occurred in three events after 80 events had been subject to further organogenesis in tissue culture. Flanking MARs did not significantly elevate transgene expression or stability, but reduced variance in expression among the events. MARs increased the correlation of expression between events for *GFP* and *BAR* genes in the same T-DNA. A majority of transformants (85%) carried single copy inserts; transgene copy number was positively correlated with expression (r = 0.46 for

GFP, and 0.35 for *BAR*). Formation of direct repeats was frequently observed in transgenic events containing multiple inserts of T-DNA, but did not adversely affect transgene expression.

RNAi using inverted repeats (IR) directed at the coding sequence gave a high degree of gene suppression of a resident *BAR* transgene; 80% of transgenic events showed more than 90% suppression. IR directed at the promoter sequence, however, was very inefficient in inducing gene suppression; only 6% of transgenic events showed more than 90% suppression. RNAi efficiency was unaffected by the presence of MARs. The degree of RNAi suppression was stable over two years in the field, as well as during seasonal development. Copy number of integrated IR loci was also unassociated with frequency of gene suppression. DNA methylation was observed in the promoter region of the highly suppressed events containing an IR of the promoter sequence, and in the coding region of highly suppressed events containing IR directed at the coding sequence; however, there was no clear relationship of methylation to the level of gene suppression in coding region-directed RNAi.

Stability of Reporter Gene Expression and RNAi in Transgenic Poplars over Multiple Years in the Field under Vegetative Propagation

by

Jingyi Li

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for
the degree of

Doctor of Philosophy

Presented September 20, 2006
Commencement June 2007

Doctor of Philosophy dissertation of Jingyi Li presented on September 20, 2006

APPROVED:

_____

Co-Major Professor, representing Forest Science

_____

Co-Major Professor, representing Forest Science

_____

Head of the Department of Forest Science

_____

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

_____

Jingyi Li, Author

ACKNOWLEDGEMENTS

I thank my major professors, Dr. Steve Strauss and Dr. Amy Brunner, for their support, guidance, and inspirations for excellence in science and great personality. I especially thank Amy for her coming back for my defense, and her understanding and support at difficult moments of my life. I am grateful to my other committee members, Dr. Glenn Howe, Dr. Tony Chen, Dr. David Mok, Dr. Jeffrey Morrell, and Dr. Terry Brown, for their generous time and advice; to Dr. Stephen Difazio for his key role in granting proposal wiring; Dr. Richard Meilan for his early involvement and advice on the projects; Sheena Powell and Rian Phund for their significant contribution to lab work and field measurements; Dr. Lisa Ganio for her advice on statistical analysis; Olga Shevchenko for cloning and sequencing, and for always being there for me when I needed help and encouragements; to Caiping Ma for her help with plant transformation and field work; Yanghuan Bao for his generous time and help with the completion of my degree; Dr. Christopher J. Bayne for sharing the microtiter plate reader. I wish to thank other members of the Strauss lab for their help during my classwork and research, and the Forest Science Office Staff for their support.

I also would like to thank Plant Genetic Systems for kindly providing the vector pTTM8, Monsanto for the use of *CP4* gene, Dr. C.N. Stewart for the use of *mGFP5er* gene, Dr. S. Spiker for providing MAR, Dr. J. Davis for providing poplar *rbc*S cDNA, and Dr. P. Waterhouse for the use of pHANNIBAL. The research project was funded and supported by Initiative for Future Agriculture and Food Systems Grant no. 00-52100-9623 from the USDA Cooperative State Research, Education, and Extension Service, TBGRC/TGERC Research Cooperative, OSU Foundation Forest Biotechnology Fund, Forest Science Department, College of Forestry, and OSU. I am very grateful to the founders of the scholarships that I received: Henry and Mildred Fowells Fellowship, Oregon

Sport Laurel Fellowship, Forestry Graduate Fellowship, Conard P. Wessela Fellowship, Mary McDonald Fellowship, Forestry Legacy Fellowship, OSU Foundation Fellowship Tuition Support Scholarship, Sauber Fund Scholarship, and Student Travel Award.

CONTRIBUTION OF AUTHORS

Dr. Stephen Difazio played an important role in writing the grant proposal. Dr. Steve Strauss and Dr. Amy Brunner were involved in every aspect of the project.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF APPENDIX FIGURES

LIST OF APPENDIX FIGURES (Continued)

LIST OF APPENDIX TABLES

DEDICATION

I dedicate this thesis to my parents, Duoping and Ximei, whom I have a deep admiration and love for, and who taught me to be strong, to embrace life with enthusiasm, and to be generous and forgiving; to my husband Jie who makes me a better person, gives me strength, and supports me all the way with his unselfish sacrifices and enduring love; to my daughter Katherine who brings so much joy to my life.

Stability of Reporter Gene Expression and RNAi in Transgenic Poplars over Multiple Years in the Field under Vegetative Propagation

Chapter 1: GENERAL INTRODUCTION

GENETIC ENGINNERING IN FOREST TREES

Genetic engineering (GE), *the use of recombinant DNA and asexual gene transfer methods to alter the structure or expression of specific genes and traits* (FAQ, 2004), has been successfully used to improve crop yields and agricultural productivity. GE crops have been widely adopted by farmers since they first became commercially available in 1996 (reviewed in Fernandez-Cornejo and Caswell, 2006). The global GE crop area increased more than fifty-fold in the first decade of commercialization (James, 2005, ISAAA), and reached 222 million acres grown by 8.5 million farmers in 21 countries in 2005. Herbicide tolerance has consistently been the dominant trait followed by insect resistance and stacked genes for the two traits. In 2005, herbicide tolerance, deployed in soybean, maize, canola and cotton, occupied 71% of the global GE crop acres. *Bt* crops accounted for 18% of the global GE total, and crops with stacked genes contributed 11%.

Transgenic trees have been considerably slower to develop, as a result of their biological recalcitrance, more limited and delayed value, and social and regulatory concerns (Sedjo, 2004). By 2004, forestry GE (also called GM or genetic modification) activities have taken place in at least 35 countries, 16 of which host some form of experimental field trials (reviewed in FAQ, 2004). *Populus* was the first tree to be genetically engineered (1986) and is by far the most commonly studied tree genus for genetic modification purposes (Brunner

et al., 2004; Peña et al., 2001; Strauss et al., 2001). *Pinus*, *Eucalyptus, Liquidambar,* and *Picea* make up the majority of the remaining experimental studies. Field trials of GE trees are largely restricted to those four top genera. To date, only two cases of GE trees have been released commercially. The first one was virus resistant papaya in Hawaii in 1997, and the second one was insect resistant *Populus nigra* in China in 2002.

A number of traits have been studied and tested in tree species (reviewed in FAQ, 2004). Among them, herbicide resistance, insect resistance, and lignin modification have been most strongly pursued and demonstrated to have significant commercialization potential in the near future. Other studied traits include plant development, sterility, phytoremedation, nitrate reductase synthesis, disease resistance, and salt resistance. The recent release of the *Populus trichocarpa* genome sequence, along with recent advances in molecular biology techniques such as gene expression profiling, will dramatically enhance the efficiency of functional and comparative genomics research in trees (Brunner et al., 2004).

Herbicide resistance

Rigorous weed control is required in young tree plantations to ensure successful establishment and maximal development (Meilan et al., 2000a). Herbicide tolerant transgenic trees may make control of weeds in plantations easier and more economical. One strategy is to introduce a mutated gene encoding the enzyme targets for various herbicides, such as *CP4* for glyphosate, an active ingredient of the herbicide Roundup®. A second strategy is to introduce a gene encoding an enzyme that detoxifies herbicides, such as glyphosate oxidoreductase (*GOX*) for glyphosate, and *BAR* for glufosinate. Transformation of *CP4* and *GOX* in hybrid poplar clones has provided stable, commercially

useful levels of herbicide resistance in the field (Meilan et al., 2000a; Meilan et al., 2002b). High levels of resistance were also achieved in transgenic eucalypts (Harcourt et al., 2000), and in *Pinus radiata* and *Picea abies* using a *BAR* gene (Bishop-Hurley *et al.* 2001).

Insect resistance

Control and management of insect pests in tree plantations can be costly and, where pesticides are used, environmentally undesirable. Insect-resistant transgenic trees offer advantages over control of insects by pesticide applications (Meilan et al., 2000b). Most insect-resistant GE trees tested to date contain a *Bacillus thuringiensis* (*Bt*) endotoxin gene. High levels of insect resistance have been demonstrated in *Bt* toxin gene *Cry3A*-containing transgenic poplars (Meilan et al., 2000b; Hu et al., 2001) and eucalypts (Harcourt et al., 2000). Transgenic *Populus nigra* transformed with the *Bt* gene was approved for commercial planting in China in 2002. Other tested genes include anionic peroxidase in sweetgum (Dowd et al., 1998), and a protease inhibitor in poplar (Leplé et al., 1995; Heuchelin et al., 1997).

Lignin modification

Lignin has been one of major targets for genetic engineering in the past decade (Whetten et al., 1998; Baucher et al., 1998; Baucher et al., 2003). Genetic modification of lignin composition to facilitate its removal during chemical pulping can provide both economical and environmental benefits. Expression of several genes involved in lignin biosynthesis has been modified in transgenic poplars (Hu et al., 1999; Franke et al., 2000; Pilate et al., 2002; Baucher et al., 2003; Jouanin et al., 2000). Downregulation of caffeic acid O-methyltransferase

(COMT) by an antisense strategy did not affect lignin content of transgenic poplar lines, but induced a dramatic alteration of lignin structure (Lapierre et al., 1999). Lignin structure was strongly altered in an overexpression transgenic poplar line where COMT activity was close to zero in woody tissues due to induced gene silencing, resulting in higher cellulose content and a higher degree of condensation of the lignin (Jouanin et al., 2000). Such changes positively affected pulp yield, but made lignin less amendable to industrial removal. Poplars with reduced cinnamyl alcohol dehydrogenase (CAD) by expressing an antisense transgene had improved characteristics, allowing easier delignification using smaller amounts of chemicals, while yielding more high-quality pulp (Pilate et al., 2002). Such improved pulping performance in transgenic trees was maintained over four years in the field. Lignin of transgenic poplars overexpressing ferulate 5-hydroxylase (*F5H)* is significantly enriched in S units (Franke *et al*., 2000), and had significantly increased chemical (kraft) pulping efficiency from greenhouse-grown trees (Huntley et al., 2003). Other modified and tested genes include *POX, CCoAOMT*, *CCR*, *LAC*, and *4CL* (summarized in Baucher et al., 2003).

BIOSAFETY OF GE TREES

Although the potential benefits of GE trees appear great, development of transgenic trees is still controversial. The main issue associated with GE trees is possible environmental disruptions resulting from gene flow or gene escape from transgenic plantations (DiFazio et al., 2004). Many tree species, such as poplars, have pollen and seed that can move very large distances, and they are interfertile with wild species that often dominate the ecosystems they occupy. Ecological impacts of transgenic trees will primarily depend on the traits conferred by the transgene and the environment in which the trees occur (Strauss

et al., 1995; DiFazio et al., 2004). Traits that can stably increase fitness might pose a greater impact on wild or managed ecosystem than those that reduce fitness. The former might include insect and herbicide resistance, and the later should include wood modification (lignin content and structure) and reproductive sterility. The longevity of trees and their large size make many kinds of containment, such as physical removal of flowers, virtually impossible (Valenzuela et al., 2006). Therefore, some form of biological confinement might be desirable before some kinds of GE trees are deployed for commercial uses (Strauss et al., 1995; Strauss et al., 2004). Simulation models can be used to estimate levels of gene flow and ecological impacts over long time frames (DiFazio et al., 2004).

CONTROL OF FLOWERING IN TREES

Stability of flowering control mechanisms is a critical biosafety consideration. Most of the current knowledge about flowering comes from the annual plant *Arabidopsis* (Komeda, 2004, Moon et al., 2005). The molecular background of the floral transition in *Arabidopsis* has been well documented (reviewed in Araki, 2001; Battey and Tooke, 2002). Floral homeotic genes can be excellent candidates for genetically engineering sterility because they are often specifically or preferentially expressed in floral tissues. Constitutive expression of the *Arabidopsis* genes *LEAFY (LFY*) and *APETALA1* (*AP1)* reduced the time of vegetative growth in citrus and poplar (Pena et al., 2001; Rottmann et al., 2000; Weigel and Nilsson, 1995). Identification of flowering genes in poplar has been mainly based on sequence conservation and phylogenetic analyses with *Arabidopsis* flowering genes (Brunner et al., 2000; Brunner et al., 2004). The availability of the poplar genome sequence and genome-wide expression analysis has dramatically facilitated the identification of floral genes in poplar.

However, analysis of gene function using transformation is challenging in trees due to their prolonged vegetative stage. Several studies in trees have focused on expression patterns of MADS-box genes, and in more limited cases, gene function has been studied using transgenic methods (e.g., Brunner et al., 2000; Rottmann et al., 2000; Elo et al., 2001; Carlsbecker et al., 2004; Keinonen and Sopanen, 2004; Mohamed, 2006).

Several approaches have been tested for floral modification in poplar (Skinner et al., 2000; Meilan et al. 2001; Strauss et al, 2004). Cell ablation utilizes a floral-specific promoter to direct expression of a cytotoxin that leads to death of floral tissues. Dominant negative mutants suppress the function of a gene at the protein level by overexpression of a mutant version of the protein. RNAi studies of the poplar homolog of the *Arabidopsis* floral repressor *TERMINAL FLOWER 1* showed that strong reduction in target endogene expression caused early flowering in field grown poplars (Mohamed, 2006). Therefore, if stable, RNAi technology combined with early screening of silencing levels may offer a potentially powerful means for floral modification.

STABILITY OF TRANSGENE EXPRESSION

Whether phenotypes of genetically engineered trees can be adequately maintained during their life cycles has been one of the major issues associated with the release of GE trees at a commercial level. Characteristics of integrated transgenes, epigenetic effects, and various stresses (both biotic and abiotic) have been implicated in instability of transgene expression.

T-DNA integration and transgene expression

Although great advances have been made in understanding genes/proteins involved in T-DNA transfer process (Zupan et al., 2000; Gelvin, 2003), the mechanics of integration into host genome remain largely unknown. The prevalent model is that illegitimate recombination occurs between the T-strand and host genome sequence, which requires only "microhomology (< 5 bp)." Therefore, transgenes tend to be more or less randomly integrated into the host genome. Variation in transgene expression levels among independent transformants is often attributed to position effect (reviewed in Kohli et al., 2003). When transgenes are integrated into transcriptionally inert regions, such as repetitive DNA and heterochromatin, transgenes are prone to be silenced. When transgenes are inserted near transcriptional enhancers, expression of transgenes might be elevated. Recent studies suggest that T-DNA may preferentially integrate into transcriptionally active, gene-rich region (Somers and Makarevitch, 2004; Alonso et al., 2003; Chen et al., 2003). But the results of these studies might be biased by marker selection during the transformation process, because genes integrated into transcriptionally inactive regions are prone to be silenced and therefore selected against (Francis and Spiker, 2005).

Multiple copies of T-DNA can be simultaneously integrated into the host genome at single or multiple loci. Direct DNA transfer methods (e.g., electroporation or particle bombardement) often result in the integration of many copies of transgenes (Kohli et al., 1999), while *Agrobacterium*-mediated transformation usually results in fewer copies. There is little or no correlation between transgene copy number and transgene silencing or expression stability (e.g., Hawinks et al, 2003; Gallo-Meagher and Irvine, 1996; Leibbrandt and Synman, 2003). When inserted in the same locus, multiple copies are often arranged as direct repeats, although inverted repeats can also occur (Kumar and Fladung, 2000a). Formation of inverted repeats and associated transgene silencing has often been observed (e.g., Cluster et al., 1996; Morino et al., 1999).

However, Meza et al. (2002) and Lechtenberg et al. (2003) did not find a correlation between inverted repeats and transgene silencing in *Arabidopsis*.

Matrix Attachment Regions (MARs)

Matrix attachment regions (MARs) are operationally defined as DNA elements that bind specifically to the nuclear matrix in vitro (Allen et al., 2000). MARs may be able to islolate tansgenes from the effects of nearby genome sequences, i.e., minimize "position effects." Numerous studies have investigated the relationship between the presence of flanking MARs in transgene constructs and subsequent transgene expression levels and stability. In most cases, MARs result in higher and more stable expression in transformants (e.g., Allen et al., 1993; 1996; Maximova et al., 2003; Han et al., 1997; Peterson et al., 2002), although the degree depends on transformation methods, transgenes studied, and the origins of the MARs. The effects of MARs on variability of transgene expression in populations of independent transformants are less consistent among the different studies (summarized in Allen et al., 2000). MARs can either reduce expression variability or have a very limited or undetectable effect. Most of those studies used a small number of events (4 - 30) for comparing expression and variability between transgenic populations with and without MARs. MARs have also been shown to prevent or reduce gene silencing. MARs reduced the loss of *GUS* expression in transgenic tobacco plants from one generation to the next during sexual reproduction (Ülker et al., 1999). The chicken A1 element was shown to stabilize *GUS* expression over time in both the hemizygous and homozygous plants in three independent transgenic tobacco events (Mlynárová et al., 2003). After removal of the A1 element, the *GUS* gene became silenced over time in two of the three events.

Mechanisms of transgene silencing

Transgene silencing can occur at either transcriptional or post-transcriptional levels. When homologous copies are present most of observed cases of transgene silencing are attributed to homology-dependent gene silencing (Meyer, 1996). When the homology is confined to the coding region, it often leads to post-transcriptional gene silencing (PTGS), and when the homology occurs at promoter sequence, it usually leads to transcriptional gene silencing (TGS).

Although the precise factors that trigger TGS are largely unknown, TGS is often associated with changes in DNA methylation and chromatin structure (Kilby et al., 1992; Kooter et al., 1999; Mette et al., 1999; Baulcombe, 2004). But emerging genetic evidence suggests that changes in DNA methylation might not be as tightly linked to TGS as previously thought (reviewed in Paszkowski and Whitham, 2001). For example, mutations in a novel gene *Morphens molecule* (MOM) released TGS without a change in DNA methylation (Amedeo et al., 2000). PTGS shares similarities with RNAi in animals (Fire et al., 1998), and results in the degradation of homologous RNAs (Vaucheret et al., 2001). TGS and PTGS are functionally linked. RNA-mediated gene silencing pathways are involved in both TGS and PTGS. All pathways involve the cleavage of a dsRNA into short interfering RNAs (siRNAs), which guides degradation or modification of homologous sequences. RNA-directed DNA methylation (RdDM) has been implicated in a type of TGS that is initiated by dsRNAs containing promoter sequences in plants (Cigan et al., 2005; Mette et al., 2000; Aufsatz et al., 2002), and human cells (Kawaskaki and Taira, 2004). RdDM is also assumed to be the source of methylation observed in coding regions in many cases of PTGS (Wang and Waterhouse, 2000; Ebbs et al., 2005). Artificial introduction of dsRNA-producing transgenes in plants has been shown to achieve a high degree and frequency of sequence-specific gene silencing (Chuang and Meyerowitz, 2000; Waterhouse et al., 1998; Stoutjesdijk et al.,

2002; Kerschen et al., 2004; Wagner et al., 2005). When dsRNAs share sequence identity with promoter regions, derived siRNAs can direct DNA methylation to the promoter sequences, leading to TGS. When dsRNA targets coding regions, siRNAs act as guides to degrade homologous sequences.

Stability study of transgene expression

For successful deployment of transgenic trees at commercial levels, genetically engineered traits should maintain their phenotypes until they reach rotation ages. This is particularily important when traits such as sterility and *Bt* insect resistance are deployed. Genomic instability induced by gene transformation, tissue culture, and other biotic and abiotic stress, and subsequent expresson stability have been one of the major concerns in application of GE plants (reviewed in Hoenicka and Fladung, 2006). Meiosis seems to promote transgene loss or inactivation. As a result, the early work on transgene expression stability mainly focused on annual plants to investigate inheritance of expression over sexually propagated generations. More recently studies on the stability of transgene expression in perennial plants under vegetative propagation have recieved increased attention (e.g., Meilan et al., 2000; Cervera et al., 2000; Hawkins, 2003).

Stability of transgene expression under sexual propagation

A number of studies have been conducted on transgene expression in annual plants undergoing sexual propagation. Most of those studies focused on a small number of selected transgenic lines for the purpose of elucidating mechanism and factors triggering transgene silencing (e.g., Lechtenberg et al., 2003; Iglesias et al., 1997; Meyer and Heidmann, 1994; Linn et al., 1990). Several studies

focused on the meiotic stability of gene expression. The frequency of meiotic instability differs greatly among species. Evaluation of transgene expression stability was mainly based on phenotyptic changes, such as antibiotic resistance. High meiotic stability of kanamycin resistance was observed in progeny derived from two tobacco transgenic lines containing a nos-*NPTII* gene (Schműlling and Röhrig, 1995). Loss of kanamycin resistance occurred among backcross progeny from homozygous transgenic lines with frequencies of 1.3×10^{-4} and 5.6×10^{-4} for the two different lines, respectively. Similarly, Conner et al. (1998) reported high meiotic stability of transgene expression in eighteen homozygous single-locus transgenic tobacco lines carrying the *NPTII* gene. More than a million seedlings, were screened for kanamycin resistance, and kanamycin-sensitive seedlings were detected in selfed progeny at a frequency of $0.5 – 5.9 \times 10^{-4}$. However, in the study by Meza et al. (2001), high frequency of transgene silencing was observed in *Arabidopsis* transgenic lines harboring the *NPTII* gene. In the T3 generation, transgene silencing occurred in 56% of the 111 lines. The similar frequency was reported in another study by Scheid et al. (1991), where 50% of the 28 *Arabidopsis* transgenic plants generated with PEG-mediated direct gene transfer failed to transmit the hygromycin resistance phenotype to the progeny.

Stability of transgene expression under vegetative propagation

Transgene expression stability under vegetative propagation has been evaluated in both herbaceous and woody perennial plants. Although instability was reported in several cases, mostly stable expression of transgenes was observed.

Herbaceous perennial plants

Chrysanthemum

In the study by Pavingerová et al. (1997), 17 transgenic lines carrying the *GUS* gene were produced using somatic embryogenesis and *Agrobacterium*-mediated transformation. Transgene expression was monitored in 17 transgenic lines during their growth in the greenhouse, and lines appeared to be stable.

Tall fescue (Festuca arundinacea Schreb.)

Two independent transgenic lines containing a *GUS* reporter gene driven by actin1 promoter were generated via PEG-mediated transformation of protoplasts (Bettany et al., 1998). *GUS* expression was studied over four to five generations of vegetative propagation (tillering) in a containment growth room. Fluorometric assays of *GUS* activity showed that expression was unstable during early generations of tillering, but more stable in the fourth or fifth tiller generation.

Sugarcane (Saccharum spp. Hybrids)

In a sugarcane transgenic line produced via bombardment, expression of the BAR herbicide resistant gene was studied for three rounds of vegetative propagation in the greenhouse (Gallo-Meagher and Irvine, 1996). The transgenic line was estimated to contain approximately five copies of the transgene. Cutting-propagated plants over three generations showed comparable herbicide resistance to the original transformant, and herbicide resistance was also maintained during meristem culture. Leibbrandt and Snyman (2003) studied a transgenic sugarcane line containing nine copies of the PAT herbicide resistance gene over three generations in the field. Herbicide resistance levels determined by herbicide spray were stably maintained during the period of the study.

Woody perennial trees

Citrange (*Citrus sinensis* L. Osbeck x *Poncirus trifoliate* L. Raf)

Expression of a *GUS* gene was studied in 70 independent transgenic lines in a screenhouse over 4-5 years (Cervera et al., 2000). Patterns of expression were comparable for each line in successive histochemical analyses, and no changes in transgene expression were found.

Poplar (Populus spp.)

Expression of a phenotypic marker, *rolC* gene, was studied in poplar transgenic lines under in vitro, greenhouse, and field conditions (Kumar and Fladung, 2000). A stable *rolC* phenotype was observed for 15 aspen hybrid transgenic lines grown in vitro for 5-6 years. However, under the same growth conditions, morphological reversions with incomplete/complete suppression of *rolC* phenotyptic expression were observed in some of the seven transgenic lines produced using wild aspen genotypes. Instability of *rolC* expression was more obvious after transgenic plants were transferred from in vitro conditions to the greenhouse or field. In the greenhouse, three of the 15 hybrid transgenic lines showed incomplete/complete suppression of transgene expression, and three of the five wild aspen transgenic lines showed altered or reverted expression. Among four hybrid transgenic lines planted in the field, one line showed variable transgene expression. Changed expression was also found in one of two field-grown wild aspen transgenic lines. The study also demonstrated that variable *rolC* expression was associated with the formation of direct repeats of the transgene. However, inactivation of *rolC* was also observed in single-copy transgenic lines. The observed variation in *rolC* expression might be attributed to its toxic effect on plant development, where even minor variations in expression can cause profound developmental changes.

In another study using a larger number of transgenic lines, Melian et al. (2002) reported high stability of herbicide resistance and *GUS* expression in 40 hybrid cottonwood (*P. trichocarpa* x *P. deltoids*) transgenic lines over four

years in the field. Only one of the 40 lines studied showed inadequate transgene stability.

Similarly, high stability of transgene expression in poplar was reported in the study by Hawkins et al. (2003). Expression of a *GUS* reporter gene driven by either the 35S promoter or a vascular-specific promoter (EuCAD) was evaluated in transgenic hybrid poplars (*P. tremula* x *P. alba*) under different growth conditions (Hawkins et al., 2003). In the first group, four transgenic lines containing the 35S-*uidA* gene were evaluated in the field for a period of six years. *GUS* activity was measured using a fluorometric assay and showed stable expression, which was also supported by histochemical assay. In the second group, 20 transgenic lines containing the 35S-*uidA* gene, and 20 transgenic lines containing the EuCAD-*uidA* gene were studied for transgene expression under in vitro conditions. From those lines, nine 35S-uidA lines and nine EuCAD-*uidA* lines were selected and transferred to the greenhouse, and subsequently to the field. Transgene expression was regularly monitored for those 18 lines under greenhouse and field conditions. Stable expression was observed under in vitro condition. The level of transgene expression appeared to be reduced in some transgenic lines upon transfer to greenhouse conditions, but the transgene continued to be expressed when grown in field conditions. No case of gene silencing was observed. To investigate the effect of stress on transgene expression stability, two 35S-*uidA* lines and two EuCAD-*uidA* lines were grown for a period of six months on medium containing different concentrations of propionic acid. The stress treatment had little effect on transgene activity. The authors also explored the relationship between copy number and transgene activity, and found either little or no correlation.

Stress-induced transgene silencing

Biotic and abiotic stress can cause genome instability, which might change gene expression of both endogenous genes and transgenes (reviewed in Hoenicka and Fladung, 2006). Heat-induced inactivation of foreign genes in transgenic plants has been reported in several studies (reviewed in Broer, 1996). Single-cell suspension cultures of a *Medicago sativa* transgenic line frequently lost herbicide resistance confered by a *PAT* gene under a high temperature (37^0C) (Walter et al., 1992). Up to 12% of the suspension culture cells grown at a constant temperature of 25^0C lost herbicide resistance within 150 days, while the frequency went up to 95% under a heat treatment of 37^0C for 10 days. Heat-induced transgene silencing was also observed in whole plants. Neumann et al. (1997) reported that heat treatment (37^0C) of transgenic tobacco (*Nicotiana tabacum*) plants led to a reversible reduction or complete loss of transgene-encoded activities in about 40% of the 10 independent transformants carrying the *LUC* and *NPTII* genes, whereas the other lines had temperature-tolerant gene expression. Temperature sensitivity or tolerance of transgene-encoded activities was heritable. In some of the lines, temperature sensitivity of the transgene-encoded activities depended on the stage of development, occurring in either seedlings (40% luciferase and 50% neomycin phosphotransferase) or adult plants (both 40%). The phenomenon did not correlate with copy numbers or the homo- or hemizygous state of the transgenes. In another study by Köhne et al. (1998) heat- induced transgene inactivation was influenced by the transgene sequence. The herbicide resistance encoded by the GC-rich *pat41* gene driven by a 35S promoter was strongly reduced in all of the 27 independent transgenic tobacco lines. In contrast, the expression of the AT rich synthetic *patS* coding for the same protein driven by the 35S promoter was stable under heat treatment.

Several studies also reported that drought affected transgene expression (summarized in Sousa-Majer et al., 2004). Studies on transgenic maize (Zea

mays L.) resistant to European corn borer (Ostrinia nibilalis L.) showed that water stress affects the level of *Bt* proteins (Traore et al., 2000). The effects of water deficit and high temperature on the activity of a-amylase inhibitor 1 (a-AI-1) were studied in transgenic peas (*Pisum sativum L.*) that were developed to control the seed-feeding pea weevil (Sousa-Majer et al., 2004). Water stress imposed during the seedling stage increased the level of a-AI-1 in seeds on a dry weight basis, while high temperatures imposed during the seedling stage reduced the level of a-AI-1 on average by 36.3% in transgenic peas, allowing 39% of adult's pea weevil to emerge compared to 1.2% in the transgenic peas grown at 27/22 °C. Transgenic petunia (*Petunia hydrida*) plants carrying a single copy of the maize A1 gene encoding a dihydroflavonol reductase were planted in the field in Germany. Transgene inactivation was observed in about 60% of the plants after a period of high light intensity and temperatures up to 36 °C (Meyer et al., 1992).

The various stresses that occur during plant transformation and tissue culture process can cause genetic instability, including cytosine methylation, structural rearrangements, and repeat-induced point mutation (reviewed in Phillips et al., 1994). Increased chromosomal variation was observed in transgenic barley plants (Choi et al., 2000). Out of 59 independent transgenic lines, 27 (46%) were tetraploid or aneuploid, while non-transgenic plants regenerated by in vitro culture had a much lower percentage of tetraploids (0–4.3%). However, an AFLP study of four transgenic aspen lines showed high genomic stability (reviewed in Hoenicka and Fladung, 2006). Out of 889 AFLP-bands, 886 were common in the non-transgenic plants and in the four transgenic aspen lines. Different taxa may be more or less susceptible to stress-induced genome change. The stress imposed by tissue culture conditions tends to induce genomic instability at particular loci, and these may be preferential targets for the integration of foreign DNA (Gould 1986; Romano et al., 2005). Transgene

expression could therefore change as a consequence of plant tissue culture and regeneration.

Stability of RNAi suppression

Although numerous studies have investigated RNAi gene silencing mechanisms and pathways, information about stability of RNAi suppression is very limited. In *Arabidopsis*, the endogenous Δ12-desaturase gene (*FAD2*) was targeted for silencing using seed-specific cosuppression (CS), hairpin RNA (hpRNA), and intron-spliced hairpin (ihp) constructs (Stoutjesdijk et al., 2002). One highly silenced ihp line was propagated for five generations and showed no reversion or reduction in its degree of silencing. Several studies have reported that temperature affects the level of RNAi silencing in both plants and animals (Szittya et al., 2003; Sós-Hegedűs et al., 2005, Fortier and Belote, 2000, Kameda et al., 2004). At low temperature (15 ^{0}C), both virus and transgene triggered gene silencing were inhibited in *Nicotiana benthamiana* (Szittya et al., 2003). When protoplasts were infected with Cymbidium ringspot virus (CymRSV) at different temperatures (15, 21, 24, and 27 ^{0}C), it was found that virus-derived siRNA were abundant at 27 ^{0}C, but not detectable at 15 ^{0}C. Agroinfiltration of wild type or *GFP* expressing *N. benthaminana* plants with 35S-*GFP* or 35S-ds*GFP* at different temperatures also showed reduced siRNA level and increased mGFP levels at low temperature. In the same study, temperature-dependent RNAi silencing was observed in transgenic *N. benthaminana* plants expressing a CrmRSV-derived RNA as well as antisense-mediated endogene inactivation in *Arapidopsis* and potato. Sós-Hegedűs et al. (2005) reported that temperature-dependence varied among different antisense transgenic lines targeted for the expression of endogenes StubGAL83, PKIN1 or,

StubSNF1, and both PKIN1 and StubSNF1. Nine of 24 lines showed temperature-insensitive RNAi silencing.

Those studies suggest that RNAi silencing induced by dsRNAs may be unstable in field environments. Nonetheless, performance of four transgenic poplars with altered lignin by CAD or COMT antisense transgenes was stably maintained in the field at two different sites over four years (Pilate et al., 2002). To our knowledge, no study of stability of dsRNA induced gene suppression has been reported in trees over years in the field environment. Until such studies are conducted, it is difficult to predict the success of RNAi-induced commercial (e.g., lignin) and biosafety (e.g., sterility, dwarfism) traits. The goal of this study was to add to the very modest literature on stability of gene expression and RNAi in trees. In particular, the study was designed to be as relevant as possible to commercial GE programs in trees by using primary transformants, screening a large number of events, studying events in a field environment, and focusing on vegetative growth and propagation. This study includes many more plants and events than any other published study of stability of transgene expression in a vegetatively propagated plant.

The main questions we sought to answer are:

1) What proportion of plants and events show unstable transgene expression or RNAi suppression?
2) Does organogenesis promote unstable gene expression or physical stability?
3) Does a field environment, pruning, and dormancy cycles promote instability compared to in vitro/greenhouse expression?
4) Do matrix attachment regions promote gene expression levels or stability among events or plants?
5) Does a native vs strong heterogonous promoter influence stability of expression?

6) Are transgene copy number, structure, or methylation associated with expression level or instability?

Chapter 2: HIGH LEVEL OF TRANGENE EXPRESSION STABILITY IN FIELD-GROWN POPLAR

ABSTRACT

High stability of transgene expression is essential for application of genetic engineering to commercial forestry. We studied the stability of expression of two reporter genes, green fluorescent protein (*GFP*) and the herbicide resistance gene *BAR*, in 2,256 transgenic poplar trees derived from 404 primary events and 106 subevents over three years in the greenhouse and field after vegetative propagation. The effects of matrix attachment regions (MARs), transgene copy number, and repeat formation in multiple insertion events on expression level, variation, and stability were examined. No gene silencing (complete breakdown of expression) was observed for *GFP* or *BAR* expression in any of the primary transgenic events over three years of study. Approximately 1 % of the transgenic events showed reduced *GFP* expression over time, but none of events showed reduced *BAR* expression. Transgenic cassettes were physically eliminated in four subevents (2.5 %) derived from three different primary events during re-organogenesis. Three subevents (2 %) from two primary events showed significantly elevated *GFP* expression, and this change was stable over time. MARs did not increase mean transgene expression level or expression stability, but did reduce variance of expression levels among events. MARs also increased coordinated expression of *GFP* and *BAR* assembled in the same T-DNA. A majority of transformants (85%) carried single copy transgenes. Transgene copy number was positively correlated with transgene expression level, but not with stability of expression. Approximately 38 % of the events containing two-copy inserts had repeats formed at the same chromosomal position, with direct repeats (87%) as the major type of repeat observed. All events containing more than two

copies of transgenes had repeat formation at a single locus, and direct repeats were predominant (77%). Loci with two direct repeats had substantially greater transgene expression level than other types of two copy T-DNA configurations, but did not affect stability of transgene expression. The use of a homologous promoter, the poplar *rbc*S promoter, which drove *BAR* in the transgenic constructs, did not adversely affect transgene expression levels or 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 (reviewed in Stam et al., 1997; Kooter et al., 1999; Fagard and Vaucheret, 2000). Although the rates and causes of instability vary widely among species, environments, and transformation systems, the detection and removal of unstable transgenic events are essential steps for commercial development (Meza et al., 2001; Kohli et al., 2003).

Unstable expression of transgenes are of particular concern for trees as a result of their long life cycles, difficult transformation, and the general absence of sexual propagation prior to field and commercial deployment (Bradshaw et al., 2000; Han et al., 1997). Meiosis is widely known to promote gene silencing, allowing unstable transgenic lines to be rapidly recognized and removed from further deployment (Scheid et al., 1999; Metz et al., 1997; Iglesias 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). High levels of stability may therefore be dictated by social as well as commercial needs.

Expression of a transferred gene can be initially silenced or inactivated over time or generations by either transcriptional gene silencing (TGS) or post-

transcriptional gene silencing (PTGS) triggered by diverse host defense responses (Matzke et al., 2000). TGS is often associated with the methylation of the promoter region (Kilby et al., 1992), but emerging genetic evidence suggests that changes in methylation might not be as tightly linked to TGS as previously anticipated (reviewed in Paszkowski and Whitham, 2001). For example, mutations in a novel gene *Morphens molecule* (MOM) released TGS without change in DNA methylation (Amedeo et al., 2000). PTGS involves degradation of RNAs directed by short inference RNAs (siRNAs) produced from double stranded RNAs (Baulcombe, 2004). Recent studies support the idea that two processes are mechanismly and functionally related because they are correlated with some of the same events, including DNA methylation and RNA mediated silencing pathways (Paszkowski and Whitham, 2001; Bender, 2001; Baulcombe, 2004).

Several factors that might serve as the sources triggering those silencing mechanisms have gained a great deal of attention in studies of transgene silencing. Those include transgene copy number, T-DNA structure, and integration sites. Presence of multiple copies of transgene transcripts might cause homology-dependent transgene silencing at TGS or PTGS levels. Different studies on the correlation between transgene silencing and copy number have shown conflicting results; and in most cases no correlation was observed (Hawkins et al., 2003; Leibbrandt and Synman, 2003; Meza et al., 2001; Gallo-Meagher and Irvine, 1996; Scheid et al., 1991; Kohli et al., 1999; Iglesias et al., 1997). Formation of transgene repeats and subsequent effects on stability have been reported in both annual plants and trees (reviewed in Stam et al., 1997; Kumar and Fladung, 2000a; Lechtenberg et al., 2003). Transgenes that are organized as inverted repeats can show low or silenced expression (Morino et al., 1999), but this is not always the case (Meza et al., 2002; Lechtenberg et al., 2003)

Integration sites can vary considerably among independent transformants, and have a profound effect on the level and stability of transgene expression (reviewed in Kohli et al., 2003). Early studies suggested that T-DNA integration is a random process, but recent studies have shown that transgenes tend to be integrated into gene-rich regions in plants (Somers and Makarevitch, 2004; Alonso et al., 2003; Chen et al., 2003). However, Francis and Spiker (2005) reported that the integration is considerably more random when transgenic plants were not identified by marker selection. It is believed that transgenes integrated in or near repetitive DNA or heterochromatin are prone to silencing, and are therefore selected against during transformation.

Matrix Attachment Regions (MARs) may act as boundary elements to reduce the influence of nearby host sequences on transgene expression (Allen et al., 1993). The effect of MARs from different sources on expression levels and stability has been studied in different species, including poplars (reviewed in Allen et al., 2000; Butaye, 2005; Maximova et al., 2003). Most reports agree that MARs increase transgene expression, but such effects are far less prominent in plants transformed with *Agrobacterium* vs biolistics (Allen et al., 1993; 1996; 2000; Maximova et al., 2003; Schöffl et al., 1993; Mlynárová et al., 1994; 1995; van der Geest et al., 1995; Han et al., 1997; Vain et al., 1999; Ülker et al., 1999; Levee, 1999; Peterson et al., 2002). Those studies showed that MARs can either reduce expression variability or have no, or a limited effect. Butaye et al. (2004) reported that MARs from a chicken lysozyme gene caused significant elevation of *GUS* activity in transformants in a silencing mutant background, while no elevation was observed in the transformants in the wild type background. The effects of MARs on transgene silencing have also been investigated in several studies (Maximova et al., 2003; Brouwer et al., 2002; Vain et al., 1999; Ülker et al., 1999; Conner et al., 1998). Most of those studies showed that the presence of MARs tended to reduce or prevent the occurrence of transgene silencing.

Conner et al. (1998), however, reported that MARs had no effect on meiotic instability in kanamycin resistance in *Arabidopsis*.

Stress-induced transgene inactivation has been reported in plants (reviewed in Broer, 1996). A heat treatment (37 ^{0}C) lasting for 10 days resulted in an almost complete (95%) loss of the phosphinothricin resistance in supension culture cells derived from a single *Medicago sativa* transgenic line (Walter et al., 1992). When transgenic tobacco plants carrying the *LUC* and *NPTII* genes were subjected to a heat treatment, 40% of 10 independent transgenic lines showed a reversible reduction or complete loss of transgene expression. In a field environment, 60% of transgenic petunia (*Petunia hydrida*) plants carrying a single copy of the maize A1 element showed phenotypic alterations after a period of high light intensity and temperatures up to 36 ^{0}C (Meyer 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 *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 plants. Out Of 59 independent transgenic lines, 27 (46%) were tetraploid or aneuploid around the tetraploid level, while non-transgenic plants regenerated after in vitro culture alone had a much lower percentage of tetraploids (0 - 4.3%). However, the AFLP study of four aspen transgenic lines and wild type plants showed high genomic stability (reviewed in Hoenicka and Fladung, 2006).

Transgenic instability was frequently observed in annual plants undergoing sexual propagation (Kilby et al., 1992; Scheid et al., 1991; Meza et al., 2001; Kohli et al., 1999; Morino et al., 1999; Müller et al., 1987; Iglesias et al., 1997; Meyer et al., 1992). More recently, several studies have been conducted in perennial plants, including poplar trees over multiple seasons

under vegetative propagation (Gallo-Meagher and Irvine, 1996; Bettany et al., 1998; Kumar and Fladung, 2001; Meilan et al., 2002a; Hawkins et al., 2003; Leibbrandt and Synman, 2003; Cervera et al., 2000). Most of those studies have reported that transgene expression under vegetative propagation is highly stable and predictable. For instance, Meilan et al. (2002a) reported high stability of herbicide resistance in 40 poplar transgenic lines over four years in the field. Long-term expression stability of the *GUS* gene was also demonstrated in another study in poplar (Hawkins et al., 2003). Four transgenic lines carrying 35S-*GUS* maintained stable expression over six years in the field. Another 18 lines carrying either 35S-*GUS* or EuCAD-*GUS* were evaluated under in vitro, greenhouse, and field environments, and no case of gene silencing was observed. A similar observation was reported by Cervera et al. (2000) in a study in which the expression of a *GUS* gene was assessed in 70 independent citrus transgenic lines over 4-5 years in the screenhouse. However, phenotypic instability of a *rolC* gene was observed in some of transgenic aspen lines grown in vitro, greenhouse or field (Kumar and Fladung, 2001). Instability of *rolC* expression was more obvious after transfer from in vitro to greenhouse.

Here we report a study of stability of transgene expression in trees using a large number of transgenic events. Expression of 404 poplar transgenic events containing two reporter genes, *GFP* and *BAR*, were studied under vegetative propagation in the greenhouse and field over three years. We investigated effects of several factors on transgene expression levels, variability, and stability, including the effects of secondary organogenesis, flanking MARs, transgene copy number, and T-DNA structure. We report extremely low level of instability with respect to all of these factors.

MATERIALS AND METHODS

Construction of selectable marker and reporter gene cassettes

We chose two reporter genes based primarily on economy and speed of assay. We used a modified *GFP* gene (mgfp5er, provided by C.N. Stewart, University of South Carolina) that was altered to increase sensitivity to blue light, maximize expression and proper processing in plants, and target the encoded protein to the endoplasmic reticulum (Haseloff et al., 1997). Expression of *GFP* gene can be visualized when illuminated by a hand-held UV lamp or blue light in the field (Harper et al., 1999), and later on it was shown that fluorescence levels can be easily quantified with a hand-held GFP-meter (Millwood et al., 2003), which allows more precise evaluation of expression variation. The *BAR* gene from *Streptomyces hygroscopicus* encodes phosphinothricin acetyltransferase and confers resistance to the herbicide glufosinate that inhibits glutamine synthetase (Riemenschneider, 1997). Our initial expectation was that silencing of the *BAR* gene could be easily screened in the field by visualizing leaf damage after leaf painting with glufosinate herbicide, which acts in a non-systemic manner. But we chose not to apply herbicide as silencing was very rare and preliminary leaf painting experiment showed that it was not sufficiently repeatable (data not shown), and can be greatly affected by leaf age and structure.

Because gene silencing may be very different with a native versus a foreign promoter, we used the cauliflower mosaic virus 35S promoter and the promoter from the poplar gene for the small subunit of ribulose bisphosphate caboxylase (*rbc*S) promoter to drive *GFP* and *BAR*, respectively. The 35S promoter had been used widely in transgenic poplars and is known to give high levels of expression in leaves as well as in many other plant organs. It has also been shown to drive high level of *GFP* expression in tobacco (Harper et al.,

2000). An *Arabidopsis rbc*S promoter fused to the *BAR* gene has been shown to confer high levels of herbicide resistance in transgenic poplars in our laboratory (unpubl. data). We expected the poplar *rbc*S promoter to behave similarly due to the high conservation of regulatory properties among *rbc*S promoters from plants (Argüello-Astora and Herrera-Estrella, 1996).

The binary vector pGreenII (Hellens et al., 2000) was used for assembling the selectable marker *NPTII* gene and reporter genes *GFP* and *BAR* (Fig. 2.1). An *Asc*I linker was inserted at *Hpa*I and *Stu*I sites of the pGreenII backbone to produce pG3. Two versions of the vector were used: one with flanking MAR elements derived from tobacco RB7 gene (provided by S. Spiker, North Carolina State University), and one without MARs. For cloning of MARs, 1,167 bp MAR fragment was removed from the vector pHK10 with the restriction enzymes *Not*I and *Spe*I, blunted, and cloned at *Fsp*I and *Sap*I (blunted) sites of the pG3 to produce pG3M.

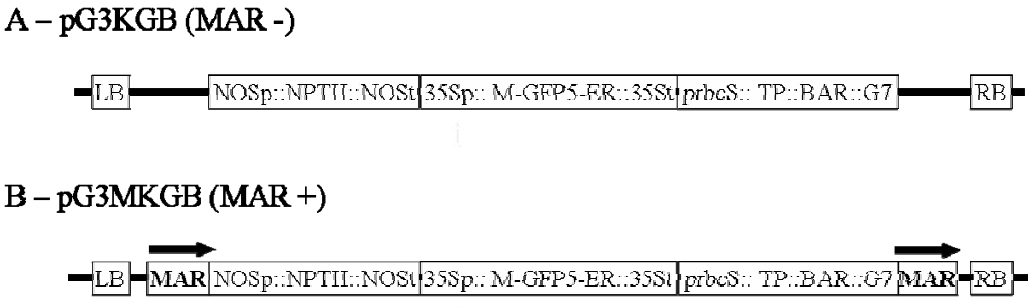


Fig. 2. 1. Schematic maps of the T-DNA region of binary vectors.
Schematic maps of the T-DNA region of binary vectors. Both constructs contained backbone and border sequence from the pGreenII binary vector, selection marker *NPTII* gene driven by NOS promoter, *GFP* gene driven by 35S promoter, and *BAR* gene driven by poplar *rbc*S promoter. A: Construct pG3KGB without flanking MARs. B: construct pG3MKGB with flanking MARs. Direction of MAR fragment is indicated by arrow. The T-DNA regions are not drawn to scale.

The selectable marker cassette consists of the promoter from the nopaline synthase gene (*NOS*), the coding region of the *NPTII* gene conferring resistance to the antibiotic kanamycin, and the 3' untranslated end from the *NOS* gene. For cloning of the cassette, an 869 bp coding region of the *NPT II* gene was cut from pJIT134 with *Pst*I and *Xba*I, and cloned into the corresponding *Pst*I and *Xba*I sites of the *NOS* promoter cassette from the pGreenII system. The whole cassette (1,396 bp) was then cut with *Eco*RV, and cloned into the blunted *Xho*I site of pG3 and pG3M to produce pG3K and pG3MK, respectively.

The *GFP* cassette contains the 35S promoter and terminator from cauliflower mosaic virus, and the *mGFP5er* coding sequence The 818 bp m*GFP*5er coding sequence was cut from pBIN m*GFP*-5-ERer with *Bam*HI and *Sac*I, and cloned into the corresponding *Bam*HI and *Sac*I sites of the 35S promoter cassette from the pGreenII system (Hellens et al., 2000) . The whole cassette was then cut with *Eco*RV, and cloned into the *Eco*RV site of pG3K to produce pG3KG.

The *BAR* cassette consists of the poplar *rbcS* promoter, the *BAR* coding sequence, and the terminator sequence from the Agrobacterium g7 gene. The poplar *rbc*S promoter was cloned from *Populus trichocarpa* genomic DNA using the GenomicWalker method (Universal Genomic Walker kit, Cat # K1807-1, Clontech) following the manufacturer's instructions. Gene-specific primers were designed from the poplar *rbc*S cDNA sequence (provided by J. Davis, University of Florida, Gainesville). The gene-specific primers for the primary and secondary PCR amplifications were: 5' GCATGCATTGAACTC GTCCACCATTGC3', and 5' ATGTCATTAGCCTTTCTGGTACTGGCT3', respectively. A 1,300 bp fragment was amplified from the secondary PCR from the *Pvu*II library at an annealing temperature of 72 ^{0}C. The amplified promoter fragment was then cloned into the pGEM Easy vector (Promega, Madison, WI), and sequenced from both ends. A 1,100 bp promoter fragment was removed

with *Acc*I and *Hind*III, and cloned into the corresponding sites of the intermediate vector pBluscript KS (Stratagene, La Jolla, CA). An 971 bp fragment encoding the *Arabidopsis rbc*S transit peptide, the *BAR* coding sequence, and the g7 terminator was amplified from the vector pTTM8 (Plant Genetic System, Belgium) with the primers 5'GTCTGCAGGAACAATGG CTTCCTCTATG 3', and 5' AGACTAGTG ATGTTAATTCCCATCTTG 3'. The amplified fragment was digested with *Spe*I and *Pst*I, and cloned into the corresponding sites of pBluscript KS containing the poplar *rbc*S promoter. The entire *BAR* cassette was then cut from the intermediate vector with *Xho*I (blunted) and *Spe*I, and cloned into *Sma*I and *Spe*I sites of the pG3KG to produce pG3KGB (Fig. 2.1A). The fragment containing the *BAR* and *GFP* cassettes was removed from pG3KGB with *Spe*I and *Sal*I (partial digestion due to an internal *Sal*I site in the *BAR* cassette), and then cloned into the corresponding sites of pG3MK to produce pG3MKGB (Fig. 2.1B).

Plant transformation and organogenesis treatments

Two hybrid poplar clones 353-53 (*P. tremula x P. tremuloides*) and 717-1B4 (*P. tremula* x *P. alba*) were transformed using *Agrobacterium tumefaciens* strain C58/pMP90 (GV3101) harboring the transgenic constructs pG3KGB and pG3MKGB following an established protocol based on the *NPTII* gene as described previously (Filichkin et al., 2006). All transformants were subsequently grown on herbicide-containing medium to confirm the incorporation of a functional *BAR* transgene. For each of the four clone and construct combinations, approximately 100 transgenic events were produced, for a total of 404 independent transgenic events (Appendix T2.1).

A total of 80 primary transformants, 20 per clone x construct

combination were randomly selected for secondary organogenesis. Leaf discs from those selected events were reintroduced into the tissue culture process following the similar tissue culture protocols used to produce the transgenic events, except that the leaf explants were not co-cultivated with *Agrobacterium* nor was kanamycin or basta added to the culture medium (Appendix F2.1). 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 x 4 (clone x construction) x 2 subevents] were regenerated from 80 primary events.

Six to eight ramets for all transgenic events and subevents, plus 32 non-transgenic (NT) plants for each of two clones, were propagated in vitro, potted in soil, and grown in the greenhouse for three to four months. Together with NT plants, four ramets from the individual transgenic events and subevents were planted at a field site near Corvallis, Oregon, in October 2003 (Fig. 2.2).

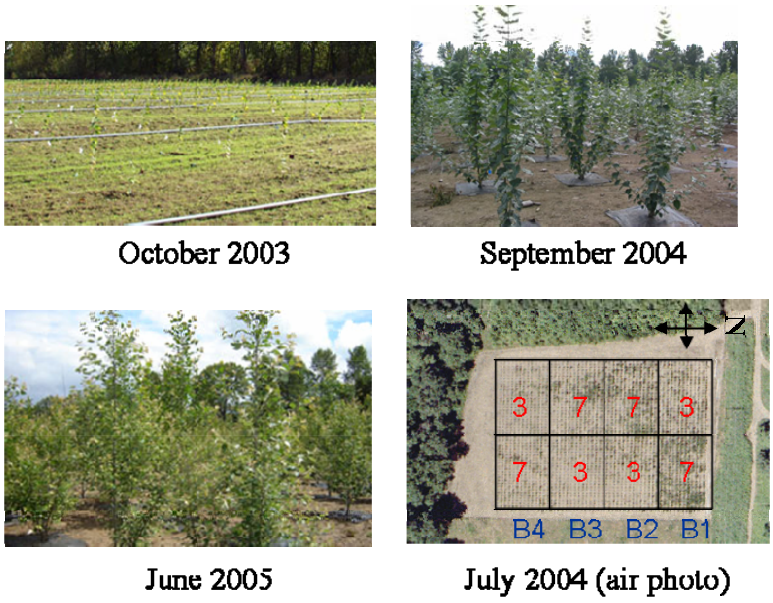


Fig. 2. 2. Field trial of reporter gene stability study. The transgenic poplars were planted at a field site near Corvallis, Oregon, USA, in October 2003. In the air photo, B1- B4 represent four different blocks; 3: 353-53 transgenic events; 7: 717-B4 transgenic events.

The plants were distributed in four blocks with a random split-plot design (Appendix T2.2), where two clones were separated and randomly assigned to subplots of each block, and transgenic events from one of two clones were randomly distributed within corresponding subplots. The plants were coppiced in spring 2004 to stimulate growth of multiple shoots, a common practice for poplar.

Measurement of transgene expression levels

GFP quantification

GFP expression was quantified with a hand-held *GFP*-Meter (Opti-Sciences, Hudson, NH; http://www.optisci.com; Fig. 2.3A). The GFP-Meter is a self-contained, hand-held fluorometer that utilizes a modulated system to compensate for light stray and temperature drift (more detailed descriptions see Millwood et al., 2003). *GFP* quantification was made on intact leaf samples and surfaces. When powered on, the meter generates an excitation light that travels through a bandpass filter to a fiber-optic cable, and then is delivered to the leaf sample through an attached leaf clip. The light emitted from the leaf sample enters back to the optic cable and is directed through a bandpass filter into a low noise preamplifier. The fluorescence signal is processed and displayed in units of counts per seconds.

For each leaf sample, two measurements were taken at two spots located on both sides of the mid-vein (Fig. 2.3, B&C). When the same optic cable was used, fluorescence levels of transgenic plants measured at different times were scaled to background fluorescence levels of non-transgenic controls.

Fig. 2. 3. Measurement of *GFP* fluorescence with a GFP-meter (A) in the greenhouse (B) and field (C). White spots on the leaves (B & C) are approximate spots that were measured for *GFP* expression. Red arrows indicate two leaves from two shoots of a plant sampled for fluorescence measurements.

BAR quantification

Expression of *BAR* transgene was quantified with the Enzyme-Linked Immunosorbent Assay (ELISA). For total protein extraction, about 50 mg leaf tissue was ground in 400 µl extraction buffer (50mM NaHPO4, PH7.0; 10mM EDTA) in a 1.5 ml microtube with a disposal pestle and grinder. The samples were centrifuged for 15 minutes at 4 ^{0}C. Approximately 200 µl of supernatant was transferred to a new 1.5 ml tube, freshly frozen in liquid nitrogen, and stored under -80 ^{0}C until assayed. The total protein concentration was measured

with the Bradford method (Bio-Rad Protein Assay Kit, Hercules, CA; Cat # 500-0001) using a microtiter plate reader (Molecular Devices, Sunnyvale, CA) following the instructions provided by the manufacturers. Triple and duplicate reactions were used for BSA standards and experimental samples, respectively.

The relative concentration of the phosphinothricin acetyl transferase (PAT) encoded by *BAR* was quantified using the commercial LibertyLink PAT/*BAR* ELISA kit (Envirologix Inc, Portland, Maine; Cat # AP013) according to "HIGH SENSITIVITY PROTOCOL". Non-transgenic controls gave the same background levels as blanks (protein extraction buffer without any samples added), thus, only blanks were included in each of assayed plates. A reference sample was run in all assayed plates for data normalization, and duplicated wells were used for all tested samples. The OD (optical density) was determined at a wavelength of 450 nanometers (nm) 20 minutes after adding the stop solution to the tested wells. The mean OD from the blank wells was subtracted from all samples.

Preliminary *GFP* expression study

Preliminary *GFP* expression studies were performed on a subset of transgenic poplars grown in in-vitro, the greenhouse, and the field to determine the best sampling method.

GFP expression variation within a leaf

GFP expression variation within a leaf was studied in 8 transgenic events and two non-transgenic plants grown in the greenhouse for approximately four months. The fourth easily observed leaf from the apex was sampled for *GFP* measurement (Appendix F2.2).

GFP expression variation within a plant

Vertical *GFP* expression variation within a plant was studied with plants grown in three different environments: in-vitro, greenhouse, and field. *GFP* measurement was made on two one-month old, in-vitro grown transgenic plants in 2003, six four-month old, greenhouse grown transgenic and NT plants in 2003, and four two-year old, field grown transgenic and NT plants in 2004.

GFP expression variation over time (daily)

GFP measurements were made on plants grown in the greenhouse and the field. Together with two NT plants, 19 greenhouse grown transgenic events (approximately four-months old) were measured in August 2003. The fourth or fifth leaf from the apex was measured for *GFP* expression at three different time points of a single day. Leaves of four transgenic and NT plants that were studied for within plant variation as discussed above were also measured at four different time points to simultaneously investigate daily *GFP* expression changes.

GFP expression measurements for stability study

The expression levels of *GFP* and *BAR* were measured in three different years: 2003 in the greenhouse, and 2004 and 2005 in the field. Measurements in the greenhouse were taken before plants were planted in the field in October 2003. Based on the preliminary expression study results (see results section), the fourth and fifth leaves from the top of each plant were sampled for the measurement of *GFP* expression (Fig. 2.3B). For the field plants, the leaf

immediately above the first fully expanded leaf from two different shoots was used for the measurement (Fig. 2.3C).

Leaf samples measured for *GFP* expression were taken and stored at -80^{0}C until the ELISA assay was performed for quantification of *BAR* expression. One ramet for each of the events and subevents was used for the PAT assay.

Comparative real-time PCR for transgene copy number

To estimate transgene copy number in transgenic plants, comparative real-time PCR was performed on 396 transgenic events. The endogenous, single copy gene *PTLF* (poplar *LEAFY/FLORICAULA*) was used as reference gene (Rottmann et al., 2000). The transgene *GFP* was chosen for the copy number analysis as it was closer to the post-transferred left boarder of the T-DNA than the *BAR* gene. Real-time PCR was performed using dual-labeled TaqMan® probes (Biosearch Technologies, Novato, CA). All primers and probes were designed using the program Primer 3 (Rozen and Skaletsky, 2000; http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). For the *GFP* transgene, the primers used for the amplification were: 5' TTAAGGGAATCGATTTCAAG 3', and 5'ACGTTGTGGGAGTTGTAGTT 3'. The hybridization probe was dual labeled with HEX and BHQ-1: 5' HEX d(CCTCGGCCACAAGTTGGAATAC) BHQ-1 3'. For the *PTLF* gene, PCR amplification primers were: 5' GGTTTCT CTGAGGAGCCAGTACAG 3', and 5' GCCTCCCATGTCCCTCTTC 3'. The hybridization probe was labeled with FAM and TAMARA: 5' FAM d(CAAGGAGGCAGCAGGGAGCGGT) TAMARA 3'. The primer and probe sets were tested for their amplification efficiency using two-fold serial standard dilutions of a transgenic sample, and were optimized to have comparable efficiency (Appendix F2.3). For the *PTLF* gene, the optimized concentrations for both forward and reverse primers were

0.4 μM, and the probe was 0.2 μM. For the *GFP* gene, optimal primer concentration was 0.4 μM for both primers, and the probe was 0.3 μM.

All PCR reactions were carried out using QuantiTect Multiplex PCR buffer (Qiagen, Valencia, CA, USA; Cat # 204545) in a volume of 25 μl. For each sample, 100 ng genomic DNA purified with the Dneasy Plant Mini Kit (Qiagen, Valencia, CA; Cat # 69106) 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 Mx3000 P real-time PCR machine (Stratagene, La Jolla, CA) with the following cycles: 95 ^{0}C for 15 min, and 45 cycles of: 94 ^{0}C for 30 s, and 60 ^{0}C for 1 min. The threshold cycle (Ct) was determined using the MX3000P$_{TM}$ RT-PCR System software (version 2). Transgene copy number of *GFP* transgene was determined by the formula: $2^{(1-\Delta Ct\ (GFP\text{-ptlf})}$.

Southern blot to estimate transgene copy number

Southern blot analysis was performed to validate the real-time PCR method. Probes were labeled using PCR amplification with DIG- 11- dUTP (Digoxigenin- 11- 2'-deoxy-uridine-5'-triphosphate, alkali-labile; Roche, Indianapolis, IN; Cat # 11573152910, 25 μl). A 494 bp *GFP* probe fragment was amplified and labeled using the primers: 5' TGGCCAACACTTGTCA CTAC 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 plasmid DNA of the transgene construct 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' GCCAGTACAG CAAGACAAGG 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 cycles used for labeling were: 95 ^{0}C for 2 min; 30 cycles of: 95 ^{0}C – 30 s, 56 ^{0}C – 30 s, and 72 ^{0}C – 40 s; 72 ^{0}C for 7 min.

A total of 16 µg genomic DNA extracted from each transgenic sample was digested with *Hind*III and *Sca*I, respectively. Both enzymes cut within the T-DNA but outside of the *GFP* probe. The digested genomic DNA was electrophoresed along with 25 ng Digoxigenni-labeled DNA molecular weight marker II (Roche, Indianapolis, IN; Cat # 11218590910) on a 0.8% agarose gel (16 cm x 14 cm) for 20 hrs at 22 V. The DNAs were transferred to the Nytran SuperCharge membrane using TURBOBLOTTER^{TM} Rapid Downward Transfer System (Schleicher & Schuell Bioscience, Keene, NH; Cat # 10416304). The membrane was baked at 80 ^{0}C for 2 hrs to fix the DNA.

The probe hybridization and DIG luminescent detection were carried out using the DIG Luminescent Detection Kit (Roche, Indianapolis, IN; Cat # 11363514910). The prehybridization and hybridization of the *GFP* probe to the target was performed at 45 ^{0}C using 4 µl PCR labeling reaction per ml of DIG Easy Hyb buffer (Roche, Indianapolis, IN; Cat # 1603558) 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) was used. Membranes were exposed to X-ray films for 30 minutes.

After exposure, membranes were stripped twice for 15 min at 37 °C in 0.2 M NaOH containing 0.1% SDS to remove the DIG-labeled *GFP* probe, rinsed for 5 min in 2 × SSC, and stored in 2x SSC buffer. Stripped membrane was then prehybridized and hybridized with the *PTLF* probe at 42.5 ^{0}C using the same procedure for the *GFP* probe hybridization and detection. Developed films were scanned with 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).

Inverse PCR for T-DNA structure

To investigate T-DNA structure for inserts with multiple copies, five different primers located along the T-DNA sequence were used for PCR amplification (Fig. 2.4). Presence and size of amplified bands were used for infer T-DNA

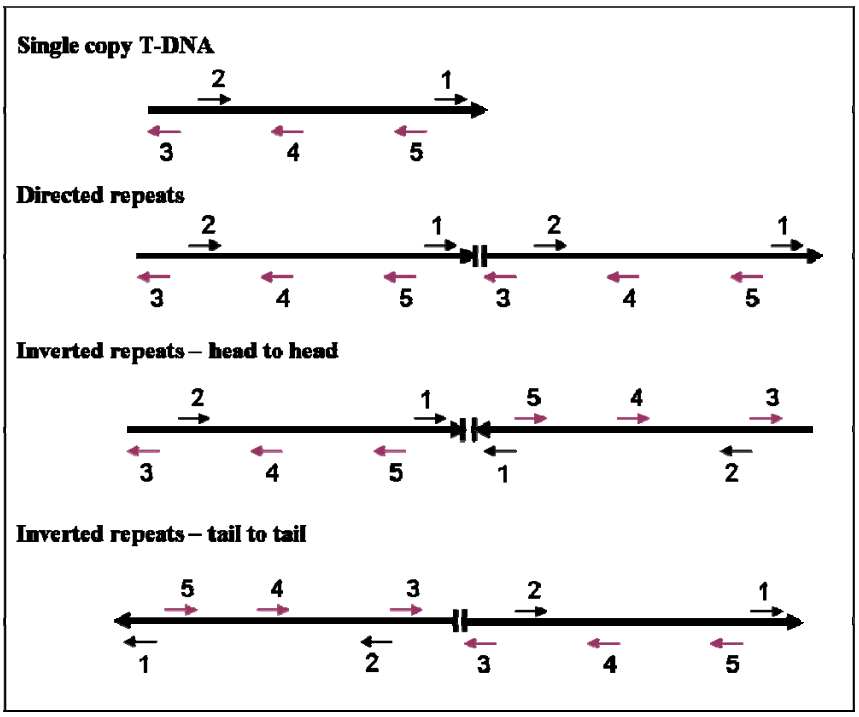


Fig. 2. 4. Diagrams of T-DNA repeats and PCR primers to determine T-DNA structure. Relative positions of five PCR amplification primers (1 ~ 5) along the T-DNA are indicated. Larger arrows represent T-DNA and point toward the right border.

structure (Table 2.1). The five primers used were: 1 - 5' TTTCTGGCAGCT GGACTTCAG 3'; 2 - 5' TAGAAA AGGAAGGTGGC TCCTACA 3'; 3 - 5' CCAAGCTCTTCAGCAATAT CAC 3'; 4 - 5' AGAAGGACCATGT GGTCTCT 3'; 5 - 5'AACCCACGTCA TGCCAGTTCC 3'. The PCR was performed using Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA, Cat # 00304-029) in a volume of 25 µl with the following concentrations of reaction reagents: primer - 0.4 µM, dNTPs - 0.2 mM, $MgSO_4$ - 2 mM, DNA - 20 ng, Taq - 2 U, and 1x High fidelity buffer. The PCR cycles were: 95 ^{0}C for 2 min; 30 cycles of: 94 ^{0}C - 20 s, 59 ^{0}C - 30 s, and 72 ^{0}C - 30 s per kb; 72 ^{0}C for 4 min.

Table 2. 1. Primer pairs used for PCR amplification to determine T-DNA repeat structure. Letters a, b, and c represent band size (kb), and number in parentheses indicates expected band size when there is no genomic DNA separating two repeats of T-DNA . Abbreviations: NR: no repeats; DR: direct repeats; IR-TT: inverted repeats arraigned as tail to tail; IR-HH: inverted repeats arraigned as head to head.

Primer Pair		NR	DR	IR-TT	IR-HH
P1	1 + 3	-	a (2.1)	b (2.5)	c (1.6)
P2	1 + 4	-	a + 1.6 (3.7)	b + 3.2 (5.7)	c (1.6)
P3	1 + 5	-	a + 3.8 (5.8)	b + 7.4 (9.9)	c (1.6)
P4	2 + 3	-	a + 3.1 (5.2)	b (2.5)	c + 6.2 (7.8)
P5	1 + 2	-	-	-	c, c+6.2, c+3.2
P6	3 + 4	-	-	b, b + 3.2, b+1.6	-

Notes: estimated size of amplified bands shown in the table is based on the pG3KGB construct. For pG3MKGB, the size is 2.4 kb larger for all corresponding bands. Amplification patterns listed here were based on presence of complete repeats. Other possible combinations resulting from T-DNA truncations are not listed here, but explained in the text.

Statistical data analysis

GFP fluorescence levels measured in the field in 2004 and 2005 were normalized according to the ratio of mean background levels of non-transgenic plants in two different years. After normalization, the two years have the same background levels. Due to different optical cables used for the greenhouse and field measurements, the measured *GFP* fluorescence levels in the greenhouse were not normalized in the same way; the sample values and background levels responded differently to different optical cables (Appendix F2.4). After normalization, quantitative values were log2 transformed for statistical tests.

To account for variations among different ELISA plates, measured optical density (OD) of the samples run on one plate was divided by the OD of the reference sample on the same plate. Normalized OD was then square root transformed for statistical tests. Except for testing the effect of MAR and transgene copy number on transgene variation among ramets and ANOVA analysis of expression levels, mean expression values averaged over four ramets for each event were used for statistical analysis. Overall ANOVA analysis of *GFP* expression was performed using a random split-block design (Appendix T2.3) with field expression data (year 2004 and 2005). Event was considered a random effect, while block, clone, and construct were treated as fixed effects. ANOVA of *BAR* expression was performed on quantified PAT levels from all three years using a mixed model with event as the only random effect. ANOVA was performed in the SAS system (v8).

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 three 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 ordered years. The slopes and pooled residual variances of regression line for each event were used

to identify unstable events over time. Normality of regression slope distribution was tested with a 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-test (unequal variance) and Wilxion rank tests were used to test effect of MARs, transgene copy number, and T-DNA structure on regression slope and residual variance, respectively. The analysis was performed with the graphical and statistical programs SigmaPlot (9.0) and S-PLUS (7.0). Levene's test on equality of variances was used to examine effect of MARs and transgene copy number on variance of expression levels among events and ramets (Table 2.2). To perform Levene's test on variance among events, an initial ANOVA was performed with averaged *GFP* or *BAR* expression level over three years as response variable and the test factor (MARs or transgene copy number) as independent variable. Absolute values of regression residuals were then used as response variable of the second ANOVA analysis with the test factor again as independent variable (A section in Table 2.2). To perform Levene's test on variance *GFP* expression among ramets, the averaged expression level of each ramet over three years was used as response variable, and the initial ANOVA models included event as random effect in addition to fixed effect of MAR or copy number (B section in Table 2.2) To perform Levene's test on interaction effect of MARs and transgene copy number on variance, the initial ANOVA models included MARs, copy number, and their interaction term as independent variables (plus event as a random effect for variance among ramets), but the second ANOVA model used the interaction term of MARs and copy number as an independent variable (section C and D in Table 2.2). All tests were performed in SAS (8.0).

Table 2. 2. Levene's test on effect of MAR and transgene copy number on variance of transgene expression among events or ramets.

A: effect of MAR or COPY on variance of expression among events	
Initial model	$Y_{ij} = \mu + MAR/COPY + \varepsilon_{ij}$ Y_{ij}: mean expression level of each event averaged over three years
2nd model	$R = \mu + MAR/COPY + \varepsilon$ R : absolute values of regression residuals of the initial model
B: effect of MAR or COPY on variance of expression among ramets (for *GFP* only)	
Initial model	$Y_{ijk} = \mu + MAR/COPY + EVENT\ (MAR/COPY) + \varepsilon_{ijk}$ Y_{ijk}: mean expression level of each ramet averaged over three years; EVENT within MAR or COPY as random effect
2nd model	$R = \mu + MAR/COPY + \varepsilon$ R : absolute values of regression residuals of the initial model
C: interaction effect of MAR and COPY on variance of expression among events	
Initial model	$Y_{ijk} = \mu + MAR + COPY + MAR*COPY + \varepsilon_{ijk}$ Y_{ijk}: mean expression level of each event averaged over three years
2nd model	$R = \mu + MAR*COPY + \varepsilon$ R : absolute values of regression residuals of the initial model
D: interaction effect of MAR and COPY on variance of expression among ramets (for *GFP* only)	
Initial model	$Y_{ijkt} = \mu + MAR + COPY + MAR*COPY + EVENT\ (MAR, COPY) + \varepsilon_{ijkt}$ Y_{ijkt}: mean expression level of each event averaged over three years; event within MAR and COPY as random effect
2nd model	$R = \mu + MAR*COPY + \varepsilon$ R : absolute values of regression residuals of the initial model

RESULTS

Preliminary study of *GFP* expression

Preliminary studies on *GFP* expression variation within a single leaf, spatial variation within a plant, and temporal variation during a day were performed in the greenhouse and field to determine sampling choices for *GFP* measurement. Expression levels were very similar among eight measured leaf areas of a given leaf, and did not differ significantly from each other (Table 2.3; Appendix A).

Table 2.3. *GFP* expression variation within a leaf. Expression level at mid-vein (V) and eight different leaf areas (L1 - L8) delimited by mid-vein and smaller veins of a single leaf for each of eight transgenic events and two non-transgenic plants (NT) was measured. Values represent mean *GFP* expression levels of a certain leaf area with standard deviation indicated. Lm: mean over different leaf areas (L1 - L8); L-V: t-test on difference of mean *GFP* expression between leaf areas and mid-vein. Significant differences ($p < 0.05$) are indicated by *.

Event	L1	L2	L3	L4	L5	L6	L7	L8	Lm	V	L-V
3MGB242	387 ± 9	388 ± 5	398 ± 2	378 ± 5	382 ± 11	383 ± 8	-	-	385± 9	461 ± 47	*
3GB183	441 ± 10	429 ± 6	417 ±11	405 ± 5	433 ± 15	420 ± 9	421 ± 17	428	426 ±14	484 ± 22	*
3GB141	439 ± 6	456 ± 6	449 ±20	443	470 ±13	434 ±13	443 ± 3	438	447 ± 17	524 ± 56	*
3MGB149	365 ± 12	365 ± 11	368 ± 6	366	388 ± 7	382 ± 8	394	-	373 ± 13	427 ± 73	*
7MGB291	361 ± 5	350 ±14	348 ± 11	346	374	357 ± 9	351 ± 27	362 ± 18	355 ± 16	400 ± 21	*
7MGB565	468 ± 17	482 ± 39	478 ± 26	496 ± 11	471 ± 11	443 ± 30	483 ± 19	482 ± 20	475 ±26	516 ± 32	*
7GB381	351± 11	351 ± 12	354 ± 31	359 ± 11	356 ± 13	347 ± 12	343 ± 12	370 ± 6	352 ±15	370 ± 17	*
7GB77	415 ± 10	392 ± 9	393 ±25	391	395 ± 21	397 ± 13	397 ± 12	391 ± 2	396 ±16	400 ± 30	
353 NT	199	206	222	207	208	221	-	-	210 ± 9	207 ± 9	
717 NT	245 ± 8	236	229	256	234	234	206	248	237 ±15	245 ± 15	

However, for seven of the eight studied transgenic plants, the mid-vein had significantly higher *GFP* expression than other portions of the leaf, while this difference was not significant in NT plants (Table 2.3; Appendix A). Therefore, we avoided the mid-vein when taking *GFP* measurements. Although *GFP* expression was very similar among other parts of a leaf, we sampled similar spots for all subsequent *GFP* measurement on different leaves.

GFP fluorescence variability among different leaves on the same plant has been previously described (Halfhill et al., 2003). For in-vitro grown plants (Fig. 2.5A), *GFP* expression levels appeared to be higher in older leaves of the same plant. In contrast, in the greenhouse and field, the highest fluorescence was in young leaves near the apical meristem and declined in the older leaves. In our study, similar fluorescence variation patterns were observed for leaves along a shoot for transgenic plants grown in the greenhouse and field environment (Fig. 2.5B&C). Background levels of non-transgenic control leaves obviously lacked such a pattern. The degree of fluorescence decline was more dramatic among the first two to four leaves, depending on expression levels, and generally tended to be more stable after the third leaf down from the apical meristem. This tendency of change in *GFP* expression within a plant appeared to be steeper in the field plants than in the greenhouse plants.

Based on these expression patterns among leaves within a shoot, we measured *GFP* expression on leaves at plastachron four and five of each plant in the greenhouse, and on the leaf before the first fully expanded leaf from each of two branches of a single plant in the field (5 - 7). Our sampling choices were intended to achieve both sensitivity so that we could differentiate absolute expression levels among transgenic events, and to minimize developmental variability by avoiding highly variable young leaves and mid-ribs.

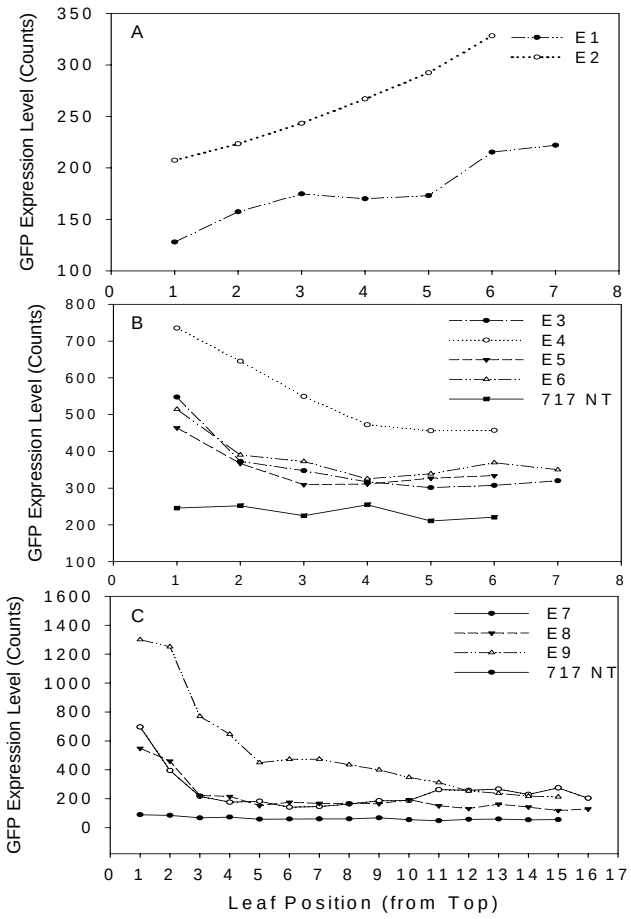


Fig. 2. 5. *GFP* expression variation within a plant grown in an in-vitro (A), greenhouse (B), and field environment (C). Each line represents one event (E1 ~ E9) or non-transgenic plant (NT). When *GFP* measurement was taken, in-vitro plants were one-month old, greenhouse plants were four-months old, and field plants were two-years old.

Due to the large number of transgenic events for *GFP* quantification, we needed to take measurements as long as possible each day. We therefore assessed the extent of diurnal variation in *GFP* expression. We measured a single leaf of 19 transgenic events and two control plants at three different times

(7 AM to 5 PM) in the greenhouse, and found that *GFP* fluorescence tended to be stable (Fig. 2.6A). Similar results were obtained for field grown plants from

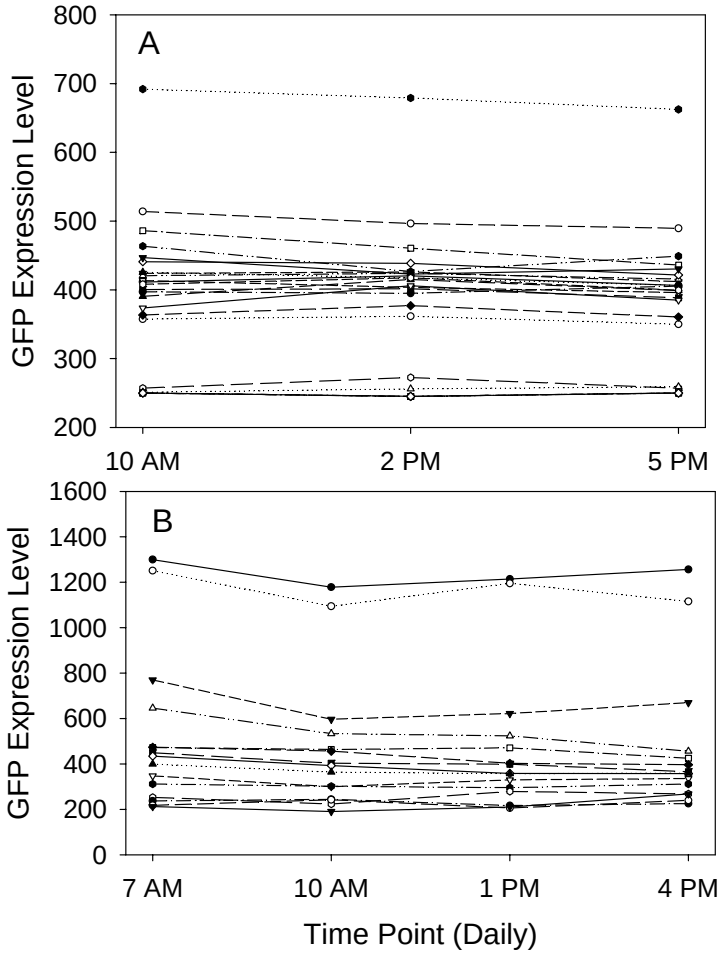


Fig. 2. 6. *GFP* expression change over time of a day in the greenhouse (A) and field (B). For the greenhouse, each line represents one leaf of each of 19 studied transgenic events and two non-transgenic plants (bottom). For expression the field, each line represents one leaf of 15 measured leaves from a single transgenic plant of three events studied.

different events, where all leaves from three different individual plants were measured for *GFP* change over time in the field, and the change of all leaves from one event is shown in Fig. 2.6B. Based on these results and practical considerations, we did not discriminate between daily sampling times when assessing gene expression.

Variation of transgenic expression among transgenic events

Expression levels of *GFP* transgene were measured on 404 transgenic events and 160 subevents for three consecutive years (2003-2005). The measurement in 2003 was done in the greenhouse before the plants were transferred to the field. A strong association was observed between quantified *GFP* values and visualized *GFP* intensities (Appendix F2.5). Expression levels for *BAR* were measured on 168 transgenic events and 78 subevents produced from the poplar clone 717-1B4 using a commercial ELISA kit. Based on the ELISA assay performed on *BAR* transgenic poplars produced for another study, quantified protein levels correlated well with visualized herbicide damage (data not shown).

A wide range of transgene expression levels were observed for both *GFP* and *BAR* transgenes (Fig. 2.7). Out of 404 transgenic events, four events (7gb41, 7gb319, 7gb227, 3gb102) were initially silenced and remained so in the field (*GFP* expression same as for non-transgenic controls). Three were transformants of 717-1B4, and one was of 353-53. None of these three 717-1B4 transgenic events showed initial silencing of *BAR* gene based on ELISA results, but initial silencing of *BAR* occurred for the event 3gb102 (data not shown).

No significant difference in mean *GFP* expression was observed between the clone 353 and 717-1B4 ($p > 0.05$, Appendix T2.3). In addition, the two clones had similar distributions of *GFP* expression in the first two years (Appendix F2.6), and did not differ significantly from each other (Kolmogorov-

Smirnov test, $p > 0.05$). However, the distributions were significantly different in year 2005 ($p < 0.05$).

The distribution of *GFP* expression followed a similar pattern in the three different years (Fig. 2.7A&B&C). There was no significant difference in

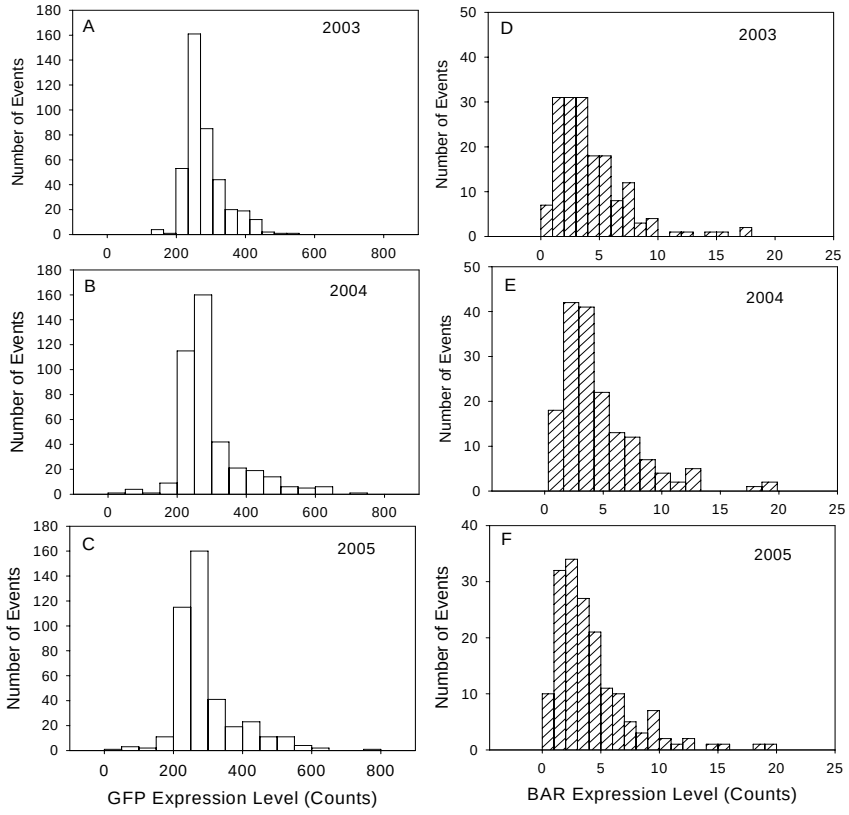


Fig. 2. 7. Histograms of *GFP* and *BAR* expression levels in three different years. Quantified expression values in the different years were scaled to their corresponding median values. A-C: *GFP* expression levels of 404 transgenic events produced from poplar clones 717-1B4 and 353-53-53 in year 2003 (greenhouse), 2004(field), and 2005(field), respectively. *GFP* expression 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 year 2003, 2004, and 2005, respectively. *BAR* expression was quantified with the ELISA assay.

distributions between any two years (Kolmogorov-Smirnov test, $p > 0.05$). Similarly, no significant difference in *BAR* expression distributions was observed among different years (Kolmogorov-Smirnov test, $p > 0.05$; Fig. 2.7D&E&F).

Stability of transgene expression over time

A strong correlation in expression levels was observed between years for both transgenes (Fig. 2.8). Correlation in *GFP* fluorescence levels between two years in the field appeared to be stronger than the correlation between the greenhouse and field, while no clear effect of this kind was seen in the correlation in quantified PAT protein levels.

No instances of complete breakdown (gene silencing) over time were observed for either transgene. We used regression approaches, i.e., regression of mean expression levels of event on median expression levels of year, to detect events with significant changes in expression levels. Large absolute values of regression slopes indicated high sensitivity of transgene expression to the general environment (year of measurement), and large regression residual variances indicated high random variation not explained by the general environment in transgene expression. For *GFP* expression, regression slopes approximated a normal distribution (Fig. 2.9A). Mean slope value was 0.81 with a standard deviations of 1.5. Four events with a slope value beyond 3 standard deviation of the mean were excluded from the estimate of the mean and standard deviation of the distribution. Statistics based on the normal distribution was 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 year of measurement. Thus, a total of eight events (2%) showed significant changes in expression over years. Five of these showed a consistent increase in

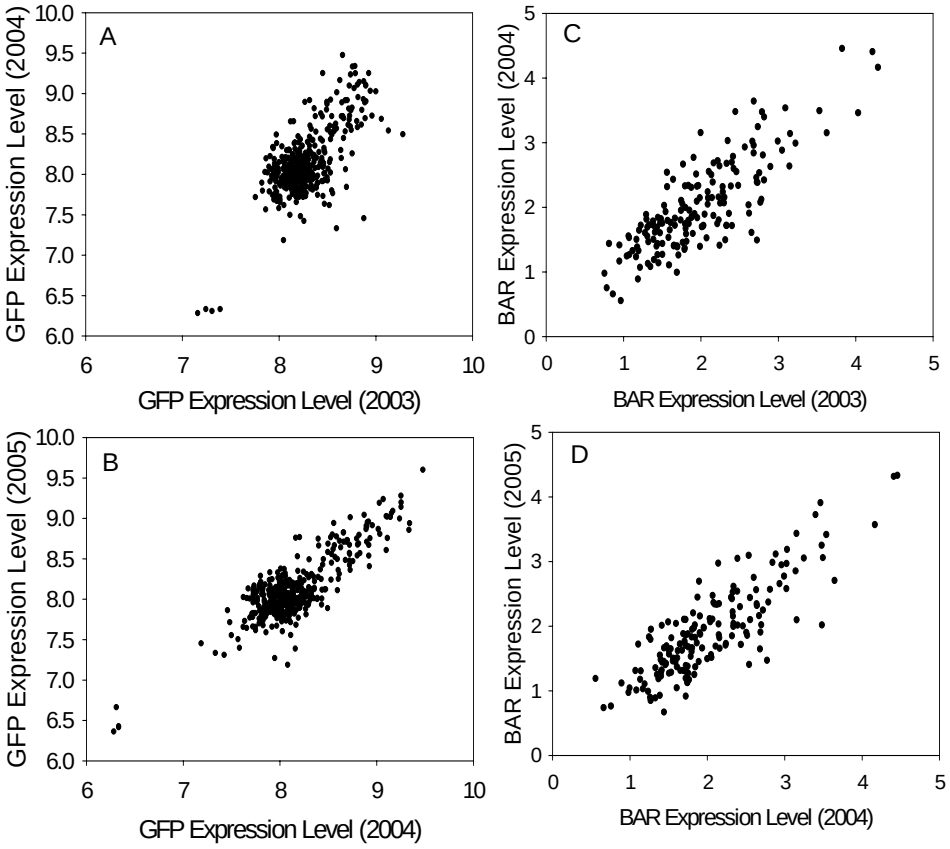


Fig. 2. 8. Correlation of transgene expression levels between different years. A: correlation of *GFP* expression levels between year 1 (2003) and year 2 (2004) (N = 404, r = 0.70). B: correlation of *GFP* expression levels between year 2 (2004) and year 3 (2005) (N = 404, r = 0.86). C: correlation of *BAR* expression levels between year1 (2003) and year 2 (2004) (N = 168, r = 0.81). B: correlation of *BAR* expression levels between year 2 (2004) and year 3 (2005) (N = 168, r = 0.84). For *GFP,* measured values were log2 transformed, and for *BAR, values* were square root transformed. All correlations were statistically significant ($p < 0.01$).

GFP expression, and the remaining three (0.7%) showed a consistent decrease over years. None of these three reduced expression events showed signs of complete gene silencing; all had an expression level at least two times more than background levels (NT controls). Based on chance alone, we

expected a significant change in expression in 4 of 404 total events. For *BAR* transgene expression, regression slopes also followed a normal distribution (Kolmogorov-Smirnov test, $p > 0.05$) (Fig. 2.9C). The calculated means and standard deviations of the distribution were then used to identify events that were beyond the 99% distribution area ($p = 0.01$). One out of 168 (0.6%) events was identified as significantly sensitive to the environment with increased expression over years, well within expectation due to chance alone.

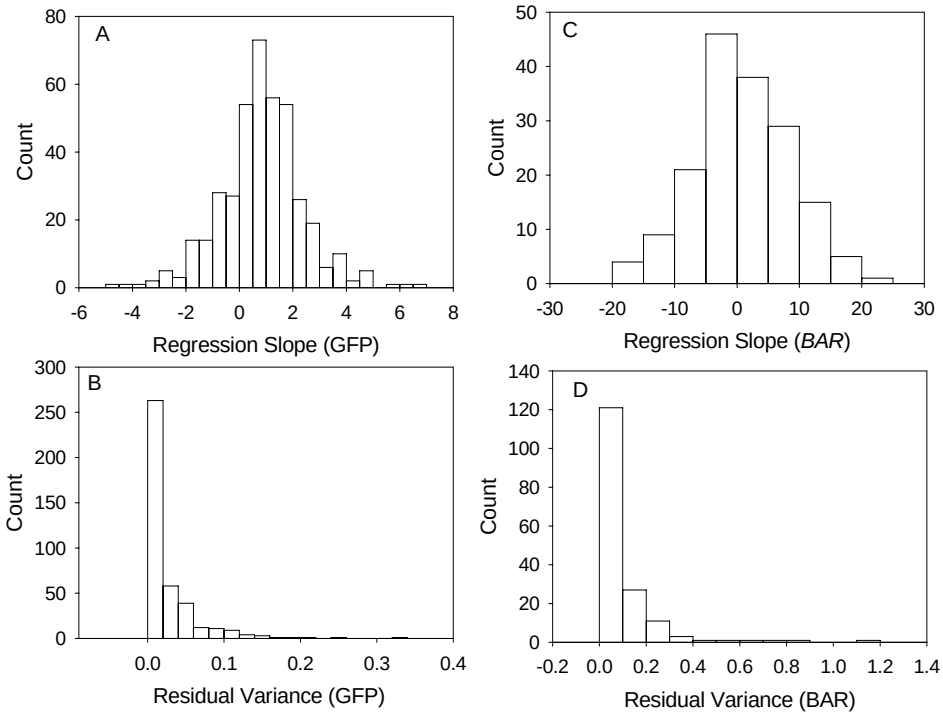


Fig. 2. 9. Regression slopes and residual variance for stability evaluation of transgene expression over three years in the greenhouse and field.
Mean expression level of each event was plotted against median expression levels of each of three years. A, B: distributions of regression slopes (A) and residual variances (B) of *GFP* expression levels of 404 transgenic events. C, D: distributions of regression slopes (C) and residual variances (D) of *BAR* expression levels of 168 transgenic events.

Pooled residual variances of each regression line are an indication of changes in transgene expression caused by random variation which was not explained by average effect of years. A majority of events had a residual variance around zero as expected (Fig. 2.9B, D), indicating that the majority of variation in transgene expression was explained by regression or 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 404 (1%) events were identified for *GFP* expression, and five of 168 (3%) events for *BAR* expression. Only the latter result exceeds expectation due to chance alone.
In summary, approximately 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 approximately 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 third year compared to the first year. There were no events that had statistically significant expression changes in both genes, nor any newly and strongly silenced event at the end of the study.

Effect of organogenesis on stability of transgene expression

A total of 80 randomly selected transgenic events, with 20 from each of four clone x construction combinations, were subjected to a further round of organogenesis to produce subevents. Two independent subevents from each of 80 reintroduced events were produced (Appendix F2.1), therefore a total of 160 subevents were studied. *GFP* expression quantification was made on all 160 subevents, while PAT ELISA assay was performed on 76 subevents in the 717-1B4 background. Expression levels of regenerated subevents generally correlated well with their corresponding events (Fig. 2.10; Fig. 2.11), but there

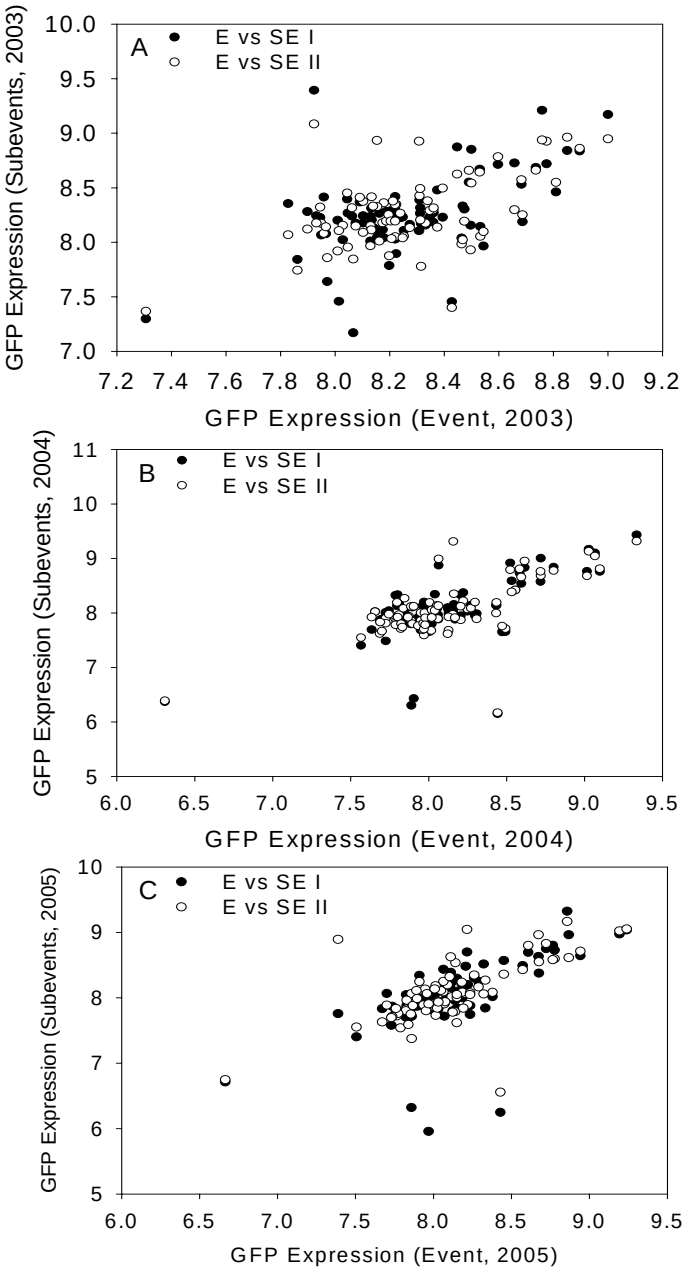


Fig. 2. 10. Scatter plots of *GFP* expression between parent events and their corresponding subevents in three different years.160 subevents were plotted against their corresponding events. In the legend: E – event, S I – subevent I, and S II – subevent II.

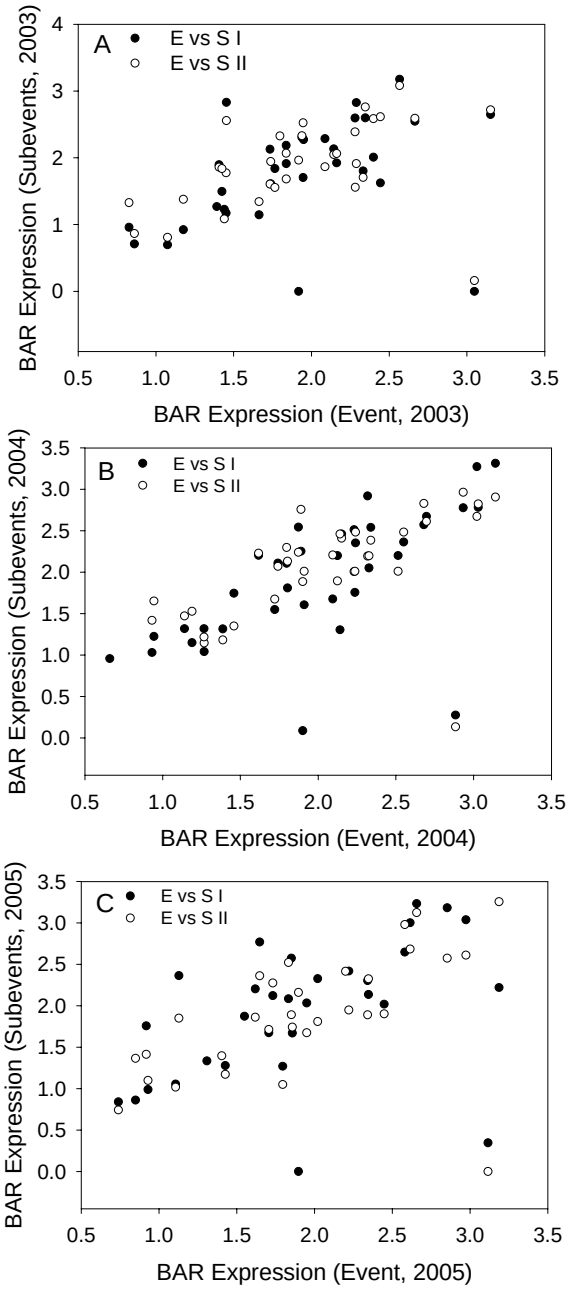


Fig. 2. 11. Scatter plots of *BAR* expression between parent events and their corresponding subevents in three different years. 76 subevents were plotted against their corresponding events. In the legend: E – event, S I – subevent I, S II – subevent II.

were several subevents that showed large change in transgene expression levels compared to their corresponding events based on visual inspection of scatter plots. Four subevents from three different events failed to express *GFP* and *BAR* transgenes above background levels characteristic of non-transgenic controls, while expression levels of their corresponding parental events were in the middle or top half of the expression ranges. PCR amplification 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 physically lost in these four subevents, while their corresponding parental events all had the transgenes present (Appendix F2.7).

`To further identify 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 of lines as potential unstable subevents. The four subevents that lost the transgenic cassettes were excluded from the analysis. Significant outliers which affected normality of the regression residuals were excluded from the construction of prediction intervals, but considered as potential unstable subevents as well. Three subevents: 3mgb161S1 and S2, and 3mgb281S2 always showed substantially elevated *GFP* expression in subevents in all three years, and were considered as outliers in regression analysis. The normality of regression residuals was checked with a normal QQ plot and tested with a Kolmogorov-Smirnov normality test ($p > 0.05$, Appendix F2.8). Five additional subevents (3%) were identified to have significant changes in transgene expression levels by 99% prediction intervals (Fig. 2.12A), but instability was not observed in all three years. Among them, two subevents (1%) showed reduced *GFP* expression compared to their events, while the other three (2%) showed increased expression. Based on the number

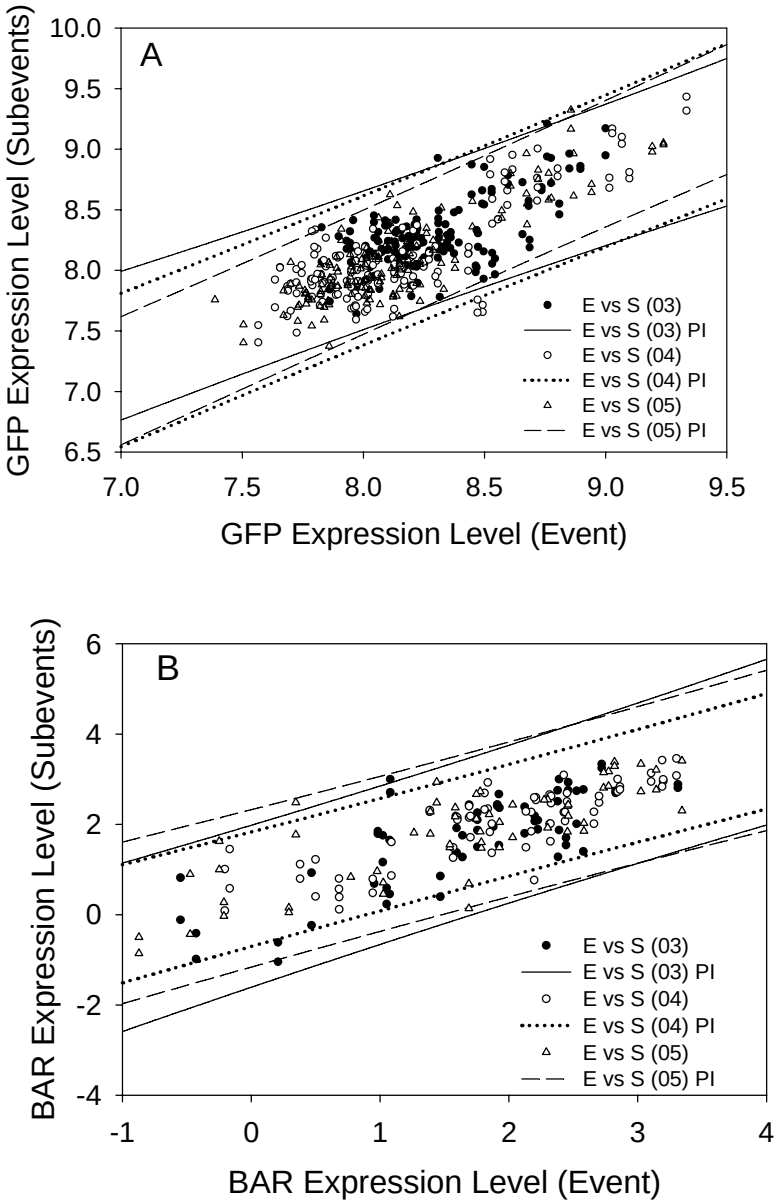


Fig. 2. 12. Organogenesis treatment effect on stability of transgene expression. A: regression of *GFP* expression of 160 subevents on expression of their corresponding events and 99% prediction intervals in three different years. B: regression of *BAR* expression levels of 76 subevents on expression of their corresponding events and 99% prediction intervals in three different years.

of regression points assessed, we expected two of 160 total points to be statistical outliers.

For *BAR* transgene expression, three subevents (4%) were identified as unstable by regression analysis (Fig. 2.12B), 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 three years, and only one subevent, 7gb247S1, had significant expression changes for both genes in one of the three years. None of the other identified subevents showed simultaneous expression changes in both transgenes.

Effect of MARs on transgene expression and stability

Effect of MARs on expression level and variance

To examine the effect of the RB7 MAR from tobacco on transgene expression level, variance, and stability, these MARs flanked the T-DNAs in one of the two transgene constructs. MARs did not significantly affect mean expression level of the *GFP* gene; the means for the two construct classes differed by only 3% (Table 2.4). In contrast, the mean *BAR* expression level without MARs was significantly higher ($p < 0.05$) than that of MAR-containing transgenic events, a difference of 23%.

MARs did not have a significant effect on distributions of *GFP* expression between events with and without MARs in any of the three years studied (Kolmogorov-Smirnov test, $p > 0.05$; Fig. 2.13; Appendix F2.9). However, only for *GFP* without MARs did there appear to be any silenced events (Fig. 2.13). Correspondingly, the distribution appeared to have a greater dispersion for events without MARs. This was reflected in a significantly greater

Table 2. 4. Effects of MARs, copy number, and T-DNA structure on transgene expression and variances among events. Test of equality of variances were performed with Levene's test. Abbreviations: M: MAR; NM: no MAR; NR: no repeats; DR: direct repeats; IR: inverted repeats; SD: standard deviation among event means; * : Statistically significant ($p < 0.05$); ns: not significant.

Factor	Level	*GFP*					*BAR*				
		N	Mean	Mean Test	SD	Variance Test	N	Mean	Mean Test	SD	Variance Test
MAR	M	200	293	ns	64	*	83	3.7	*	2.4	*
	NM	191	302		80		82	4.8		3.6	
Copy	1	334	279	*	53	*	144	3.89	*	2.78	ns
	2+[a]	57	402		82		21	6.89		3.72	
MAR *Copy	1M	173	275	ns	41	*	73	3.6	ns	2.28	ns
	2+M	27	410		63		10	5.56		2.72	
	1NM	161	284		63		71	4.39		3.21	
	2+NM	30	395		97		11	8.38		1.44	
T-DNA repeats	NR[b]	27	377	*[c]	89	ns [c]	7	5.3	ns [c]	2.1	ns [c]
	DR[b]	11	467		72		8	7.8		2.8	
	IR[b]	4	310	-	99	-	2	4.8	-	3.5	-

[a]: 2+ represents two or more copies
[b]: values were only based on two-copy transgenic events.
[c]: tests were performed between NR and DR.

variance and standard deviation in the non-MAR group (Table 2.4). However, after the four transgenic events with silenced *GFP* expression were removed from variance analysis, the two groups did not significantly differ in expression variance ($p = 0.11$). MARs also did not significantly affect variance of expression at the within event level ($p > 0.05$, Appendix T2.4).

Effect of MARs on expression distributions appeared to be greater for *BAR* than for *GFP*, as such effect was statistically significant (Kolmogorov-Smirnov test, $p < 0.05$; Fig. 2.13; Appendix F2.10). Similar to their effect on *GFP* expression variance among events, MARs significantly reduced variance

of *BAR* expression ($p < 0.05$; Table 2.4), and resulted in a smaller dispersion of the distributions for MAR-containing events.

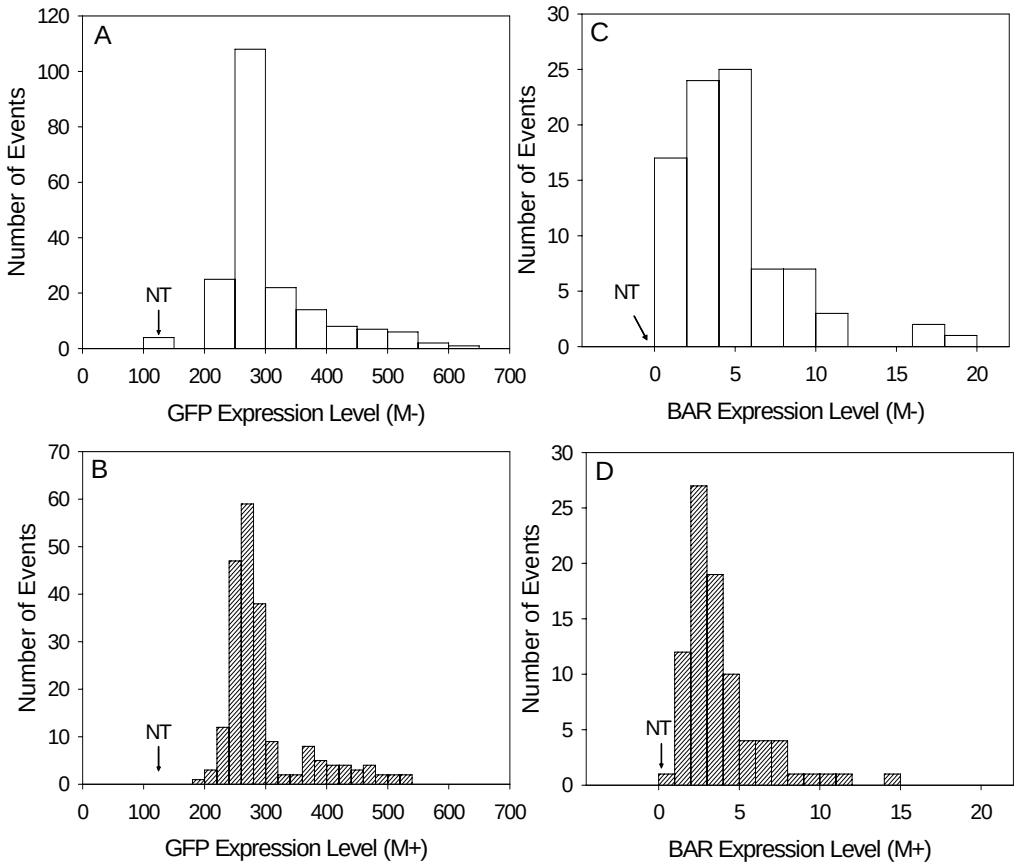


Fig. 2. 13. Histograms of *GFP* and *BAR* expression levels of transgenic events with and without MARs. A, B: *GFP* expression distribution of transgenic events without (A) and with (B) MARs, respectively. C, D: *GFP* expression distribution of transgenic events without (C) and with (D) MARs, respectively. For each event expression level was averaged over three years for making histograms. Mean background levels of non-transgenic plants (NT) are indicated.

Effect of MARs on correlated expression of linked GFP and BAR transgenes

Since the *GFP* and *BAR* transgenes were assembled in the same T-DNA, it is reasonable to expect that the two transgenes might show correlated expression levels among transgenic events. Expression levels of *GFP* were therefore plotted against the corresponding *BAR* expression levels of transgenic events transformed with the two constructs pG3KGB and pG3MKGB (Fig. 2.14). There was a weak positive association ($r = 0.15$) between expression levels of *GFP* and *BAR* in transgenic plants without MARs, but the association was not significant ($p = 0.11$). Presence of extreme points could potentially affect calculated correlation strength and statistical significance. After three extreme points as indicated in Fig. 2,14A were excluded from the correlation analysis, the correlation between *GFP* and *BAR* was higher and significant ($r = 0.21$; $p = 0.044$).

In the MAR-containing transgenic events, the correlation between *GFP* and *BAR* expression levels was much stronger ($r = 0.42$), and highly statistically significant ($p < 0.0001$). Therefore, MARs tended to increase correlated expression between two genes linked in the same T-DNA.

Effect of MARs on stability of transgene expression

Among eight events which were identified as highly sensitive to the general environment (year of measurement) for *GFP* expression, four contained MARs and the other four did not. Out of the four events with large random variation in *GFP* expression over time, one contained MARs and the other three did not. Therefore, no obvious association between MARs and stability was

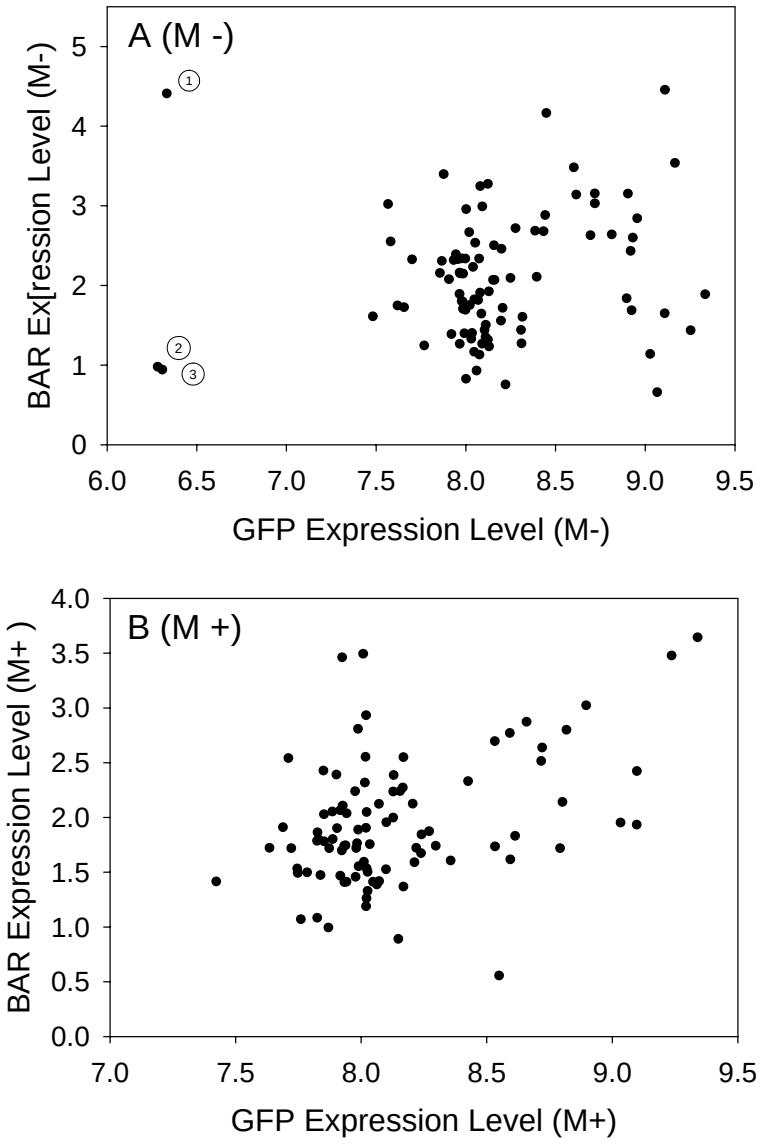


Fig. 2. 14. Correlation of expression levels of *GFP* and *BAR* transgenes. A: correlation of *GFP* and *BAR* expression levels of 92 transgenic events in clone 717-1B4 transformed with the construct pG3KGB without flanking MAR sequences ($r = 0.15$; $p = 0.11$). B: correlation of *GFP* and *BAR* expression levels of 95 transgenic events transformed with the construct pG3MKGB with flanking MAR sequences ($r = 0.42$; $p < 0.0001$).

observed based on those identified events. MARs also did not have a significant effect on mean regression slopes caused by general environmental effects (two small t-test, $p > 0.05$) or regression residuals caused by random variation over time (Wilcoxon rank test, $p > 0.05$).

Similarly, MARs did not have a significant effect on stability of *BAR* expression over time. Out of the five events that showed significant variation in *BAR* expression, three contained MARs and the other two did not. in addition, no significant effects of MARs were observed on mean regression slope or regression residual variance ($p > 0.05$).

Molecular Characterization of Transgenes

Transgene copy number with comparative real-time PCR

Quantitative real-time PCR has been previously used to estimate transgene copy number (Ingham et al., 2001). In our study, we used an endogenous gene *PTLF*, the poplar homolog of *LEAFY*, as the reference gene. *PTLF* exists as a single copy gene in poplar genome, which was confirmed by Southern blot analysis (Appendix F2.9). The 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 full T-DNA transfers. The TaqMan® probe and primer sets were optimized to achieve high and comparable PCR efficiency for both genes (Appendix F2.3). DNA extracted from each of 396 transgenic events was run, in duplicate, in a duplex TaqMan® assay. Transgene copy number was calculated by comparing the Ct of the *GFP* transgene to that of endogenous *PTLF*, as described in the methods section. We used duplicated instead of triplicated reactions as the Ct of two duplicated assays was generally very close ($\Delta Ct < 0.2$, data not shown). There was also a good correlation in calculated copy number

between different runs, i.e., with different reaction mixes and plates run on different days (Fig. 2.15).

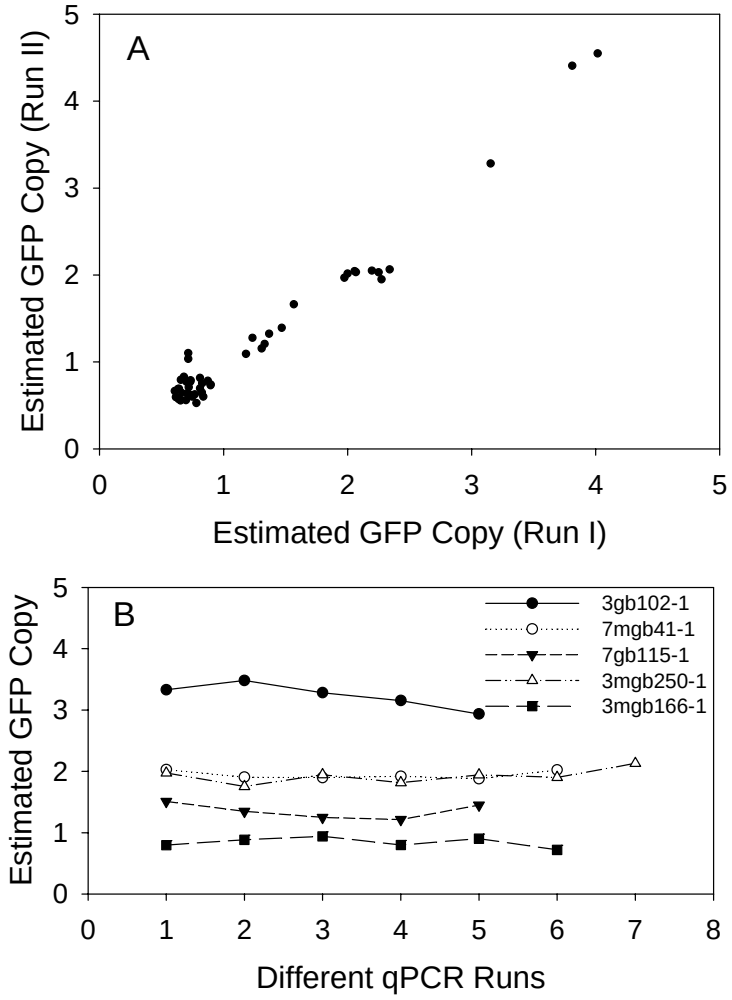


Fig. 2. 15. Reproducibility of estimated transgene copy number among different quantitative PCR (qPCR) runs. A: Scatter plot of estimated copy number of 47 transgenic events between two different qPCR runs (R = 0.9935; Y = -0.1131 + 1.0752*X). B: Variation of estimated *GFP* transgene copy number of five transgenic events among different runs. Each line represents one event .

We estimated copy number for 396 events (Fig. 2.16), and one of these events had multiple copies but the exact number was not determined due to unclear results from Southern. As shown in the figure, the quantified values (QV) were not necessarily integers. Therefore, we estimated copy number using the following ranges: QV<1.5: one copy; 1.5<QV<2.5: two copies, and so on.

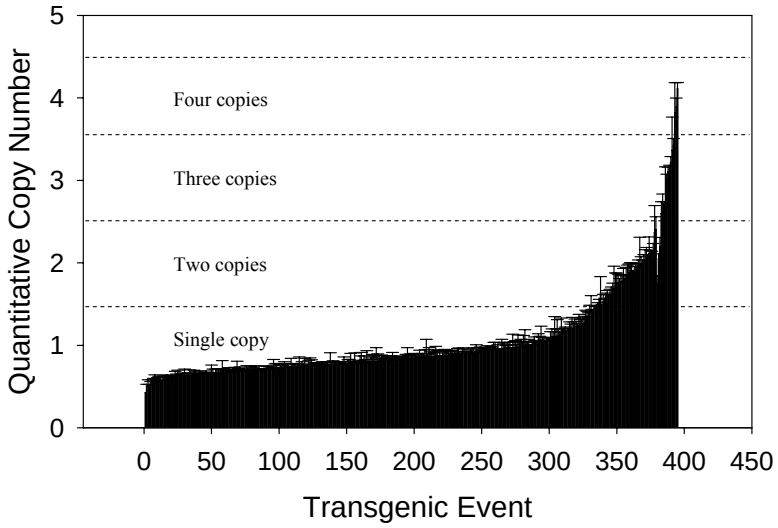


Fig. 2. 16. Quantitative transgene copy number of 396 transgenic events by comparative real-time PCR. Bars represent one standard error of mean copy number of duplicated or multiple reactions of the same sample

We found that our *Agrobacterium-mediated* transformation system produced a majority of single copy transformants (Table 2.5), with 85 % of the transformants contained a single copy, 11 % had two copies, 3% had three copies, and the remaining 1% contained four copies. None of studied events contained more than four copies of the insert. These results are generally consistent with previous observations of *Agrobacterium* transformed

dicotyledonous and monocotyledonous plants (Hansen & Chilton, 1999; Ingham et al., 2001)

Table 2. 5. Estimated transgene copy number of 396 transgenic samples, and correlation with Southern blot results.

Copy No.	No. of Samples Determined by RT Assay (% of Total)	Correlation with Southern			
		(*Sca*I)		(*Hind*III)	
		No. Assayed	No. Consistent (%)	No. Assayed	No. Consistent (%)
1 copy	338 (85%)	17	16 (94.1%)[a]	2	2 (100%)
2 copy	44 (11%)	11	9 (82%)[b]	2	2 (100%)
3 copy	10 (3%)	4	4 (100%)	4	4 (100%)
4 copy	4 (1%)	3	2 (67%)[c]	2	0 (0%)[d]
Total	396	35		10	

[a] The only inconsistent event had a QV of 1.4, and displayed two bands on the Southern blot.
[b] Two inconsistent events had a QV of 1.6, and displayed a single band on the Southern blot.
[c] One sample determined to have four copies by RT 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).
[d] Two samples determined to have four copies by RT assay displayed a signal intensity of three copies by Southern.

To validate our RT 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 genomic DNAs. Both enzymes cut outside of the *GFP* probe sequences (Fig. 2.17A). The number of hybridization bands from *Sca*I digestion indicated the number of transgene loci, while signal intensities of hybridization bands from *Hind*III digestion indicated total number of transgene copies in genome (Fig. 2.17B).

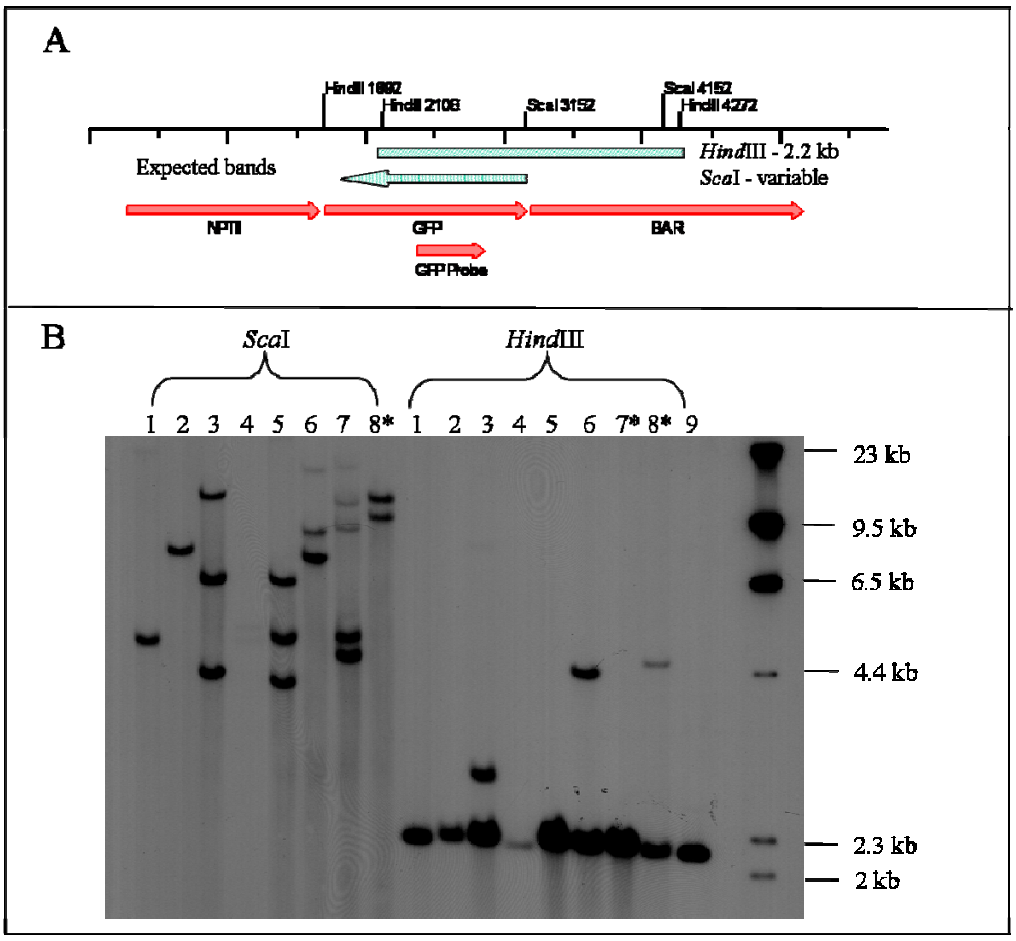


Fig. 2. 17. Southern analysis of transgene copy number. A. graphic representation of the T-DNA of transgene construct pG3KGB. The relative positions of two restriction enzymes *Sca*I and *Hind*III, three transgene (NPTII, *GFP*, *BAR*), and *GFP* probe region are shown. Expected hybridization bands are indicated by blue arrows. B. Southern blot of genomic DNA extracted from eight transgenic events (1~8) plus plasmid DNA of transgene construct pG3KGB (9). Genomic DNA was digested with *Sca*I and *Hind*III respectively, and probed with a DIG labeled *GFP*-specific probe. Inconsistent events between Southern blot and RT are indicated by *. DIG labeled DNA molecular weight marker is shown at right.

We had 15 events with a QV between 1.4 and 1.6, two events with a QV of 2.6, and two events with a QV of 3.4. We performed Southern blot on 15 events with a QV between 1.4 and 1.6, and two events with a QV of 3.4. Out of the 17 events studied, 14 events (82.3%) had a consistent copy number between Southern and quantitative PCR (Table 2.5; Appendix F2.12). In addition to those 17 events, another 18 events with a more decisive QV were randomly chosen for the Southern blot. We found an overall 95% correlation between RT and Southern blot when *Sca*I was used for restriction digestion. Nine of those events, plus an additional event, were further confirmed with *Hind*III digestion. We found 100% correlation when RT determined copy number was less than four copies. But hybridization signal intensities tended to reach saturation at three copies, as two samples with four copies determined by RT had the same intensity as three copies. One of these two events was also digested by *Sca*I, and was shown to have four copies (#7, Fig. 2.17B).

Correlation of transgene copy number with transgene expression levels

Positive correlations between copy number and expression levels were observed for both *GFP* and *BAR* transgenes (Fig. 2.18). The events with multiple copies of inserts had a higher mean expression level of transgenes (~1.5 fold for *GFP*, and ~ 2 fold for *BAR*); the difference in means was significant for both transgenes (Table 2.4; $p < 0.05$). The regression of expression on 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 multiple-copy groups. The presence of multiple copies increased variance in *GFP* expression by 139% among events ($p < 0.05$; Table 2.4). Although a similar effect was observed for *BAR*

expression (79% greater for high copy events) it did not reach statistical significance ($p > 0.05$). A large part of the high variance for high copy number event was a result of one three-copy event that had silenced *GFP* expression, while the rest of the multiple-copy events had high expression (Fig. 2.18A). 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 *GFP* expression among ramets ($p < 0.05$; Appendix T2.4).

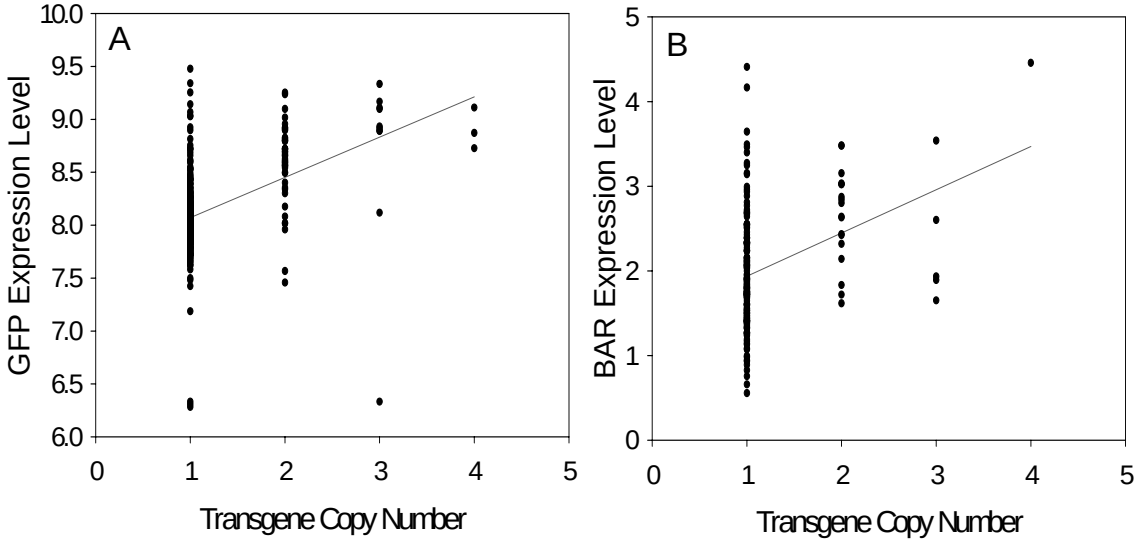


Fig. 2. 18. Correlation of transgene copy number with transgene expression levels. A: correlation of *GFP* expression with transgene copy number ($N = 394$, $r = 0.46$, $p < 0.001$). B: correlation of *BAR* expression with transgene copy number ($N = 187$, $r = 0.35$, $p < 0.001$).

Interaction effect of MARs and transgene copy number and expression

For events transformed with the construct without MARs, 82.7% contained a single copy of the insert, and a similar frequency was seen for events transformed with the construct containing MARs (83.6%; Fig. 2.19).

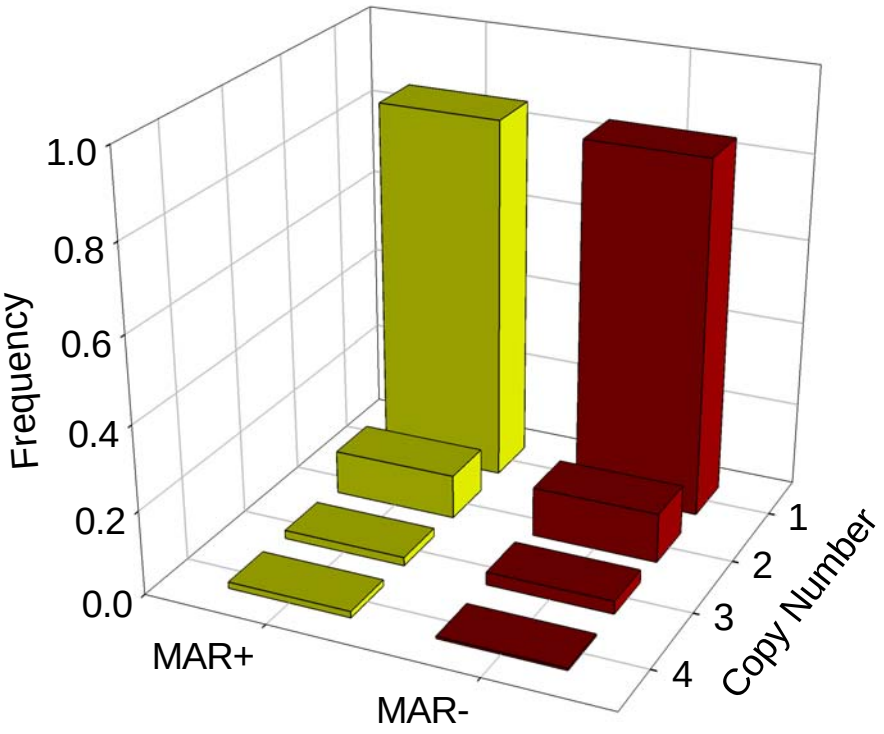


Fig. 2. 19. Distribution of transgene copy number for transgenic events with and without MARs. A: copy number distribution of 201 transgenic events with MARs. B: copy number distribution of 195 transgenic events without MARs.

MARs and transgene copy number did not interact to affect mean transgene expression levels ($p > 0.05$; Table 2.4). However, MARs interacted with copy number in affecting variance of *GFP* expression among events. The presence of multiple copies of transgenes in the events without MARs had the highest variance in *GFP* expression among events (Table 2.4). Therefore, the presence of MARs tended to attenuate the effect of transgene copy number on variance among events. In contrast, such interaction effect was not observed on variance in *BAR* expression among events ($p > 0.05$; Table 2.4).

T-DNA repeat formation

We investigated repeat formation for the 42 events containing two copies of transgenes, and the 13 events with three or four copies. Approximately half of the multi-copy events contained MARs. PCR was performed using primers facing away from each other in the T-DNAs (Fig. 2.4). Therefore, no PCR amplification should be achieved in the case of single copy integration or when multiple copies are integrated at different loci. However, certain amplified bands will be observed when multiple copies of T-DNA are arranged as direct or inverted repeats at the same chromosomal position. The presence and size of PCR products from different pairs of primers may determine the types of repeats, as well as whether the repeats are complete or incomplete (Table 2.1). A similar strategy has been described previously (Kumar and Fladung, 2001).

Analysis of two-copy events

Out of 42 two-copy events, 26 (62%) events showed no amplification of any expected PCR fragments, a result consistent with two copies inserted at

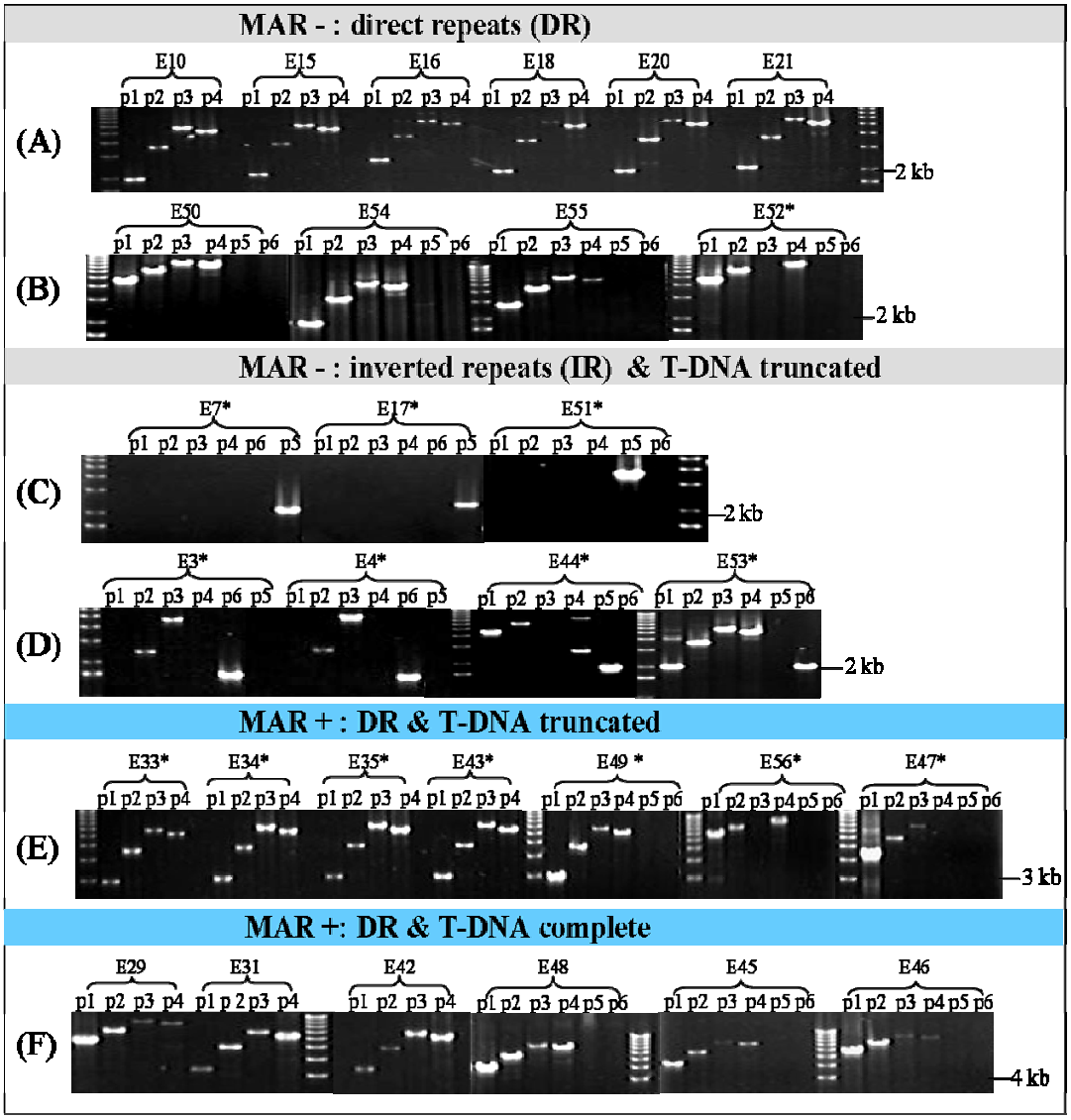


Fig. 2. 20. PCR amplified bands of transgenic events with two or more copies of inserted T-DNAs. E10 ~E21: two-copy transgenic events without MARs. E29 ~ E43: two-copy transgenic events with MARs. E45 ~ E 49: three copy transgenic events with MARs. E50~E55: three copy transgenic events without MARs. E56: four copy transgenic event. p1 ~ p6: different primer pairs (Table 2.1). Events with T-DNA truncated are indicated by "*". A: Two-copy non-MAR events with DR; B: three-copy non-MAR events with DR; C: multiple-copy non-MAR events with IR; D: multiple-copy non-MAR events with both DR and IR; E: multiple-copy MAR events with truncated DR; F: multiple-copy MAR events with complete DR. 1 kb DNA ladder are shown.

different loci. All of these DNAs had also supported successful PCR reactions during copy number determinations. Noneof the five single-copy events tested supported PCR amplification using the primers for determining repeat structure. Two events E3 and E4 showed an amplification pattern indicating the presence of both direct and inverted repeats with large T-DNA truncations occurring (Fig. 2.20D). The presence of a largely truncated T-DNA was not detected by our quantitative PCR due to loss of the primer target region. These two events were therefore considered as three-copy events for the subsequent analysis.

Out of the 19 non-MAR events, eight events (42%) showed positive PCR amplification (Fig. 2.20; Table 2.6), indicating integration of the two copies at the same locus. Among these eight events, six events formed direct repeats (Fig. 2.20A), and two events (E7 & E17) formed inverted repeats arranged as head-to-head (Fig. 2.20C). Seven (33%) of the 21 MAR-containing events had repeat formation at the same locus, and all of them were arranged as direct repeats. A total of 13 events (32.5%) showed a PCR band indicating direct repeat formation at the same locus, with the repeats separated by intervening DNA ranging in size from 0.4 to 2.6 kb (Fig. 2.20A, Table 2.6). For six non-MAR events with direct repeats, five showed complete direct repeats without the obvious presence of intervening DNA sequence (the size and patterns of amplified fragments followed predictions), while the other event had a complete direct repeat with approximately 400 bp of intervening DNA between the repeats. Frequent truncations were observed for MAR-containing events with direct repeats. Out of seven events with direct repeats, three had complete direct repeats, and the remaining four had a large truncation (~1.4 kb) in the MAR sequences at the junctions of two repeats (e.g.Fig. 2. 20E).

A large truncation was also observed in events E7 and E17 containing head-to-head inverted repeats. About 1.9 kb of T-DNA sequence was truncated on the left border of one copy, and 2.7 kb on right border of the other copy. In

Table 2. 6. Amplified PCR products and repeat formation of multiple copies of the T-DNA. DR: direct repeats. TT: tail-to- tail inverted repeats. HH: head-to- head inverted repeats.

Event No.	Event Name	Copy No.	Amplified PCR fragments with different primer pairs (in kb)						Remarks
			1 + 3	1 + 4	1+ 5	2 + 3	1+2	3+4	
7	3gb193	2	-	-	-	-	2.1		HH; truncated
10	3gb236	2	2.1	3.7	5.8	5.2	-	-	Complete DR
15	3gb72	2	2.5	4.2	6.9	6.2	-	-	Complete DR
16	7gb234	2	2.1	3.7	5.8	5.2	-	-	Complete DR
17	7gb261	2	-	-	-	-	2.3		HH; truncated
18	7gb298	2	2.1	3.7	5.8	5.2	-	-	Complete DR
20	7gb99	2	2.1	3.7	5.8	5.2	-	-	Complete DR
21	7gb262	2	2.1	3.7	5.8	5.2	-	-	Complete DR
29	3mgb30	2	7	8.5	~11	~10	-	-	Complete DR
31	3mgb58	2	4.4	6.1	8.2	7.6	-	-	Complete DR
33	7mgb325	2	3	4.6	6.5	6			DR; truncated (MARs)
34	7mgb329	2	3	4.6	6.5	6	-	-	DR; truncated (MARs)
35	7mgb339	2	3	4.6	6.5	6	-	-	DR; truncated (MARs)
42	7mgb553	2	4.4	6.1	8.2	7.6	-	-	Complete DR
43	7mgb85	2	3	4.5	6.5	6	-	-	DR; truncated (MARs)
45	3mgb128	4	5	6.7	8.9	8			Complete DR
46	3mgb116	3	6.5	8.2	10	9	-	-	Complete DR
47	3mgb158	3	4.4	6.1	8.2	-	-	-	DR; truncated (MARs)
48	3mgb218	4	5	6.7	8.9	8	-	-	Complete DR
49	3mgb77	3	3	4.5	6.5	6	-	-	DR; truncated (MARs)
56	7mg546	M*	6.6	8.3	-	10	-	-	DR; truncated
3	3gb141	**3**	-	2.4	4.5	-	-	1.5	DR and TT; truncated
4	3gb142	**3**	-	2.4	4.5	-	-	1.5	DR and TT; truncated
50	7gb281	3	4.4	6.1	8.2	7.6	-	-	Complete DR
51	7gb248	3	-	-	-	-	3.6	-	HH; truncated
52	7gb86	3	4.4	6.1	-	7.6	-	-	DR; truncated
53	7gb392	4	2.1	3.7	5.5	5	-	2	DR & TT; truncated
54	7gb247	3	2.1	3.7	5.5	5	-	-	Complete DR
44	3gb102	3	4.4	6.1	-	7.8/2.6	1.7	-	DR & HH; truncated
55	3gb65	3	2.8	4.5	6.5	6	-		Complete DR

*Contained multiple copies (2~4), but the exact number was not determined.

addition, we noticed that when two copies were arranged as inverted repeats, T-DNA truncations occurred in both copies and on different borders, i.e., if the left border of one copy was truncated, then a truncation of the other copy occurred on the right border.

Analysis of events with three and four copies

When more than two copies are integrated, there are certain cases in which the PCR strategy is not able to determine T-DNA structures for all integrated copies. For instance, if three copies of the insert were arranged as direct repeats with the same size of intervening DNAs between two repeats, then the amplified patterns are the same as two direct repeats on the same locus. However, the amplified bands and size still enable a determination of whether at least two copies were arranged as certain types of the repeats at the same locus. To our surprise, PCR amplified bands were present on all transgenic events with more than two copies (Fig. 2.20, Table 2.6), indicating that for all events at least two copies of the insert formed repeats at the same chromosomal locus.

Out of nine non-MAR containing events, four events (44.5%) contained direct repeats (Fig. 2.20B), one event E51 (Fig. 2.20C) formed inverted repeats arranged as head-to-head, and the other four (44.5%) had the presence of both direct and inverted repeats (Fig. 2.20D). All six events with flanking MARs formed direct repeats on the same locus. T-DNA truncations were also seen in about 50% of the events with repeat formation, especially when MARs or inverted repeats were present (Table 2.6). In event E44 which was silenced for both *GFP* and *BAR* expression, one of the amplified bands from primer pair P4 was not explainable with any kind of repeat formation, and it seems likely that T-DNA rearrangements might have occurred in the inverted repeat copies, or non-specific amplification might be occurring.

In summary, the presence of flanking MARs did not seem to significantly affect the frequency of integration of multiple copies on the same locus. For two-copy transgenic events, the frequency for with and without MARs was 0.42 and 0.33, respectively. The difference was not significant (Chi-square Test, $p < 0.05$). For more than two copies, the frequency was 1 for both groups. The overall frequency for the integration of multiple copies on the same locus was 0.55; direct repeat was the major type of repeat formation (77%). However, there was a strong association between the presence of MARs and occurrence of DNA truncations when multiple copies were arranged as directed repeats. T-DNA truncation occurs at a frequency of 54% when MARs were present, and 10% without MARs. The observed frequencies were statistically significant ($p < 0.05$; Chi-square test). In addition, MARs reduced the formation of inverted repeats at a single locus. When MARs were present, none of the 13 multiple-events with repeat formations at a single locus contained inverted repeats. When there was no flanking MARs, seven of 17 multiple-copy events with T-DNA repeats at one locus contained either inverted repeats or both direct and inverted repeats. The association was statistically significant ($p < 0.01$; Chi-square Test). We also found that formation of inverted repeats at a single locus was always accompanied by T-DNA truncations. All seven events containing inverted repeats, arranged as either head-to-head or tail-to-tail, had T-DNA truncations. In contrast, T-DNA truncations occurred in one of the ten non-MAR events arranged as direct repeats. Difference in frequencies of T-DNA truncations between these two types of repeats was also highly significant (Chi-square Test, $p < 0.001$).

Correlation of transgene expression levels with repeat formation of T-DNAs

We plotted *GFP* expression levels of 40 two-copy events against determined repeat types to see if repeat formation and types of repeats affected transgene expression (Fig. 2.21). Events with direct repeats had a 24% higher *GFP* expression level than events without directed repeats ($p < 0.05$). Due to a small number of events with inverted repeats (2), no strong statistical inferences can

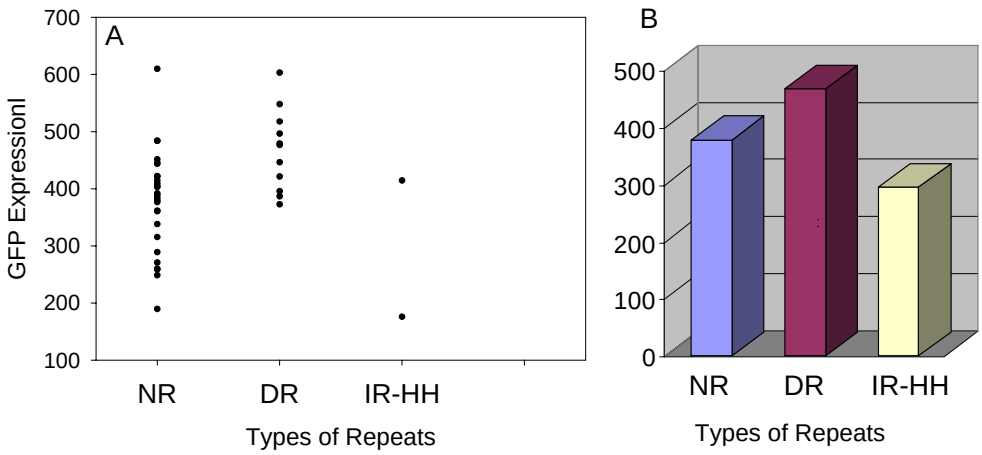


Fig. 2. 21. Correlation of *GFP* expression levels with repeat formation of T-DNAs. The scatter plot (A) was based on 40 two-copy transgenic events whose T-DNA structure was determined with PCR amplification. Mean *GFP* expression for different types of repeats is shown in B. NR: no repeats; DR: direct repeats; IR-HH: head-to-head inverted repeats.

be made regarding their effect on transgene expression. Both of these events showed *GFP* expression at the lower end of the range, and were significantly lower in expression than the direct repeats ($p = 0.004$). However, the mean expression level of these two events did not differ significantly from that of the

events with no repeats ($p = 0.14$). There was no significant difference in stability between the direct repeats and 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 (Table 2.4).

DISCUSSION

GFP expression variation

One major issue that exists for almost all stability studies is that additional variation resulting from experimental procedures, sampling strategy, and physiological and developmental stages can confound actual transgene expression change (James et al., 2004). *GFP* fluorescence was shown to vary within leaf canopies of 35S-*GFP* transgenic oilseed rape grown in an environmental chamber (Halfhill et al., 2003). The fluorescence intensity of individual leaves was highest when a leaf was relatively young, and intensity decreased over time as a leaf aged. As a result, a fluorescence intensity profile existed within a leaf canopy. The fluorescence was highest in young leaves near the apical meristem and decreased in older leaves. We observed a similar pattern of fluorescence change in transgenic poplars grown in both greenhouse and field environments. In contrast, in-vitro grown plants did not display such an obvious pattern and had a slight increase in fluorescence in older leaves. The observed difference between in vitro and greenhouse-field grown plants is likely due to a lower extent of physiological and morphological differentiation induced by environment in the in-vitro grown plants. Therefore, for *GFP* expression measurements of greenhouse and field plants, youngest leaves near the apical meristem should be avoided due to high developmental variability. Based on our results, the optimal leaf positions for study of *GFP* expression in transgenic

poplar plants were the fourth to seventh leaves from the apical meristem, which were generally stable for *GFP* expression or decreased slowly.

GFP fluorescence did not vary significantly among different spots within a given leaf when mid-vein was excluded. *GFP* fluorescence was generally significantly higher and more variable when measured on the mid-vein, which can be attributed to unevenness of a mid-vein in thickness, i.e., varied numbers of cells measured, and non-flat surface, especially for field plants. Therefore, in any case, the mid-vein of a leaf should be avoided. We also did not observe obvious diurnal variation in *GFP* expression, which will facilitate the utility of *GFP* in transgenic poplars when a large number of samples need to be studied.

Stability of transgene expression

One of major biological concerns of transgenic trees is whether transgene expression will be stable over a full rotation cycle that spans years to decades. Several previous studies in poplar transgenic trees have suggested that expression of transgenes in poplar under vegetative propagation can be highly stable over multiple years in the field (Meilan et al, 2002; Hawkins et al. 2003). In our study we monitored expression of 2,256 transgenic poplar trees derived from 404 primary events and 160 subevents over three years in the greenhouse and field. Although some variation in transgene expression levels 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 expression), 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 their rotation cycles. In our 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 both of two years in the field, which could occur by chance alone. None of those events showed sign of complete gene silencing as their expression levels were more than two times that of the background level characteristic of non-transgenic plants. In contrast, no events showed significant reduction in *BAR* expression over years. The difference in observed instability in *GFP* and *BAR* may be due to greater accuracy of the ELISA procedure used in determining protein levels for *BAR*. It may also result from strong influence of leaf age and environment in which leaves develop a background fluorescence that obscures *GFP* detection.

The various stresses that occur during tissue culture process have been associated with genomic alterations (Choi et al., 2000), including changes in chromosome structure, DNA methylation, and DNA sequence (reviewed in Phillips et al., 1994). The stress imposed by tissue culture 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 our callogenic and organogenic tissue culture system to evaluate transgene stability during regeneration. Because selection and cocultivation with *Agrobacterium* were not carried out, this treatment 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. PCR amplification with different pairs of primers confirmed that both *GFP* and *BAR* genes had been lost in those subevents, but present in their corresponding parental events and expressing

subevents. Physical loss of transgenes during sexual propagation has been reported in annual plants (Srivastava et al., 1996; Joersbo et al., 1999; Feldmann et al., 1997). The three events that gave rise to these four subevents all contained a single copy gene; it is therefore unlikely that the loss of transgenes was caused by intrachromosomal recombination between transgenic copies integrated at different loci resulting in elimination, 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 explantation for transgene loss. Although transgene loss occurred somatically, it could also be the result of a yet unkown mechanism responsible for the non-mendilian inheritance observed in the *Arabidopsis hothead* mutant (Lolle et al. 2005) and in varieties of flax under certain environmental conditions (Chen et al, 2005). At any rate, if reintroduction of primary transformants into tissue culture process is desired, it is advisable to keep them under selection pressure due to the possibility of physical loss of transgenes.

Approximately 3% of subevents had elevated *GFP* expression. In particular, three MAR-containing subevents derived from two events appeared to maintain elevated expression over time. It was not determined whether a similar expression change also happened to *BAR* transgene in these three subevents as *BAR* expression was not determined for clone 353-53. Approximately 1% of events showed elevated *BAR* expression in the first year, but this was not maintained in subsequent years in the field.

In conclusion, we observed 1% of the transgenic events with unstable *GFP* expression and 2% of the events with unstable *BAR* expression over time which can not be explained by chance alone. After an additional round of organogenesis, physical loss of the transgenes occurred at a frequency of 0.025 at the subevent level and 0.038 at the event level; significant expression change

occurred at 2% for *GFP* and 1% for *BAR*, which was not explained by chance alone associated with the deployed p value (0.01).

Effect of MARs on transgene expression and stability

Most previous studies of MARs on transgene expression indicated that MARs can increase transgene expression to a varying degree, but gives little or no reduction in variance of expression (reviewed in Allen et al., 2000). We found that MARs from the tobacco root expressed gene RB7 did not cause expression levels for the two transgenes to change in a consistent manner. Mean expression of *GFP* was slightly lower in MAR-containing transgenic events, while *BAR* expression was significantly lower in MAR-containing events. The result contrasts with a previous study, where highly elevated transgene expression levels (10 fold) due to the RB7 MARs was observed in poplar (Han et al., 1997). MAR effects can be strongly influenced by developmental state and tissue studied. (reviewed in Allen et al., 2000; Ülker et al., 1999). In the previous study by Han et al. (1997), MAR effect was studied with stem tissues of in-vitro grown plants, and the present study used leaf tissues of field grown plants. MARs might act differently among different tissues as in the same previous study MAR had a much smaller effect on *GUS* expression in tobacco, whose leaf tissues were assayed. In mouse a large MAR effect on transgene expression has been observed in embryonic cells, but a much a smaller effect was seen in the differentiated adult tissues (Thompson et al., 1994). Promoter-dependent effect has been also reported in several studies (Sidorenko et al., 2003; Mankin et al., 2003). In addition, because of the stringency of basta selection, our initial selection of primary transformants on herbicide-containing medium might have removed some of the silenced or weakly expressed events from our study, and therefore contributed to observed increase in *BAR*, but not *GFP*, expression in

the groups without MARs. In addition, results can be biased when a small number of transgenic events are used for evaluating MAR effect.

MARs appeared to significantly reduce variance in expression of both transgenes. The effect was largely due to a reduction in the number of weakly expressing events by MARs. The transgenic events with no detectable *GFP* expression all lacked MARs. If these four events are removed, the effect of MARs on variance was not statistically significant. Mankin et al (2003) also reported that the number of low-expressing *GUS* transformants was greatly reduced when MAR-flanked constructs were used in bombardment transformation of tobacco. Similar results were obtained in studies of bombardment transformation of tobacco and rice (Vain et al. 1999; Ülker et al. 1999). Another intriguing effect of MARs found in present study was to prevent the integration of inverted repeats at the same locus, which might also reduce the number of weakly expressing events. The frequency of IR integration was 0.41 when the MARs were not present and 0 when MARs were present, and this difference is statistically significant.

Our finding that MARs increased the correlation in expression of two reporter transgenes *GFP* and *BAR* assembled in the same binary vector supports the hypothesis that MARs reduce the influence of nearby host genomic environment on transgene expression (Bode et al., 1996). A similar result was also reported by Mlynárová et al.(2002) in tobacco transgenic lines transformed with *Agrobacterium tumefaciens*. The study investigated coordinated expression of *GUS* and LUC transgenes assembled in the same T-DNA, and found that expression of the two genes, when flanked by the chicken lysozyme A element, showed high correlation. However, the frequent truncation of the MAR sequence observed in our transgenic events that contain multiple copies suggests that the presence of MARs led to a reduction in the truncations of transgenes, and also reduced effective gene silencing. This is consistent with previous

observation that MARs increased transformation efficiency (Han et al., 1997; Shimizy et al., 2001; Petersen et al., 2002), and supports the explanation that MARs increased the correlation in expression between *GFP* and *BAR* transgenes by reducing occurences where one of the transgenes is truncated. Therefore, the presence of flanking MARs might simply provide extra sequence that protects transgenes from truncations that occur during integration process. The extensive variation among transformants containing MARs in expression suggests that they have little, if any, direct role in buffering from chromosomal position effects.

Transgene copy number and its effect on transgene expression and stability

When it is desired to estimate integrated transgene copy number for a large number of samples, the traditional Southern blot method is very time-consuming and labor-intensive. Our results illustrate high levels of accuracy of TaqMan® assay based on high consistency among different replicates and runs, and high correlations with Southern blot results. The limitations of TaqMan® assay were also evident. There were cases where calculated copy number was in the middle of two integers, and therefore difficult to assign to a copy number class. For these, new real-time PCR runs or other assays, such as Southern blot are needed for confirmation. Truncations of T-DNA may have caused some of the problems in copy number determination. When one of multiple copies is truncated, quantitative PCR might fail to detect the truncated copy. This is also true for Southern blots, which will either fail to detect or produce weaker hybridization signals from a partial truncation. T-DNA truncations can often happen when multiple copies of inserts are present (see the section of T-DNA repeats). When there are no truncations, all samples should only produce a single hybridization

band from *Hind*III digestion. The presence of additional bands indicates that one of the *Hind*III sites on the T-DNA was lost as a result of truncations. This was the case for three events with multiple copies. It is therefore advisable not to choose PCR primers or hybridization probes from the sequence close to the two T-DNA borders. Quantitative PCR is a better choice than Southerns when more than two 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 actual number of inserts, and measuring band intensity lacks precision.

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 (Hobbs et al., 1993; McCabe et al., 1999; van der Hoeven et al., 1994; Voelker et al., 1996; Tang et al., 2003), negatively correlated (Hobbs et al., 1993; Cervera et al. 2000; Hobbs et al. 1993; Mannerlöf et al. 1997), or uncorrelated (Bauer et al., 1998; Hobbs et al., 1993; Mc-Cabe et al., 1999). Our study showed a positive and strongly significant, but weak correlation between transgene expression and copy number for both *GFP* and *BAR*. Contrary to expectation (reviewed in Stam et al., 1997), copy number was not correlated with initial transgene silencing nor with stability under organogenesis or over time. Out of four events which were initially silenced for *GFP* expression, three contained one copy, and the other one had three copies. All four subevents that lost their transgenes during organogenesis were derived from single copy transgenic events. The conventional wisdom that simple, single copy insertion events should be the ones chosen for research or commercial purposes was not supported in our study of vegetatively propagated poplars. However, larger copy numbers than can occur in bombardment transformation, could behave differently than the one to four copy range 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, although the effect was not statistically significant for the *BAR* transgene. There was also a significant interaction effect of MAR with copy number with respect to expression. The highest variance was seen in transgenic events containing multiple copies without flanking MARs. However, such effect was not seen in the *BAR* transgenic population, a possible consequence of its different selection history (discussed above).

It is generally believed that homologous promoters should be avoided, as they can lead to transgene and/or the resident 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 the use of 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 the maize ubiquitin promoter, rice glutelin promoter, rice actin promoter, rice LHCP (light harvesting chlorophyll a/b-binding protein of photosystem II) promoter, and rice cytochrome C promoter (summarized in Sunilkumar et al., 2005). The study by Sunilkumar et al. (2005) further addressed this issue by testing 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 of multiple copies of transgenes and effect on expression and stability

Multiple copies of transgenes can be arranged at one locus as direct, inverted, or combinations of both types of repeats (summarized in Krizkova and Hrouda, 1998). The presence of repeated structures has been reported to be accompanied by transgene silencing at transcriptional or post-transcriptional levels (reviewed in Vaucheret and Fagard, 2001). Out of 40 two-copy transgenic events, 32.5% carried direct repeats in the same locus, and 5% had inverted repeats arrainged as head-to-head. To our surprise, transformants carrying direct repeats at one locus had significantly elevated transgene expression compared to those with two copies inserted at the different loci. Inverted DNA repeats have been demonstrated to be associated with TGS and/or PTGS (Muskens et al., 2000). Inverted repeat structure was consistently associated with T-DNA truncations in our study, but the two events carrying inverted repeats (head-to-head) had stable transgene expression. Inverted and direct repeat arrangements or T-DNA truncations were also not sufficient to induce transgene silencing in *Arabidopsis* (Lechtenberg et al., 2003; Meza et al., 2002).

Unexpectedly, all transformants containing three or four copies had repeats with at least two copies 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 of copy number and transgene expression, repeat formation did not reduce expression. It also did not increase instability over time.

GENERAL CONCLUSIONS

Transgene silencing, both initial and subsequent silencing over time, appears to be a minor concern in *Agrobacterium* transformation in poplar. Due to the wide range but consistency of transgene expression, it is advisable to perform early screening to select desired expression levels. Once selected, expression should usually remain consistent in field environments. MARs appear to be useful in transgenic constructs because they lead to reduction in variance of transgene expression, number of weakly expressing events, truncations of transgenes, formation of inverted repeats, and increased correlation between two transgenes in the same t-DNA. We found a high percentage of single copy integration of T-DNAs, a positive correlation between transgene copy number and expression, a high percentage of strongly expressing direct repeats with multiple insertions. Selection of transformants on expression level alone therefore appears to provide an accurate prediction for future expression without detailed molecular characterization.

Although high stability of transgene expression was achieved in our study after two seasons in the field with different coding genes, promoters, constructs, and poplar clones, long-term stability might still need to be investigated further. However, due to current stringent regulations, and the many years needed to study transgenic trees, evaluation of long-term stability is difficult. In addition, when RNAi technology is used to suppress endogenous flowering genes for the purpose of bioconfinement or control of reproductive development, stability of gene suppression might need to be evaluated using the actual flowering genes. In this case, long-term stability evaluation is even more critical because expression of endogenous flowering genes is generally differentially associated with different phase changes of a long-lived tree. Stability of transgene expression can be very different when sexual propagation is used as meiosis is largely known to promote transgene inactivation. It is

challenging, but not impossible, to study meiotic stability of transgene expression in poplar, especially when early flowering GE trees can be developed.

ACKNOWLEDGEMENTS

We would like to thank Dr. C. Neal Stewart for the gift of *mgfp5er* gene, Dr. Steven Spiker for kindly providing MAR, and Dr. John Davis for the use of poplar *rbc*S cDNA. We also would like to thank Dr. Lisa Ganio for statistical advice, and Dr. Christopher J. Bayne for sharing the microtiter plate reader. The research project was supported by Initiative for Future Agriculture and Food Systems Grant no. 00-52100-9623 from the USDA Cooperative State Research, Education, and Extension Service, TBGRC/TGERC Research Cooperative, and OSU Foundation Forest Biotechnology Fund.

REFERENCE

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.

Allen, G.C., G.E. Hall, Jr., S. Michalowski, W. Newman, S. Spiker, A.K. Weissinger and W.F.Thompson. (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco. Plant Cell 8: 899–913.

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

Alonso, J.M., A.N. Stepanova, T.J. Leisse, C.J. Kim, H. Chen, P. Shinn, D.K. Stevenson, J. Zimmerman, P. *BAR*ajas, R. Cheuk. (2003) Genome-wide insertional mutagenesis of *Arabidopsis* thaliana. Science. 301: 653-657.

Amedeo, P., Y. Habu, K. Afsar, S.O. Mittelsten, and J. Paszkowski. (2000) Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. Nature. 405:203-206.

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

Baulcombe, D. (2004) RNA silencing in plants. Nature 431: 356-363.

Bender, J. (2001) A vicious cycle: RNA silencing and DNA methylation in plants. Cell 106:129-132.

Bode, J., M. Stengert-Iber, T. Schlake, and A. Dietz-Pfeilstetter. (1996) Scaffold/matrix-attached regions: topological switches with multiple regulatory functions. Crit. Rev. Euc. Gene Exp. 6: 115-138.

Boerjan, W. (2005) Biotechnology and the domestication of forest trees. Curr Opin in Biotechnol. 16: 159-166.

Bradshaw, H.D., Jr, R. Ceulemans, J. Davis, R.F. Stettler. (2000) Emerging model systems: Poplar (*Populus*) as a model forest tree. J. Plant Growth Regulators 19: 306–313.

Brouwer, C., W. Bruce, S. Maddock, Z. Avramova, and B. Bowen (2002) Suppression of transgene silencing by matrix attachment regions in Maize: a dual role for the Maize 5' *ADH1* matrix attachment region. Plant Cell 14: 251 – 2264.

Campbell, M.M., A.M. Brunner, H.M. Jones, and S.H. Strauss. (2003) Forestry's Fertile Crescent: The application of biotechnology to forest trees. Plant Biotechnol. J. 1: 141-154.

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, S., W. Jin, M. Wang, F. Zhang, J. Zhou, Q. Jia, Y. Wu, F. Liu, and P. Wu. (2003) Distribution and characterization of over 1000 T-DNA tags in rice genome. Plant J. 36: 105-113.

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.

Chuang, C.F. and E.M. Meyerowitz. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis* thaliana. Proc. Natl. Acid. Sci. USA, 97: 4985-4990.

Conner, A.J., L. Mlynárová, W.J. Stiekema, and J. Nap. (1998) Meiotic stability of transgene expression is unaffected by flanking matrix-associated regions. Mol Breed. 4:47-58.

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

FAO. (2004) Preliminary review of biotechnology in forestry, including genetic modification. Forest Genetic Resources Working Paper FGR/59E. Forest Resources Development Service, Forest Resources Division. Rome, Italy.

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.

Fernandez-Cornejo, J. and C. Margriet. (2006) The First Decade of Genetically Engineered Crops in the United States U.S. Department of Agriculture, economic Research Service, Technical Bulletin No. 11. April. 2006.

Filichkin, S.A., R. Meilan, V.B. Busov, C. Ma, A.M. Brunner, and S.H. Strauss. (2006) Alcohol-inducible gene expression in transgenic *Populus*. Plant Cell Rep. (online version: DOI: 10.1007/s00299-005-0112-3).

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

Francis, K.E. and S. Spiker. (2005) Identification of *Arabidopsis* thaliana transformants without selection reveals a high occurrence of silenced T-DNA integrations. Plant J. 41: 464-477.

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

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

Gould, A.R. (1986) Factors controlling generation of variability *in vitro*. In: Cell Culture and Somatic Cell Genetics of Plants (Vasil, I.K., ed.). Vol. 3. Academic Press, Orlando, FL, USA, pp. 549-567.

Halfhill, M.D., R.J. Millwood, and 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.P. Ma, and S.H. Strauss. (1997) Matrix attachment regions (MARs) enhance transformation frequency and transgene expression in poplar. Transgenic Res. 6: 415–420.

Han, K.-H., M.P. Gordon, and S.H. Strauss. (1997) High frequency transformation of cottonwoods (*genus Populus*) by Agrobacterium rhizogenes. Can. J. For. Res. 27: 464-470.

Hansen, G., R.D. Shillito, and M.D. Chilton. (1997) T-strand integration in maize protoplasts after codelivery of an T-DNA substrate and virulence genes. Proc. Natl. Acad. Sci. USA. 94: 11726-11730.

Hansen, G. and M.D. Chilton. (1999) Lessons in gene transfer to plants by a gifted microbe. Curr. Top. Microbiol. Immunol. 240: 21-57.

Harper, B.K., S.A. Mabon, S.M. Leffel, M.D. Halfhill, H.A. Richards, K.A. Moyer, C. Stewart, and J. Neal. (1999) Green fluorescent protein as a marker for expression of a second gene in transgenic plants. Nature Biotechnol. 17: 1125-1129.

Harper, B.K. and C.N. Stewart Jr. (2000) Patterns of green fluorescent protein expression in transgenic plants. Plant Mol. Biol. Rep. 18: 141a-141i.

Hawkins, S., J. 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.

Hellens, R.P., E.A. Edwards, N.R. Leyland, S. Bean, and P.M. Mullineaux. (2000) pGreen: a versatile and flexible binary Ti Vector for *Agrobacteruim*-mediated plant transformation. Plant Mol. Biol. 42: 819-832.

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.

Iglesias, V.A., E.A. Moscone, I. Papp, F. Neuhuber, S. Michalowshi, 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.

James, V. A., B. Worland, J. W. Snape, and P. Vain. (2004) Strategies for precise quantification of transgene expression levels over several generations in rice. J. Exp. Bot. 55(401): 1307 – 1313.

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.

Kilby, N.J., H.M.O. Leyser, and I.J. Furner. (1992) Promoter methylation and progressive transgene inactivation in *Arabidopsis*. Plant Mol. Biol. 20: 103-112.

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

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.

Kooter, J.M., M.A. Matzke, and P. Meyer. (1999) Listening to the silent genes: trangene 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(6): 673-680.

Kumar, S. and M. Fladung. (2000a) Determination of transgene repeat formation and promoter methylation in transgenic plants. BioTechniques. 28: 1128-1137.

Kumar, S. and M. Fladung. (2000b) Transgene repeats in aspen: molecular characterization 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(2): 671 – 677.

Levee, V., E. Garin, K. Klimaszewska, and A. Seguin. (1999) Stable genetic transformation of white pine (Pinus strobus L.) after cocultivation of embryogenic tissues with *Agrobacterium* tumefaciens. Mol. Breed. 5: 429–440.

Li, L.C. and R. Dahiya. (2002) MethPrimer: designing primers for methylation PCRs. Bioinformatics. 8(11): 1427-31.

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.

Mankin, S.L., G.C. Allen, T. Phelan, S. Spiker, and W.F. Thompson. (2003) Elevation of transgene expression level by flanking matrix attachment regions (MAR) is promoter dependent: a study of the interactions of six promoters with the RB7 3' MAR. Transgenic Res. 12(1): 3-12.

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.

Maximova, S., C. Miller, G. Antfflnez de Mayolo, S. Pishak, A. Young, and M. J. Guiltinan. (2003) Stable transformation of Theobroma cacao L. and influenceof matrix attachment regions on *GFP* expression. Plant Cell Rep. 21: 872-883.

McCabe, M.S., U.B. Mohapatra, S.C. Debnath, J.B. 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) Stability of herbicide resistance and *GUS* expression in transgenic hybrid poplars (*Populus* sp.) during several years of field trials and vegetative propagation. HortScience 37(2): 277-280.

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

Meyer P, Linn F, Heidmann I, Meyer H, Niedenhof I, Saedler H (1992) Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. Mol. Gen. Genet. 231: 345-352.

Meza, T.J., D. Kamfjord, A. Hakelien, 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. Transgen. Res. 10: 53-67.

Meza, T.J., B. Stangeland, I.S. Mercy, M. Skårn, D.A. Nymoen, A. Berg, M.A. Butenko, A-M. Håkelien, C. Haslekås, and L.A. Meza-Zepeda. (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.

Millwood, R.J., M.D. Halfhill, D. Harkins, R. Russotti, and C.N. Stewart, Jr. (2003) Instrumentation and methodology for quantifying *GFP* Fluorescence in intact plant organs. BioTechniques. 34: 638-643.

Mlynárová, L., A. Loonen, J. Heldens, R.C Jansen, P. Keizer, W.J. Stiekema, and J.P. Nap. (1994) Reduced position effectin mature transgenic plants conferred by the chicken lysozyme matrix-associated region. Plant Cell 6: 417-426.

Mlynárová, L., R.C. Jansen, A.J. Conner, W.J. Stiekema, and J.P. Nap. (1995) The MAR-mediated reduction in position effect can be uncoupled from copy number-dependent expression in transgenic plants. Plant Cell 7: 599–609.

Mlynárová, L., A. Loonen, E. Mietkiewska, R.C. Jansen, and J. Nap. (2002) Assembly of two transgenes in an artificial chromatin domain gives highly coordinated expression in tobacco. Genetics. 160: 727–740.

Mlynárová, L., A. Hricová, A. Loonen, and J. Nap. (2003) The Presence of a chromatin boundary appears to shielda transgene in tobacco from RNA silencing. Plant Cell 15: 2203–2217.

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

Morina, K., O. Olsen, and K. Shimamoto. (1999) Silencing of an aleurone-specific gene in transgenic rice is caused by a rearranged trangsgene. Plant J. 17: 275-285.

Müller, A.J., R.R. Mendel, J. Schiemann, C. Simoens, and D. Inzé. (1987) High meiotic stability of a foreign gene introduced into tobacco by *Agrobacterium*-mediated transformation. Mol. Gen. Genet. 207: 171-175.

Paszkowski, J. and S.A.Whitham.(2001) Gene silencing and DNA methylation processes. Curr. Opin. Plant Biol. 4: 123-129.

Petersen, K., R. Leah, S. Knudsen, and V. Cameron-Mills. (2002) Matrix attachment regions (MARs) enhance transformation frequencies and reduce variance of transgene expression in barley. Plant Mol. Biol. 49: 45–58.

Phillip, 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.

Risseeuw, E., M.E. Fanke-van Dijk, and P.J. Hooykaas. (1997) Gene targeting and instability of *Agrobacterium* T-DNA loci in the plant genome. Plant J. 11: 717-728.

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(2): 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.

Rozen, S. and H.J. Skaletsky. (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.

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

Schimizu, K., M. Takahashi, N. Goshima, S. Kawakami, K. Irifune, and H. Morikawa. (2001) Presence of an SAR-like sequence in junction regions between an introduced transgene and genomic DNA of cultured tobacco cells: its effect on transformation frequency. The Plant Journal. 26(4): 375-384.

Schöffl, F., G. Schroder, M. Kliem, and M. Rieping. (1993) An SAR sequence containing 395 bp DNA fragment mediates enhanced, gene-dosage-correlated expression of a chimaeric heat shock gene in transgenic tobacco plants. Transgen. Res. 2: 93-100.

Sedjo, R.A.. (2004) Genetically engineered trees: promise and concerns. Resources for the future. http://www.rff.org/Documents/RFF-RPT-Forest.pdf.

Sidorenko, L., W. Bruce, S. Maddock, L. Tagliani, X. Li, M. Daniels, and T. Peterson1. (2003) Functional analysis of two matrix attachment region (MAR) elements in transgenic maize plants. Transgen. Res. 12: 137-154.

Sijen, T., I. Vijn, A. Rebocho, R. van Blokland, D. Roelofs, J.N. Mol, and J.M. Kooter. (2001) Transcriptional and post-transcriptional gene silencing are mechanistically related. Curr. Biol. 11: 436-440.

Somers, D.A., and I. Makarevitch. (2004) Transgene integration in plants: poking or patching holes in promiscuous genomes. Curr. Opin. Biotechnol. 15(2): 126-131.

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

Stam, M., N.M. Joseph, and J.M. Kooter. (1997) The silencing 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.

Strauss, S.H., S.P. DiFazio, and R. Meilan. (2001) Genetically modified poplars in context. Forestry Chronicle. 77(2): 271-279.

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.

Thompson, E.M., E. Christians, M.G. Stinnakre, J.P. Renard. (1994) Scaffold attachment regions stimulate HSP70.1 expression in mouse preimplantation embryos but not in differentiated tissues. Mol. Cell Biol. 14: 4694-4703.

Ülker, B., G.C. Allen, W.F. Thompson, S. Spiker, and A.K. Weissinger. (1999) A tobacco matrix attachment region reduces the loss of transgene expression in the progeny of transgenic tobacco plants. Plant J. 18: 253-263.

Vain, P., B. Worland, A. Kohli, J.W. Snape, P. Christou, G.C. Allen, and W.F. Thompson. (1999) Matrix attachment regions increase transgene expression levels and stability in transgenic rice plants and their progeny. Plant J. 18: 233-242.

Van der Geest, A.H.M., G.E .Hall, Jr., S. Spiker, and T.C. Hall. (1994) The beta-phaseolin gene is flanked by matrix attachment regions. Plant J. 6: 413-423.

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

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, and D.J. Hawkins. (1992) Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257: 72–74.

Chapter 3: EFFICIENCY AND STABILITY OF TRANSGENE SUPPRESSION VIA RNAi IN FIELD GROWN POPLARS

ABSTRACT

RNAi induced by double-stranded RNA (dsRNA) was studied in 56 independent poplar transgenic events in the field over two years. A resident *BAR* transgene was targeted with two different types of RNAi constructs: a 475 bp inverted repeat of the promoter sequence and a 275 bp inverted repeat of the coding sequence. A Matrix Attachment Region (MAR) from the tobacco RB7 gene flanked two of the four transgenic constructs. RNAi directed at the coding sequence was a strong inducer of gene silencing; 80% of the transgenic events showed more than 90% suppression of *BAR* expression. In contrast, RNAi targeting the promoter resulted in only 6% of transgenic events showing more than 90% suppression. The MAR had no detectable effect on RNAi suppression and stability. Seasonal change of RNAi silencing was also investigated with two highly suppressed events.The degree of gene suppression was highly correlated over two years in the field, and over seasonal development within a year. Copy number of the integrated RNAi transgene was not associated with RNAi silencing levels. DNA methylation of the homologous promoter region was associated with high degree of gene suppression induced by IR of the promoter sequence. DNA methylation was also observed in the coding region of some, but not all, of highly suppressed events containing IR of the coding sequence.

INTRODUCTION

RNA interference (RNAi) is homology-dependent gene silencing induced by double-stranded RNA (dsRNA). RNAi was first discovered in the nematode *Caenorhabditis elegans* (Fire et al., 1998), and then demonstrated in *Drosophila*, *Tryphnosoma*, plants, and vertebrates (reviewed in Plasterk and Ketting, 2000). Recent findings have shown that RNAi-related gene silencing is an evolutionarily conserved gene-regulatory mechanism used for cellular defense, RNA surveillance, and in guiding development (Vance and Vaucheret, 2001; Kusaba, 2004). Current models for pathways of RNAi silencing all involve the cleavage of dsRNA into short interfering RNAs (siRNAs) or microRNAs (miRNAs) (Baulcombe, 2004). Those small RNA act as guides to direct cleavage or translational repression of complementary mRNAs to induce post-transcriptional gene silencing (PTGS), or cause DNA and chromatin modification to induce transcriptional gene silencing (TGS) (Lippman and Martienssen, 2004).

Artificial introduction of dsRNA-producing transgenes in plants has been shown to induce a high degree and frequency of sequence-specific gene silencing (Chuang and Meyerowitz, 2000; Waterhouse et al., 1998; Stoutjesdijk et al., 2002; Kerschen et al., 2004; Wagner et al., 2005). Constructs designed to express hairpin RNA (hpRNA) are more efficient inducers of silencing (Wesley et al., 2001; Smith et al., 2000; Stoutjesdijk et al., 2002). In an hpRNA-producing vector, partial or full sequences of target genes are placed as inverted repeats (IR) separated by an unrelated sequence (Kusaba, 2004). When used as a spacer, intron sequence has been shown to increase stability and efficiency of RNAi (Wesley et al., 2001). In addition to coding sequence, dsRNA of promoter sequences can result in TGS accompanied by *de novo* methylation of homologous sequences (Mette et al., 2000; Cigan et al., 2005). Therefore, RNAi

technology offers a new avenue to genome-wide study of gene function, and to crop improvement via directed, dominant gene suppression.

One of major concerns in application of genetic engineering to forest trees is the potential ecological impacts on wild or managed ecosystems from dispersal of pollen or seeds from transgenic plantations (DifaZio et al., 2004). Biological confinement mechanisms such as sterility genes may therefore be necessary before the commercial release of some forms of genetically engineered trees (Strauss et al., 1995). RNAi technology offers a potentially powerful tool to genetically engineer floral sterility by inhibiting expression of one or multiple genes required for the onset of flowering and production of fertile gametes. RNAi or antisense constructs has been extensively used to modify flowering in scientific research (e.g., Mohamed, 2006; Yan et al., 2004).

Although there have been numerous studies of RNAi silencing mechanisms, there is little information on the efficiency or long-term stability of RNAi in plants. In *Arabidopsis*, the high degree of silencing of the endogenous gene Δ12-desaturase gene (*FAD2*) induced by the intron-spliced hairpin transgene was studied in one transgenic line over five sexually propagated generations, and no reversion or reduction of gene silencing was observed (Stoutjesdijk et al., 2002). Temperature-dependent gene silencing has been reported both in plants and animals (Szittya et al., 2003; Sós-Hegedűs et al., 2005; Fortier and Belote, 2000, Kameda et al., 2004). In *Nicotiana benthamiana*, both virus and transgene (35S-*GFP*/35S-ds*GFP*) triggered RNA silencing were inhibited at a low temperature (15 ^{0}C), and the loss of silencing was accompanied by the reduction of siRNA (Szittya et al. 2003). This study also reported that temperature affected antisense-mediated endogene inactivation in *Arabidopsis* and potato. However, Sós-Hegedűs et al. (2005) reported both temperature dependence and independence of antisense-mediated gene silencing. Out of 24 potato antisense transgenic lines, nine lines were not influenced by the

low temperature (15 ^{0}C) with respect to antisense induced gene silencing. Those studies suggest that RNAi suppression may be unstable in field environments, where temperatures can vary dramatically.

Reported studies of RNAi silencing in perennial woody plants are relatively rare. In the gymnosperm species *Pinus radiata,* cinnamyl alcohol dehydrogenase (*CAD*), a gene associated with the biosynthesis of lignin, was more efficiently suppressed by an inverted repeat of the *CAD* cDNA than by either sense or antisense transgenes introduced via biolistic transformation (Wagner et al., 2005). It was also found that suppression level was positively associated with the expression level of the transgene. In the angiosperm species poplar, RNAi suppression of the endogenous flowering gene *PCENL1* caused early flowering in the field (Mohamed, 2006). These two studies suggest that RNAi can be successfully used in diverse tree species, and therefore, can be a valuable tool for introducing commercially useful traits.

Before RNAi technology can be successfully used to induce floral sterility for the purpose of bioconfinement, it must be demonstrated that RNAi is a strong inducer of gene silencing for genes required for fertility and that RNAi-induced gene silencing can be maintained over multiple years in the field. Here we reported the first study of stability of RNAi in poplar trees in a field environment. We transformed *BAR* transgenic poplars with four kinds of intron-spliced hpRNA (ihpRNA) constructs containing inverted repeats (IR) directed at promoter or coding sequence. We also evaluated how flanking MARs influenced RNAi efficiency and stability. We report that RNAi directed at the coding sequence, but not promoter, was strong and stable over years, and that MAR elements had no influence on efficiency or stability.

MATERIALS AND METHODS

Construction of RNAi stability vectors

Gene- and promoter-specific sequences in sense and antisense orientation were assembled in the suppression vector pHANNIBAL (Wesley et al., 2001; provided by Dr. P. Waterhouse, CSIRO, Australia). For the construction of inverted repeats (IR) of coding sequence, a 275 bp sequence (Appendix 3B) was amplified using forward primer 5' CACTCGAGGGTCTGCACCAT CGTCA AC 3', and reverse primer 5' TGGGTACCTCAGCAGGTGGGT GTAGAG 3', with *Xho*I and *Kpn*I sites (underlined) introduced in the forward and reverse primers, respectively. The same fragment was again amplified with the primers with different restriction sites, *Bam*HI and *Cla*I, introduced: 5' AGGGATCCGGTCTG CACCATCGTCAAC 3' and 5' CATCGATTCA GCAGGTGGGTGTGAG 3'. Both fragments were cloned into pHANNIBAL to create an IR. Sense and antisense fragments of the *Arabidopsis rbc*S promoter sequence were assembled in the same way. A 475 bp fragment (Appendix 3C) was amplified with the primers: 5' GGCTCGAGAT ATATTCCACAGTTT CACC 3', and 5' CCGGTACCTTGG TCTAGTGCTTT GGTCA3', and cloned into *Xho*I and *Kpn*I sites of the pHANNIBAL to give the sense orientation. The same fragment amplified with the primers 5' ACGGAT CCATATATTCCACA GTTTCACC 3' and 5' CCATCGATTTGGTCTAGT GCTTTGGTCA 3' was cloned into the *Bam*HI and *Cla*I sites to give the antisense orientation.

The binary vector pGreenII (Hellens et al., 2000) was used for assembling the selectable marker *CP4* cassette and RNAi cassettes. Two versions of the vector were used: one with flanking MAR element derived from the tobacco RB7 gene (ref? ; provided by S. Spiker, North Carolina State University), and one without MARs. For cloning of MARs, a 1,167 bp MAR fragment was cut from the vector pHK10 with the restriction enzymes *Not*I and

*Spe*I, filled in with T4 DNA polymerase, and cloned at *Fsp*I and *Sap*I (blunted) sites of the pGreenII to produce pG3M.

The selectable marker *CP4* cassette has been previously described (Meilan et al., 2002b). It consists of the FMV-34S promoter from figwort mosaic virus, the coding region of the *CP4* gene from *Agrobacterium tumefaciens* conferring resistance to herbicide glyphosate, and the 3' untranslated end from small subunit of ribulose bisphosphate carboxylase from pea (Coruzzi et al., 1984). A 3,044 bp cassette was dropped out with *Pme*I and *Not*I from the vector pMON17227, blunted with T4 DNA polymerase, and cloned into the *Sma*I site of pGreenII to produce pG3C, and the *Xho*I (blunted) and *Eco*RV sites of pG3M to produce pG3MC.

The IR directed at coding sequence, together with the 35S promoter and octopine synthase terminator (OCS), were cut from pHANNIBAL with *Not*I and cloned into the corresponding site of the pG3C to produce pG3CBi (BiM-, Fig. 3.1A). The same cassette was cut with *Spe*I, and cloned into the *Spe*I site of pG3MC to produce pG3MCBi (BiM+, Fig. 3.1B).

The IR directed at the promoter sequence, together with 35S promoter, were cut from pHANNIBAL with *Spe*I and *Sac*II and cloned into the same sites of the pG3C and pG3MC to produce pG3CPi (PiM-) and pG3MCPi (PiM+), respectively.(Fig. 3.1,C&D).

Fig. 3. 1. RNAi constructs used to silence a *BAR* transgene. All constructs contain backbone and border sequence from the pGreenII binary vector, and the selection gene *CP4* driven by the FMV-34S promoter. A: RNAi construct with inverted repeats (IR) of 275 bp of *BAR* coding sequence. B: RNAi construct with IR of 275 bp of *BAR* coding sequence with flanking MARs. C: RNAi construct with IR of 475 bp of *rbcS* promoter sequence. D: RNAi construct with IR of 475 bp of *rbcS* promoter sequence with flanking MARs. Abbreviations: E9t: terminator from small subunit of RUBP carboxylase from pea; CP4: EPSPS gene from *Agrobacterium tumefaciens*; 34Sp: promoter from figwort mosaic virus; OCSt: octopine synthase terminator; *BAR*: 275 bp coding sequence of *BAR* from *Streptomyces hygroscopicus*; PDKin: pyruvate orthophosphate dikinase intron; 35Sp: 35S promoter from cauliflower mosaic virus; *rbc*Sp: 475 bp of promoter sequence from the *rbc*S gene of *Arabidopsis*. T-DNA regions are not drawn to scale.

Plant transformation and field trials

Four *BAR*-containing transgenic poplar events (we refer to these four events as parent events) that were previously generated in the laboratory via *Agrobacterium* transformation were used to transform RNAi constructs. The four parent events are 353-38, 353-29, 717-A, and 717-C. 353-38 and 353-29

were made in hybrid poplar clone 353-53 (*P. tremula x P. tremuloides*), and the other two in hybrid clone 717-1B4 (*P. tremula* x *P. alba*). The original transgenic poplars had been transformed with the binary plasmid pTTM8 provided by Plant Genetic Systems (Belgium) and contained in the T-DNA: two chimeric genes conferring resistance to the antibiotic kanamycin and the herbicide glufosinate-ammonium, respectively, and another chimeric *BARNASE* gene that can impart male-sterility (Appendix F3.1). The herbicide resistance gene was driven by the promoter of a strongly photosynthesis-associated gene, the small subunit of ribulose bisphosphate carboxylase derived from *Arabidopsis*.

The four events were retransformed with *Agrobacterium tumefaciens* strain C58/pMP90 (GV3101) harboring RNAi constructs as described previously (Filichkin et al., 2006), except that glyphosate selection was used with a concentration of 0.4 mM for callus and shoot induction, and 0.012 mM for root induction. The presence of the sense fragments in the transgenic events was confirmed with a forward primer annealing to the 35S promoter and a reverse primer annealing to the PDK intron. The primers used were: 5' GCACA ATCCCACTATCCTTCGCAAG 3' and 5' GATAGATCTTGCGCTTTGTTA TATTAGC 3'. The presence of the *BAR* antisense fragment was confirmed with primers to the antisense fragment and the OCS terminator: 5' TAGTGGTTGAC GA TGGTGCAGACCG 3' and 5' CGGCCAATACTCAACTTCAAGGAATC TC 3'. The presence of the *rbc*S promoter antisense fragment was confirmed with primers to the antisense fragment and the 34S promoter: 5' GATTAGGTG AAACTGTGGAATATATGG 3' and 5' GCGCCTAACA ATTCTGCACC 3'.

Six to eight ramets of each individual transgenic event and non-transformed parent event (NT) were propagated in vitro and transferred to the greenhouse. Together with the NT controls, four ramets of each transgenic event were coppiced and planted at a field site near Corvallis, Oregon in June 2004

(Fig. 3.2). The plants were distributed in four blocks with a random split-block design because of inherent differences in rate of height growth of the two clones, and transgenic events in the 353-53 clone were separated from those in the 717-1B4. Trees were randomly assigned to one of two subplots within each block.

Fig. 3. 2. Field trial of RNAi transgenic poplar. The trees were planted in June 2004, and the picture was taken in September 2005.

Gene expression measurements

RNA extraction

Young leaves were sampled from all field plants in 2004 and 2005 and stored at -80 ^{0}C until assayed. RNA was extracted for two ramets of each individual event using the RNeasy Mini Kit (Qiagen, Valencia, CA; Cat # 74106) with modifications as described in Brunner et al. 2004c (Appendix 3D).

cDNA synthesis

Extracted RNA was treated with DNAse (Ambion, Austin, TX, Cat # 1906) to remove residual genomic DNA following the manufacturer's instructions. For the synthesis of first-strand cDNA, 1 µg of total RNA was reverse transcribed using oligo(dT) and random primers with SuperScript First Strand synthesis kit for RT-PCR (Invitrogen, Carlsbad, CA; Cat # 12371-019) in a volume of 20 µl.

Real-time PCR

Real-time PCR was performed with SYBR green (Platinum SYBR Green qPCR Super Mix UDG, Invitrogen, Carlsbad, CA; Cat # 11733-046) in a Mx3000 P real-time PCR machine (Stratagene, La Jolla, CA). An endogenous ubiquitin (*UBQ*) gene was used as a reference gene. We chose the *UBQ* gene as the internal control because poplar *UBQ* is stably expressed in a variety of tissues (Brunner et al., 2004c). Primers used for amplification of the targeted *BAR* gene were: 5' TTTCTGGCAGCTGGACTTCAG 3' and 5' ATCCTAGAACGC GTGATTCAGATC 3', with a product size of 84 bp. The *UBQ* gene was amplified with primers 5' GTTCAATGTTTCGTTTCATG 3' and 5' TAACAGGAACG GAAACATAG 3', giving a product size of 100 bp.

For each sample duplicated wells were run for both target gene and the reference gene. Synthesized cDNA of the parent event 717-C was serially diluted 5-fold and run in triplicate to create a standard curve. For every reaction, 2 µl of 50 x diluted first strand cDNA was used in a total volume of 25 µl. The cycle conditions were the same for all PCR: 50 ^{0}C for 2 min, 95 ^{0}C for 2 min, and 40 cycles of: 95 ^{0}C - 15 s, 56 ^{0}C - 30 s, and 72 ^{0}C - 30 s. For each transgenic event, the relative expression level of the *BAR* gene was calculated with

MX3000P RT-PCR software (version 2) (Stratagene, La Jolla, CA) based on the standard curves and normalized to *UBQ* quantities.

Seasonal development of gene suppression

Tissue samples (bud or leaf) were taken at different seasonal points for two highly suppressed transgenic events and one non-transformed parent event (353-29) to study seasonal changes in gene expression due to RNAi silencing. Samples were freshly frozen in liquid nitrogen, and stored at - 80 ^{0}C until analyzed.

Molecular characterization of transgenes

Transgene copy number

To estimate transgene copy number in transgenic plants, comparative real-time PCR was performed on all studied events. An endogenous, single copy gene poplar *LEAFY* (*PTLF*) was used as reference gene, which was confirmed by Southern blot (Appendix F2.11). The 35S promoter driving the IR was chosen for the copy number analysis so the same primer and probe sets could be used for all four constructs. Real-time PCR was performed using dual-labeled TaqMan® probes (Biosearch Technologies). All primers and probes were designed using the program Primer 3 (Rozen and Skaletsky, 2000; http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). For the 35S promoter, the primers used for amplification were 5' CAGTGGTGAAC ATAGTGTCG 3', and 5' TACAATGGACGATTTCCTCT 3'. The hybridization probe was labeled with HEX and BHQ-1: 5' HEX d(CACCTTCACCTTCGAACTTCCTTC) BHQ-1 3'. For the *PTLF* gene, the

primers were: 5' GGTTTCTCTGAGGAGCCAGTACAG 3', and 5' GCCTCCCATGTCCCTCTTC 3'. The hybridization probe was labeled with FAM and TAMARA: 5' FAM d(CAAGGAGGCAGCAGGGAGCGGT) TAMARA 3'. The primer and probe sets were tested for their amplification efficiency using a standard curve, and were optimized to have a high and comparable efficiency (close to 100%). For the *PTLF* gene, the optimized concentration for both forward and reverse primers was 0.4 μM, and for the probe was 0.2 μM. For the 35S promoter the optimal primer concentration was 0.5 μM for both primers,

All PCR reactions were carried out using QuantiTect Multiplex PCR buffer (Qiagen, Valencia, CA; Cat # 204545) in a volume of 25 μl. For each sample, 100 ng genomic DNA purified with the Dneasy Plant Mini Kit (Qiagen, Valencia, CA; Cat # 69106) was used in a duplexed reaction to amplify both *PTLF* and 35S targets, and run in duplicated wells. The amplification was performed in an Mx3000 P real-time PCR machine (Stratagene, La Jolla, CA) with the following cycles: 95 ^{0}C for 15 min, and 45 cycles of: 94 ^{0}C for 30 s, and 60 ^{0}C for 1 min. The threshold cycle (Ct) was determined using the MX3000P$_{TM}$ RT-PCR System software (version 2). Copy number of 35S promoter was determined by the formula: $2^{(1-\Delta Ct\ (35S\text{-}ptlf)}$. The use of comparative real-time PCR approach to estimate transgene copy number was verified by Southern blot analysis as described in Chapter II.

Methylation analysis using methylation specific PCR (MSP)

DNA methylation status of promoter and coding regions of the target *BAR* gene were investigated using methylation specific PCR (MSP). A total of 1 μg of genomic DNA was bisulfite modified with the CpGenome^{TM} DNA Modification Kit (Chemicon, Billerica, MA; Cat # S7820) following the manufacturer's

instructions. In the bisulfite reaction, all unmethylated cytosines are converted to uracils, while methylated cytosines remain unaltered. Specific primers were designed to distinguish methylated from unmethylated DNA. Methylation specific primers for both *rbc*S promoter and *BAR* coding regions were designed using the program MethPrimer (Li and Dahyia, 2002) based on the sense strand. M primers used for amplification of methylated *rbc*S promoter strand were: 5' CGGAAAAGGTATAAGTAAAATATTTAATTC 3' and 5' TAAACCGCTA AAATAATACCACGT 3'. U primers used for amplification of unmethylated *rbc*S promoter sense strand were: 5' TGGAAAAGGTATAAGTAAAATAT TTAATTT 3' and 5' CTTAAACCACTAAAATAATACCACAT 3'. M and U primers amplified the same region of the *rbc*S promoter, and yielded a product size of 726 and 728 bp, respectively (Appendix 3C). M primers used for amplification of the methylated *BAR* coding sequence were: 5' GAACGACGTT CGGTCGATATTCGTC 3', 5' AAAACGTAAAACCCAATCCCGTCCG 3', and a degenerated primer 5' AAACGTAAAACCCAGTCCCGTCCG 3' was used in case cytosine in the sequence CAG was methylated. U primers used for the amplification of the unmethylated *BAR* sense strand were: 5' AATGATGTT TGGTTGATATTTGTTG 3' and 5'AAACATAAAACCCAATCCCATCCAC 3'. Both pairs of the primers amplified the same region of *BAR* coding sequence, and amplified a product size of 309 bp (Appendix 3B).

All PCR amplifications were performed in a volume of 50 µl with the following reagents: 2 µl bisulfite modified DNA, dNTPs (1.25 mM), MgCl2 (6.75 mM), Plantium Taq (2 U, Invitrogen, Carlsband, CA), primers (each 0.4 µM), and 1x PCR buffer. The PCR cycles were: 94 ^{0}C for 2 min; 39 cycles of: 94 ^{0}C - 20 s, 59 ^{0}C - 20 s, and 68 ^{0}C - 40 s / kb; 72 ^{0}C for 3 min. The PCR products were purified with the QIAquick PCR Purification Kit (Qiagen, Valencia, CA; Cat # 28104), and were sequenced using PCR primers in the ABI

3730 sequence machine at the Center for Genome Research and Biocomputing (CGRB) at OSU.

Statistical analysis

Quantified expression values of the RNAi transgenics were normalized to that of the corresponding parent event. Stability of RNAi suppression across years was analyzed with regression analysis of log2 transformed values. Difference in efficiency of RNAi suppression between different IR transgenes was studied using Fisher's Exact Test.

RESULTS

RNAi silencing efficiency with different types of dsRNA

Four transgenic events (parent events) containing the resident *BAR* transgene were demonstrated to confer high, stable resistance to herbicide glufosinate over multiple years in the field (data not shown). These four parent events were retransformed with the four RNAi constructs. A total of 78 transgenic events were generated from the 16 (4 x 4) construct x line combinations. The number of produced events was smaller than originally planned due to unexpectedly low transformation efficiency (ranging from 0 ~ 1%; Appendix T3.1)), a possible consequence of "leaky" expression of the *BARNASE* gene in the transgenic parent lines. This is much lower than the efficiencies (ranging from 5~37%) obtained in other transformation projects using similar protocols (Meilan and Ma, unpubl. data). Among the 76 RNAi transgenic events, 16 events (20%) contained only an antisense fragment, and the sense fragment was absent. Those events were excluded from RNAi silencing analysis. An additional six events

did not survive after planting in the field. Therefore, a total of 56 transgenic events were studied (Table 3.1).

Real-time PCR primers were designed to amplify only the resident *BAR* gene, not the transcribed dsRNA from the RNAi constructs. A majority of transgenic events (72%) had two individual plants (ramets) assayed for RNAi suppression levels; the remaining events employed two technical replicates (two different runs of RT-PCR). For each of the four parent events, the expression level of the *BAR* gene was determined by pooled RNAs from four different ramets. Due to the different expression levels of the four parent events used for transformation (Appendix F3.2), quantified expression level of the *BAR* gene for

Table 3. 1. Number and RNAi silencing efficiency for each of event and construct combinations and RNAi silencing efficiency. Abbreviations: Pi: IR of promoter sequence; Bi: IR of coding sequence; M-: without flanking MARs; M+: with flanking MARs.

	Construct	No. of transgenic Events	Suppression > 90% (% of Total)[a]	Suppression > 50% (% of Total)	Suppression > 0% (% of Total)
Lines Pooled	PiM-	16	0(0%)	4(25%)	10(63%)
	PiM+	15	2(13%)	3(2%)	8 (53%)
	BiM-	16	13 (76%)	14(88%)	14(88%)
	BiM+	9	7 (78%)	8 ((89%)	8 (89%)
Lines & MAR pooled	Pi	31	2 (6%)	7 (23%)	18 (58%)
	Bi	25	20(80%)	22(88%)	22(88%)
Lines and IRs pooled	M-	32	13 (41%)	18(56%)	24(75%)
	M+	24	9(38%)	11(46%)	16(67%)

[a] Number of transgenic events with more than 90% expression suppression of the *BAR* gene in each parent event and construct combination.

each individual RNAi transgenic event was normalized as a fraction of the expression of their corresponding parent event (Fig. 3.3; Appendix F3.3). Therefore, highly suppressed or silenced events had a value close to 0, while those not suppressed had a value near to 1. Except when considering year-to-year stability, mean expression over two years for each event was used for RNAi suppression analyses.

A large difference in RNAi suppression efficiency was observed between promoter (Pi) and coding sequence (Bi) RNAi types (Table 3.1; Fig. 3.3). For transgenic events containing Bi constructs, expression of the *BAR* gene was either highly suppressed or close to that of the parental event (Appendix F3. 5A), with 80% of the transformed events showing strong suppression. However, transgenic events containing Pi had more variable suppression levels (Appendix F3.4B), with a small portion (6%) of transformed events being highly suppressed. RNAi constructs containing IR directed at the coding sequence had a much higher suppression efficiency than IR directed at promoter sequence (Table 3.1).

None of eight transgenic events transformed with PiM- in the 717-A background showed more than 90% suppression of the *BAR* gene, while six of seven (86%) 717-A transgenic events transformed with BiM- showed more than 90% suppression (Appendix T3.2). A similar difference in suppression efficiency was observed between seven 353-29 transgenic events containing PiM- and BiM-, respectively. None of the PiM group showed more than 90% suppression, while 71% of BiM- transformed events showed high degrees of suppression. When RNAi transgenics were grouped according to the RNAi constructs, the frequencies with more than 90% suppression of the *BAR* gene were: 0% for 16 PiM- transgenics (Fig. 3.3A), 13% for 15 PiM+ events (Fig. 3.3B), 76% for 16 BiM- events (Fig. 3.3C), and 78% for 9 BiM+ events (Fig. 3.3D). When transgenics were further pooled according to the types of dsRNAs,

i.e., Pi or Bi, 2 of 31 (6%) events containing Pi had more than 90% suppression of the *BAR* gene, whereas 20 of 25 (80%) Bi-containing events showed more than 90% suppression (Appendix F3.5). This difference is highly statistically significant based on Chi-square test ($p < 0.0001$).

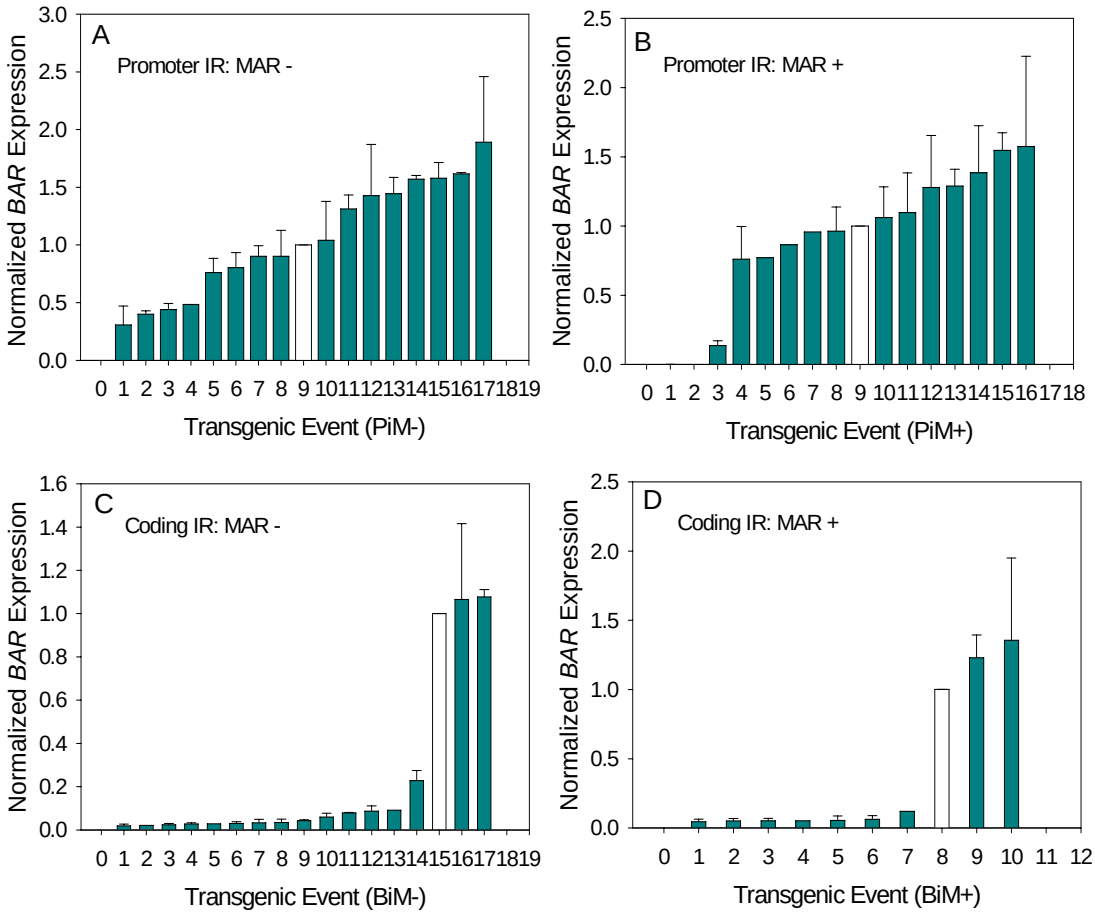


Fig. 3. 3. Expression of the resident *BAR* transgene in transgenic events containing four different RNAi constructs. Expression of the *BAR* gene was determined by real-time RT-PCR using the *UBQ* gene as a reference gene. *BAR* expression normalized as a fraction of parent event, and averaged over two years. White bars represent the parent events which have a normalized value of 1. Error bars represent one standard error of the mean over two years.

Effect of flanking MARs on suppression efficiency

There was no obvious effect of MARs on the frequency of gene suppression. In the 717-A transgenic population, BiM- and BiM+ transformed events showed similar percentages of events with more than 90% suppression (Table 3.1). All BiM- transformed events showed more than 50% suppression, compared to 80% of BiM+ events. No difference was observed in frequency of highly suppressed events in the 717-A population transformed with PiM- and PiM+.

No obvious association was observed between RNAi suppression and the presence of flanking MARs when transgenic populations from different parental events were pooled. There was not a statistically significant effect on suppression frequency (Chi-square test, $p > 0.05$).

Stability of RNAi suppression over time

Based on quantified expression of the resident *BAR* transgene in two different years in the field, the degrees of RNAi suppression was highly stable over time in all constructs ($r = 0.69$; Fig. 3.4). Events with a high level of suppression in the first year stayed highly suppressed in the subsequent year. Among 17 events with more than 90% suppression in the first year, 16 events (94%) retained the same strong degree of suppression, and the remaining one event decreased its suppression from 92% to 46%.

RNAi suppression was also studied at different seasonal points for two highly suppressed events and one non-transformed parent event. Samples were taken since early March (before buds were flushed) through mid-summer (Fig. 3.5). Temperature ranged from 14 to 27 ^{0}C at the time when the samples were taken for RNAi suppression analyses, and mean monthly temperature ranged from 9.1 to 20.6 ^{0}C in months of sampling. Plants were dormant at the first sample time, and actively growing in the last four sample times. There was

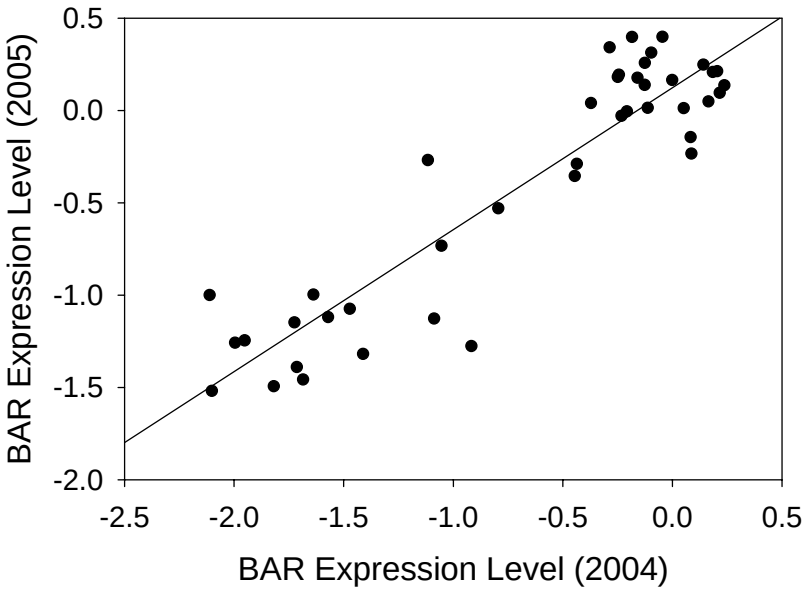


Fig. 3. 4. Stability of RNAi suppression over two years in the field. Expression levels were determined with real-time RT-PCR using *UBQ* as an internal reference gene, normalized to expression levels of corresponding parent events used for transformation, and then log transformed. $R^2 = 0.47$ ($p < 0.0001$, $n = 42$).

substantial seasonal variation in expression of the *BAR* gene in the parental event, a likely consequence of its *rbc*S promoter, which is most active in mature leaves (Sugita and Gruissem, 1987). The two highly suppressed events (>97%), which included one 353-29 event (BiM-), and one 353-38 event (PiM+), showed stable suppression over time.

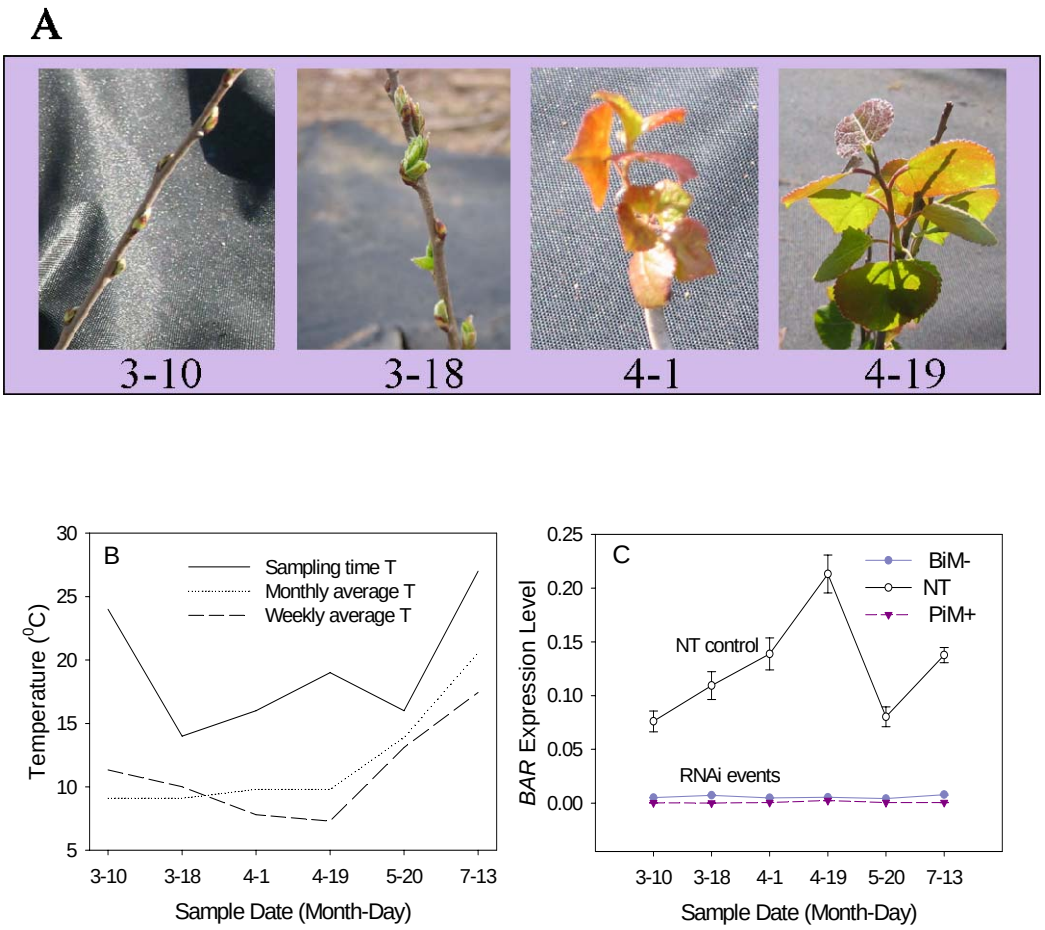


Fig. 3. 5. Seasonal changes inRNAi suppression. A: Plant status when bud/leaf samples were taken for RNAi suppression analyses. B: Temperature (T) at the time when samples were taken, mean weekly T (previous six days plus the sample day), and mean monthly T. C: Expression of the target *BAR* gene for two strongly suppressed events: one 353-29 transgenic event transformed with BiM-, and one 353-38 event transformed with PiM+, and parent event 353-29 (NT). Error bars represent one standard error of mean over technical replicates.

Molecular characterization of RNAi transgenes

Transgene copy number

RNAi transgene copy number was estimated for 53 transgenic events with comparative real-time PCR (described in the Chapter II). Due to the presence of the *BAR* gene in the parental event used for transformation, sequence within the 35S promoter that drove the sense and antisense fragments was used to estimate copy number. Among the 53 transgenic events studied, 45 (85%) contained a single copy, five (9%) had two copies, and three (6%) had three copies (Fig. 3.6). The frequency of single copy transformants was about the same as that we observed in our reporter gene populations (Chapter II, Table 2.5).

RNAi transgene copy number was not associated with the degree of RNAi suppression of the *BAR* gene ($r = 0.015$). The degrees of suppression varied widely among single-copy events. The mean degree of *BAR* suppression was 37% for single-copy integration, 14% for two copy events, and 58% for three copy events. The difference in mean suppression level between single and multiple copy events was not statistically significant ($p < 0.05$). The presence of the multiple copies of dsRNA also did not increase variability of suppression levels among transformants (Levene's test, $p > 0.05$).

DNA methylation analysis

DNA methylation is often associated with TGS or PTGS (Wang and Waterhouse, 2000; Mette et al., 2000). We therefore analyzed DNA methylation status of both promoter and coding sequence of the target *BAR* gene in 14 RNAi transgenic events and two parent events, 353-38 and 717-A (Table 3.2). Among 14 studied RNAi transgenic events, four were transformed with the

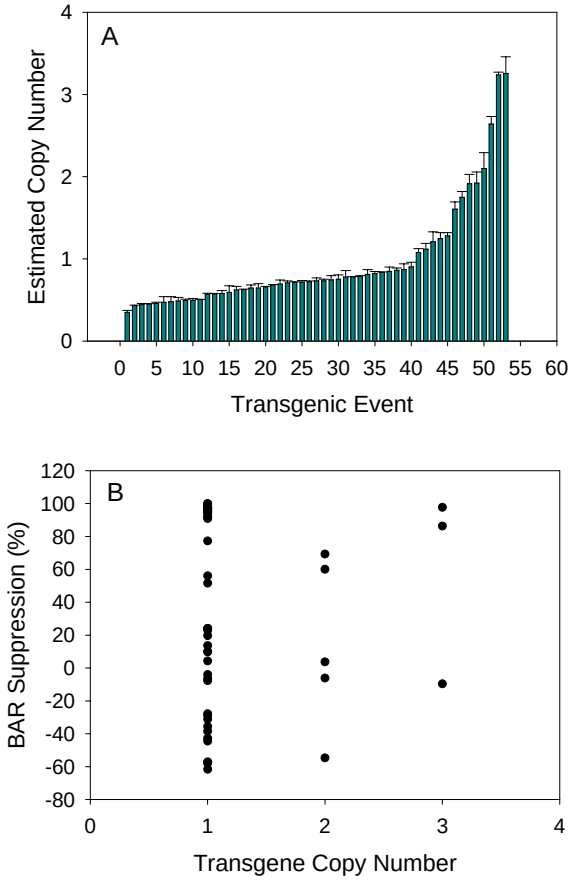


Fig. 3. 6. Estimated transgene copy number and correlation with RNAi suppression levels. A: Quantified RNAi transgene copy number of 53 transgenic events with comparative real-time PCR. Error *BAR* represents one standard error of mean of duplicated wells. B: Correlation of transgene copy number with RNAi suppression levels (%) of the targeted *BAR* gene ($r = 0.015$, $p > 0.05$).

MAR-containing promoter directed construct PiM+ (# 1~4), five with the MAR-less coding region directed construct BiM- (# 7~10, 16), and five with MAR-containing coding region directed construct BiM+ (# 11~15). Except for event #4, 15, and 16 (Table 3.2), all the studied events showed high *BAR* suppression.

Table 3. 2. Methylation status of the target *BAR* gene in RNAi transgenics with different degrees of suppression. Sign "-"indicates unmethylated, and "+" methylated. NT: non-RNAi parent transgenics.

Event No.	Event Name	Degree of RNAi Suppression	Methylation Status	
			***rbc*S Promoter**	***BAR* Coding Sequence**
1	**38-MPi-1**	**99.9**	**+**	**-**
2	**38-MPi-3**	**99.8**	**+**	**-**
3	**7C-MPi-30**	**86**	**+**	**-**
4	7C-MPi-16	0	-	-
5	353-38 (NT)	control	-	-
7	**7A-Bi-150**	**97**	**-**	**-**
8	**7A-Bi-151**	**97**	**-**	**-**
9	**7A-Bi-166**	**97**	**-**	**+**
10	**7A-Bi-218**	**96**	**-**	**-**
11	**7A-MBi-95**	**95**	**-**	**+**
12	**7A-MBi-190**	**95**	**-**	**-**
13	**7A-MBi-191**	**95**	**-**	**-**
14	**7A-MBi-192**	**94**	**-**	**-**
15	7A-MBi-15	0	-	-
16	29-Bi-225	0	-	-
17	717-A (NT)	control	-	-

Therefore, those three unsilenced events along with the two parental events served as putative negative controls for M primers, and positive controls for U primers.

Methylation status was studied with methylation-specific PCR (MSP) using M primers that amplify methylated sequence, and U primers that amplify unmethylated sequence (Fig. 3.7).

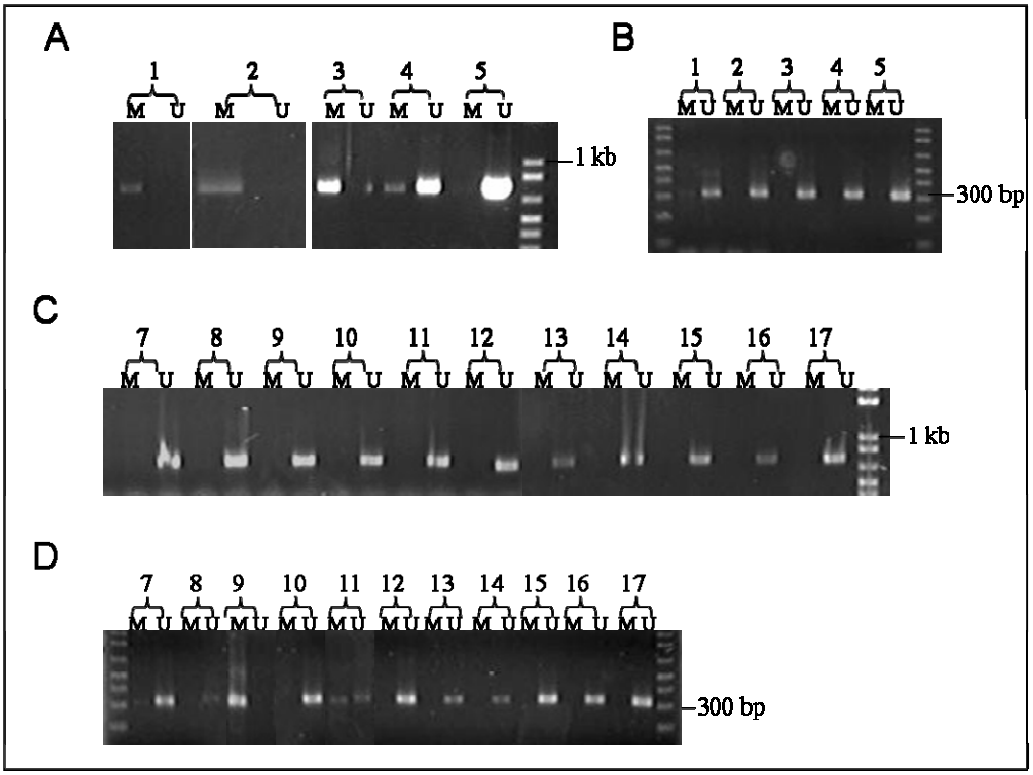


Fig. 3. 7. Methylation-specific PCR (MSP) of RNAi transgenic events and parent events. A: events containing Pi amplified with promoter-specific MSP primers. B: events containing Pi amplified with coding-specific MSP primers. C: events containing Bi amplified with promoter-specific MSP primers. D: events containing Bi amplified with coding-specific MSP primers.

Amplified PCR products of events1, 2, 3, 4, 5, 9, 11, and 15 were directly sequenced with corresponding MSP primers from both ends to confirm PCR results. For all BiM- and BiM+ transgenic events, as well as two parent events, only U primers of *rbc*S promoter were able to amplify expected bands from bisulfite treated genomic DNAs, indicating that the *rbc*S promoter region was never methylated. PCR amplified bands with M primers of the *rbc*S promoter were observed in three highly suppressed PiM- containing transgenic events, and a faint band was observed in a non-suppressed event. Subsequent

sequencing of the amplified products showed that DNA methylation occurred in the region homologous to the IR used in the promoter-directed transgenic constructs (Fig. 3.8). Cytosines in all studied CpG (5) and CAG (3) sequences, and one CCA sequence were methylated (shaded in Fig. 3.8). However, the DNA methylation was not observed outside the homologous region (indicated by * in Fig. 3.8). Sequencing of amplified bands with U primers of *rbc*S promoter from events # 4 and 5 showed an absence of methylation in all sequences studied (Appendix F3.6).

M primers of *BAR* coding sequence amplified positive bands in two of the 16 studied samples 7A-BiM-166 (# 9) and 7A-MBi-95 (#11), and amplified bands with U primers were observed in the rest of samples (Fig. 3.7D, Table 3.2). M amplified bands of # 9 and 11 as well as U amplified bands of # 11 and # 15 were further sequenced to confirm PCR amplification. Sequencing results of M amplified band of event # 9 showed that *BAR* coding sequence was heavily methylated (Fig. 3.9A). The cytosines in almost all CpG (95%) sequences were methylated and not converted to uraciles during bisulfite treatment (shaded). In addition, cytosine methylation also occurred in the symmetric sequence context CNG (N can be A, T, or C) at a frequency of 52% (underlined), and the nonsymmetric sequence CNN occurred at a frequency of 8% (double-underlined).

Similarly, sequencing of M primer-amplified products of event # 11 showed highly methylated *BAR* coding sequence (Fig. 3.9B). Most Methylated cytosines in *BAR* coding sequences were common between event # 9 and 11, but slight differences were observed in several places. Sequencing of the U primer-amplified band of the same event indicated that the amplification was false negative and showed the same methylation of the coding region (data not

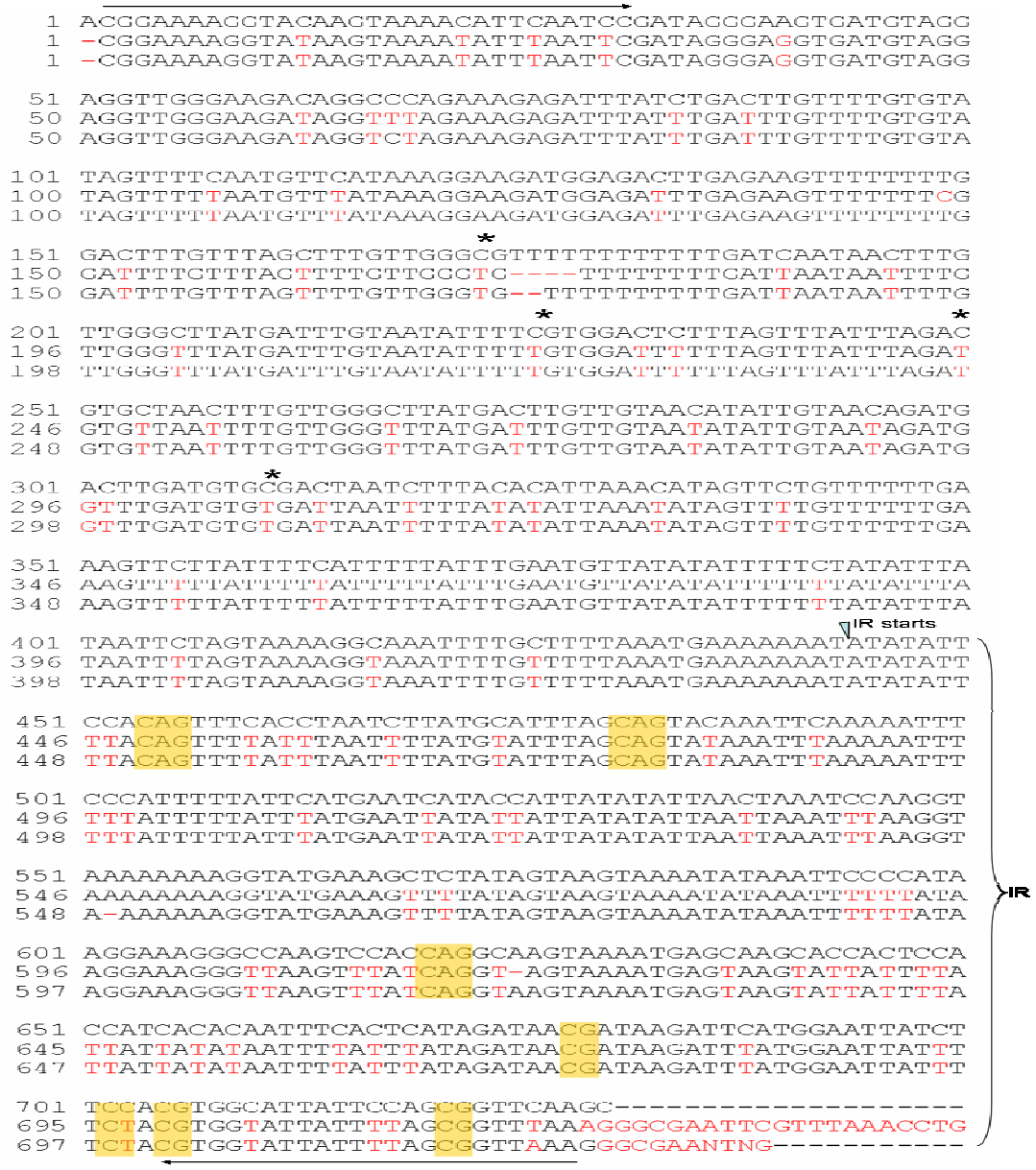


Fig. 3. 8. Methylation analysis of *rbc*S promoter sequence of a PiM+ RNAi transgenic event (38-MPi-3). The top sequence represents untreated *rbc*S promoter genomic sequence. The bottom two sequences represent PCR amplified sequence from bisulfite-treated genomic DNA (two different PCR clones sequenced). Methylated CpG and CAG sequences are shaded. Unmethylated CpG is indicated by "*". The sequence included in IR is also shown on right. Arrows indicate MSP primer sites.

shown). Sequencing of U amplified band from the event #15 clearly showed that *BAR* coding sequence was not methylated for this event, as all cytosines were converted to uraciles after bisulfite treatment (Appendix F3.7).

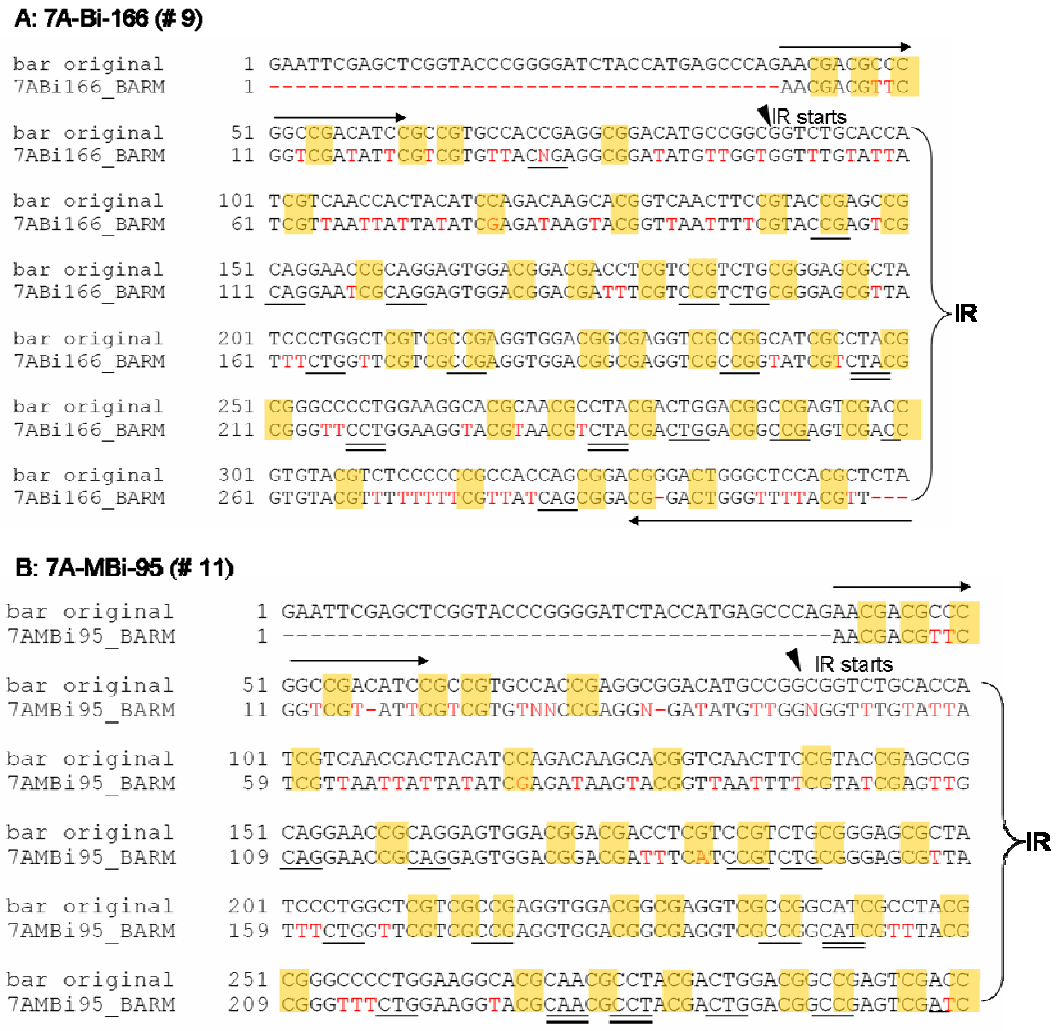


Fig. 3. 9. Methylation analysis of *BAR* coding sequence of two highly suppressed Bi RNAi transgenic events. The top sequence represents untreated *BAR* genomic sequence. The bottom sequence represents PCR amplified sequence from bisulfite-treated genomic DNAs. Methylated CpG sequences were shaded, CNG sequences underlined, and CNN sequences double-underlined. Sequence included in IR is indicated on right. MSP primers are indicated by arrows.

DISCUSSION

RNAi efficiency and associated factors

Effect of different types of dsRNAs

We have demonstrated that our ihpRNA constructs gave high and stable RNAi suppression. The molecular phenotypes of the Bi transformants generally fell into two categories: highly silenced (more than 90% suppression) and non-silenced, with the former group accounting for 80% of the Bi transgenic populations. Other studies of RNAi directed at poplar endogenes support the value of RNAi, but did not report as high a level of efficiency (Mohamed, 2006). Out of 15 transgenic events containing an IR of *PtCENL-1*, the poplar homolog of *Arabidopsis TERMINAL FLOWER 1* gene, only four events (27%) showed more than 50% suppression of the endogenous *PtCENL-1* gene. The other RNAi constructs directed at different poplar flowering genes in our laboratory were also less efficient (unpubl. data). As reviewed in Kurreck (2006), a number of factors could contribute to the efficacy of siRNA, including the thermodynamic properties of the siRNA and the target RNA structure. Therefore, RNAi efficiency might largely depend on the sequence of dsRNA and the properties and structures of target genes. In *Arabidopsis*, degree of RNAi suppression varies dramatically among 25 targeted endogenous genes, with some ranging from little or no residual transcript RNA, some showing little or no reduction, and others showing intermediate reduction (Kerschen et al., 2004).

Transcribed dsRNAs derived from promoter inverted repeats (IR) have been demonstrated to induce TGS of reporter transgenes and endogenous genes in plants (Mette et al., 2000; Cigan et al., 2005). Synthetic siRNAs targeted to the promoter of a specific gene can also induce TGS in human cells (Kawasaki

and Taira, 2004; Morris et al., 2004). Our ihpRNA constructs containing IR directed at the promoter sequence were able to induce suppression of the *BAR* gene, but with a much lower frequency than IR of the coding sequence (Bi). Although 58% of produced Pi transgenic events showed some levels of suppression, only 6% of them showed a level higher than 90%. The frequency of male-sterile phenotypes appeared to vary among different promoter IRs targeted to genes related to pollen development in maize (Cigan et al., 2005). One IR promoter construct was much less efficient in inducing mutant phenotypes than were promoters of other maize genes. The size of IR, properties of promoter sequences and regions used to make the IR including number of cytosine sites that can be methylated, might affect RNAi efficiency. The *Arabidopsis rbc*S promoter is relatively CpG poor, and only contains 22 CpG sites along 1,731 bp of its sequence. The sequence from -31 to -506 relative to translation start site was used for IR in our transgenic constructs, and includes only 7 CpG sites, a plant TATA box, and the other motifs for binding transcription factors. Further studies using different sizes of IRs, and IRs aimed at different regions of this promoter, as well as IRs of other promoters, might provide a very different picture of promoter-directed RNAi efficiency in poplar. In contrast, our coding region directed IR gave very high levels of suppression and transformants were either highly suppressed and had BAR expression levels similar to that of the parent events.

Effect of MARs

MARs could elevate RNAi efficiency if it increases transgene expression (reviewed in Allen et al., 2000) and prevents TGS of transgenes (Levin et al., 2005) or could reduce RNAi efficiency if it impedes homology-dependent gene silencing (Allen et al., 1996; Mlynárová et al., 2003). Chuang and Meyerowitz

(2000) found that expression levels of both strands of AG RNA from an AG inverted repeat transgene increased in proportion to the severity of floral phenotype. In addition, IR driven by 35S promoter achieved high suppression efficiency ranging from 87% to 99% of produced transformants, while weaker NOS promoter gave a lower efficiency (6%). Cigan et al. (2005) also reported that only constitutive expression of MS45 promoter IR under the ubiquitin promoter resulted in male-sterile phenotypes, while no mutant phenotype was observed in transformants containing the same IR but driven by a maize anther-specific promoter. We found that MARs had little effect on transgene expression (Chapter II), perhaps explaining their lack of effect on RNAi efficiency. In *Arabidopsis*, MARs from chicken lysozyme gene boosted the average transgene expression levels five- to 12-fold in a transformed PTGS-impaired background, whereas no such boost is observed in the wild type background, suggesting that MARs don't suppress PTGS but act as enhancers of expression, yet the enhancer effect is only observed when the PTGS mechanism is suppressed (Butaye et al., 2004).

Effect of transgene copy number

There is very little information on the effect of copy number of integrated IR on RNAi suppression efficiency. Transgene copy number was not related to severity of phenotypes in RNAi T1 plants in *Arabidopsis* (Chunag and Meyerowitz, 2000). Also in *Arabidopsis* Kerschen et al. (2004) reported that single copy RNAi T4 lines targeting the same endogenous gene generally reduced transcript levels to the same extent, whereas multi-copy RNAi lines differed in the degree of target reduction and never exceeded the effect of single copy transgenes. We found no obvious association between RNAi transgene copy number and suppression levels of the resident *BAR* gene. Single-copy and

multiple-copy transgenics both showed highly variable degrees of gene suppression. Two of three events that contained three copies showed a high degree of gene suppression, showing that it does not preclude high levels of suppression in poplar.

Stability of RNAi over time

There have been numerous studies on stability of transgene expression (non-dsRNA) in crop plants (reviewed in Stam et al., 1997; Kooter et al., 1999) and in trees (reviewed in Hoenicka and Fladung, 2006). However, little information is available on stability of RNAi suppression. Stable suppression of target genes over the long-term would be particularly important if RNAi technology is used to produce a biosafety trait such as genetically engineered floral sterility for the prevention of transgene dispersal (Strauss et al., 1995). Phenotypic changes of a single highly silenced RNAi transgenic line appeared to be stably inherited over five generations in *Arabidopsis* (Stoutjesdijk et al., 2002). We observed high stability of RNAi suppression with different constructs over two years in the field.

RNAi efficiency did not appear to be affected by the changing level of expression of the rbcS driven *BAR* gene over a season nor was it reduced during dormancy or cool spring temperature. Temperature-dependent RNAi silencing has been reported in plants (Szittya et al. 2003; Sós-Hegedűs et al., 2005), *Drosophila* (Fortier and Belote, 2000), and mammalian cells (Kameda et al., 2004). Different limiting temperatures were reported in those studies. In *Nicotiana benthamiana*, levels of dsRNA-associated siRNA decreased at low temperature (15 ^{0}C). In *Drosophila* the RNAi effect on sex differentiation observed at 29 ^{0}C was strongly inhibited at 22 ^{0}C. A higher limiting temperature was observed in mammalian cells, where RNAi effect was observed at 28 ^{0}C or

below. We found that two highly silenced events, one transformed with Bi, and the other with Pi, remained highly suppressed during the period from early March to July, in spite of large temperature difference on the sampling dates. Therefore, it appears that dormancy cycles and environmental variation might not have a significant impact on RNAi gene suppression in poplar. However, further study of RNAi stability with age and environment, as well as with the specific endogenous gene targets of commercial value, are needed. Nonetheless, our results suggest that the RNAi machinery in trees have evolved to be robust to the extreme environmental variation they are adapted to.

.

Role of DNA methylation

DNA methylation has been implicated in both gene regulation and transgene silencing in plants (reviewed in Wassenegger, 2000). In mammalian genomes, methylation occurs almost exclusively at cytosines in the symmetric dinucleotide context CpG. In plant genomes, the sequence CpG is also the predominant methylation context, but the symmetric context CNG and asymmetric context CNN can also be methylated (Mathieu and Bender, 2004). RNA-directed DNA methylation (RdDM) has been implicated in TGS that is initiated by dsRNAs containing promoter sequences (Aufsatz et al., 2002; Mette et al., 2000; Kawasaki and Taira, 2004; Kawasaki and Taira, 2004; Cigan et al., 2005). We also found that three highly suppressed events containing IR directed at promoter sequence showed DNA methylation in the promoter region. DNA methylation only occurred in the sequence homologous to the IR used in the transgenic constructs, and did not spread to the other regions of the same promoter. Aufsatz et al. (2002) also reported that in *Arabidopsis* NOS promoter dsRNA caused *do novo* methylation within a region of RNA-DNA sequence identity, but observed a more frequent DNA methylation, which affected

cytosines in any sequence context . In the present study, DNA methylation affected all symmetric CpG and CAG sequences, whereas cytosines in nonsymmetric contexts were rarely affected. If RdDM was involved in the present study, then it contrasts with the previous findings that RdDM affects cytosines in all possible sequence contexts (reviewed in Mathieu and Bender, 2004).

RdDM was observed in protein regions in many cases of PTGS (Wang and Waterhouse, 2000; Ebbs et al., 2005). However, the degree to which DNA methylation is relevant to PTGS remains uncertain. Out of eight highly silenced events containing IR directed at coding sequence, two events showed heavy DNA methylation in the coding sequence. DNA methylation occurred at 95% of the CpG sequence contents, 52% of the symmetric CNG sequences, and 8% of the nonsymmetric CNN sequences. This is consistent with previous finding that cytosines at both symmetric and non-symmetric sites can be methylated by RdDM (Pélissier et al., 1999). DNA methylation was not restricted to the sequence used for IR, and spread to the other regions of coding sequence including the transit peptide. But DNA methylation did not spread to the untranscribed promoter region. DNA methylation in the coding region was not observed in the other six highly silenced events, two unsilenced events, and the non-transformed parent event. Although DNA methylation of *GUS* sequence was observed in most of the silenced lines, demethylation treatments of those silenced lines did not release PTGS of *GUS*, suggesting that DNA methylation is not essential for PTGS (Wang and Waterhouse, 2000). Although previous findings and current observation do not preclude a contribution of methylation in silencing, it is clearly not required for strong RNAi-mediated PTGS.

ACKNOWLEDGEMENTS

We would like to thank Plant Genetic System for kindly providing the vector Pttm8, Monsanto for the use of CP4 gene, Dr. Steven Spiker for providing MAR fragment, and Dr. Peter Waterhouse for the use of pHANNIBAL. The research was supported by Initiative for Future Agriculture and Food Systems Grant no. 00-52100-9623 from the USDA Cooperative State Research, Education, and Extension Service, TBGRC/TGERC Research Cooperative, and OSU Foundation Forest Biotechnology Fund.

REFERENCES

Allen, G.C., G.E. Hall, Jr., S. Michalowski, W. Newman, S. Spiker, A.K. Weissinger and W.F.Thompson. (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco. Plant Cell 8: 899-913.

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

Aufsatz, W., M. F. Mette, J. van der Winden, A.J. Matzke, and M. Matzke. (2002) RNA-directed DNA methylation in *Arabidopsis*. Proc. Natl. Acad. Sci. USA. 99**:** 16499-16506.

Baulcombe, D. (2004) RNA silencing in plants. Nature. 431:356-363.

Brummell, D.A., P.J. Balint-Kurti, M.H. Harpster, J.M. Palys, P.W. Oeller, and N. Gutterson. (2003) Inverted repeat of a heterologous 3'-untranslated region for high-efficiency, high-throughput gene silencing. Plant J. 33: 793-800.

Chuang, C.H. and E.M. Meyerowitz. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA. 97: 4985-4990.

Cigan, A.M., E. Unger-Wallace, and K. Haug-Collet. (2005) Transcriptional gene silencing as a tool for uncovering gene function in maize. Plant J. 43: 929-940.

Coruzzi, G., R. Broglie, C. Edwards, and N.H. Chua. (1984).Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase. EMBO J. 3: 1671-1679.

DiFazio, S.P., G.T. Slavov, J. Burczyk, S. Leonardi, and S.H. Strauss. (2004) Gene flow from tree plantations and implications for transgenic risk assessment. In C. Walter and M. Carson (eds.) Plantation Forest Biotechnology for the 21st Century. Research Signpost. Kerala, India. 405-422.

Ebbs, M.L., L. *BAR*tee, and J. Bender. (2005) H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. Mol. and Cell Biol. 25(23): 10507-10515.

Filichkin, S.A., R. Meilan, V.B. Busov, C. Ma, A.M. Brunner, and S.H. Strauss. (2006) Alcohol-inducible gene expression in transgenic *Populus*. Plant Cell Rep. (online version: DOI: 10.1007/s00299-005-0112-3).

Fire, A., S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, and C.C. Mello. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans.* Nature 391**:** 806-811.

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

Fort, E. and J.M. Belote. (2000) Temperature-dependent gene silencing by an expressed inverted repeat in Drosophila. Genesis 26: 240-244.

Hansen, G. and M.D. Chilton. (1999) Lessons in gene transfer to plants by a gifted microbe. Curr. Top. Microbiol. Immunol. 240: 21-57.

Hawkins, S., J. 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.

Hellens, R.P., E.A. Edwards, N.R. Leyland, S. Bean, and P.M. Mullineaux. (2000) pGreen: a versatile and flexible binary Ti Vector for *Agrobacteruim*-mediated plant transformation. Plant Mol. Biol. 42: 819-832.

Iglesias, V.A., E.A. Moscone, I. Papp, F. Neuhuber, S. Michalowshi, 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.

Kameda, T., K. Ikegami, Y. Liu, K. Terada, and T. Sugiyama. (2004) A hypothermic-temperature-sensitive gene silencing by the mammalian RNAi. Biochem. Biophys. Res. Commun. 315(3): 599-602.

Kass, S.U., N. Landsberger, and A.P. Wolffe. (1997) DNA methylation directs a time-dependent repression of transcription initiation. Curr. Biol. 7: 157-165.

Kawasaki H. and K. Taira. (2004) Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature 431: 211-216.

Kerschen, A., C.A.. Napoli, R.A. Jorgensen, and A.E. MÜller. (2004) Effectiveness of RNA interference in transgenic plants. FEBS Letters 566: 223-228.

Kilby, N.J., H.M.O. Leyser, and I.J. Furner. (1992) Promoter methylation and progressive transgene inactivation in *Arabidopsis*. Plant Mol. Biol. 20: 103-112.

Kooter, J.M., M.A. Matzke, and P. Meyer. (1999) Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends in Plant Sci. 4(9): 340-347.

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

Kurreck, K. (2006) siRNA Efficiency: Structure or Sequence—That Is the Question. J. Biomedicine Biotechnol. 2006: 1-7.

Kusaba, M. (2004) RNA interference in crop plants. Curr. Opin. Biotechnol. 15: 139-143.

Levin, J.S., W.F. Thompson, A.S. Csinos, M.G. Stephenson, and A.K. Weissinger. (2005) Matrix attachment regions increase the efficiency and stability of RNA-mediated resistance to Tomato Spotted Wilt Virus in transgenic tobacco. Transgen. Res. 14: 193-206.

Lippman, Z. and R. Martienssen. (2004) The role of RNA interference in heterochromatic silencing. Nature 431: 364-370.

Mathieu, O. and J. Bender. (2004) RNA-directed DNA methylation. J. Cell Sci. 117(21): 4881 - 4888.

Meilan, R., D.J. Auerbach, C. Ma, S.P. DiFazio, and S.H. Strauss. (2002a) Stability of herbicide resistance and GUS expression in transgenic hybrid poplars (*Populus* sp.) during several years of field trials and vegetative propagation. HortScience 37(2): 277-280.

Meilan, R., K. Han, C. Ma, S.P. DiFazio, J.A. Eaton, E.A. Hoien, B.J. Stanton, R.P. Crockett, M.L. Taylor, R.R. James, J.S. Skinner, L. Jouanin, G. Pilate, and S.H. Strauss. (2002b) The *CP4* transgene provides high levels of tolerance to Roundup® herbicide in field-grown hybrid poplars. Can. J. For. Res. 32(6): 967-976.

Mette, M.F., W. Aufsatz, J. van der Winden, M.A. Matzke, A.J.M. Matzke. (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J. 19: 5194-5201.

Meyer P, F. Linn, I. Heidmann, H. Meyer, I. Niedenhof, H. Saedler. (1992) Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. Mol. Gen. Genet. 231: 345-352.

Meza, T.J., D. Kamfjord, A. Hakelien, 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. Transgen. Res. 10: 53-67.

Mlynárová, L., A. Hricová, A. Loonen, and J. Nap. (2003) The Presence of a chromatin boundary appears to shielda transgene in tobacco from RNA silencing. Plant Cell 15: 2203–2217.

Morina, K., O. 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.

Morris K.V., S.W. Chan, S.E. Jacobsen, and D.J. Looney. (2004) Small interfering RNA-induced transcriptional gene silencing in human cells. Science 305: 1289–1292.

Müller A.J., R.R. Mendel, J. Schiemann, C. Simoens, and D. Inzé. (1987) High meiotic stability of a foreign gene introduced into tobacco by *Agrobacterium*-mediated transformation. Mol. Gen. Genet. 207: 171-175.

Pélissier, T., S. Thalmeir, D. Kempe, H.L. Sänger, and M. Wassenegger. (1999) Heavy *de novo* methylation at symmetrical and non-symmetrical sites is a

hallmark of RNA-directed DNA methylation. Nucleic Acids. Res**.** 27**:** 1625–1634.

Plasterk, R.H.A. and R.F. Ketting. (2000) The silencing of the genes. Curr. Opin. in Genet. and Dev. 10: 562-567.

Preuss, D. (1999) Chromatin silencing and Arabidopsis development: a role for polycomb proteins. Plant J. 11: 765-767.

Rozen, S. and H.J. Skaletsky. (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology**.** Humana Press, Totowa, NJ, pp 365-386.

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

Smith, N.A., S.P. Singh, M.B. Wang, P.A. Stoutjesdijk, A.G. Green, and P.M. Waterhouse. (2000) Total silencing by intron-spliced hairpin RNAs. Nature 407: 319-320.

Sós-Hegedűs, A., A. Lovas, M. Kondrák, G. Kovács, and Z. Bánfalvi. (2005) Active RNA silencing at low temperature indicates distinct pathways for antisense-mediated gene-silencing in potato. Plant Mol. Biol. 59: 595-602.

Stoutjesdijk, P.A., S.P. Singh, Q. Liu, C.J. Hurlstone, P.A. Waterhouse, and A.G. Green. (2002) hpRNA-mediated targeting of the *Arabidopsis* FAD2 gene gives highly efficient and stable silencing. Plant Physiol. 129: 1723-1731.

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.

Sugita, M. AND W. Gruissem. (1987) Developmental, organ-specific, and light-dependent expression of the tomato ribulose-1,5-bisphosphate carboxylase small subunit gene family. Proc. Natl. Acad. Sci. USA. 84: 7104-7108.

Szittya, G., D. Silhavy, A. Molnár, Z. Havelda, A. Lovas, L. Lakatos, Z. Banfalvi, and J. Burgyán. (2003) Low temperature inhibits RNA silencing-mediated defence by the control of siRNA generation. EMBO J. 22(3): 633-640.

van Blokland, R., M. ten Lohuis, and P. Meyer. (1997) Condensation of chromatin in transcriptional regions of an inactivated plant transgene: evidence

for an active role of transcription in gene silencing. Mol. Gen. Genet. 257: 1-13.

Vance, V. and H. Vaucheret. (2001) RNA silencing in plants - defense and counterdefense. Science 292: 2277-2280.

Wagner, A, L. Phillips, R.D. Narayan, J.D. Moody, and B. Geddes. (2005) Gene silencing studies in the gymnosperm species *Pinus radiata*. Plant Cell Rep. 24: 95-102.

Wassenegger, M. (2000) RNA-directed DNA methylation. Plant Mol. Biol. 43: 203-220.

Waterhouse, P.M., H.W. Graham, and M.B. Wang. (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci. USA. 95: 13959-13964.

Wesley, S.V., C.A. Helliwell, N.A Smith., M.B. Wang, D.T. Rouse, Q. Liu, P.S. Goodings, S.P. Singh, D. Abbott, P.A. Stoutjesdijk, S.P. Robinson, A.P. Gleave, A.G. Green, and P.M. Waterhouse. (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J. 6: 581–590.

Yan, L., A. Loukoianov,A. Blechl, G. Tranquilli, W. Ramakrishna, P. SanMiguel, J.L. Bennetzen, V. Echenique, and J. Dubcovsky. (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303: 1640-1644.

Chapter 4: CONCLUSION

MAJOR CONCLUSIONS

Transgene expression, stability, and correlates

1. Transgene expression levels of different transformants can vary considerably. Two reporter genes, *GFP* and *BAR*, displayed a continuous, wide range of expression levels in 404 independent transgenic events transformed in two hybrid poplar clones. The range and coefficient of variation among event means was 6.3 fold and 25% for *GFP*, respectively, and 32.6 fold and 70% for *BAR*, respectively.
2. We did not observe any case of gene silencing (complete breakdown of expression) in the more than two thousand trees studied over multiple years in the greenhouse and field under vegetative propagation. A small amount of variation in *GFP* and *BAR* was observed in about 1 ~ 2% of transgenic events, which might be attributed to physiological and developmental status or experimental error. Our results suggested that multiple-year prediction of transgene expression and associated trait performance can be predicted with high confidence based on their expression in the greenhouse or the early field trial.
3. Flanking the transgene constructs with MARs failed to increase transgene expression, but tended to reduce variability among events, probably by reducing the number of weakly expressing events. The presence of MARs might act to reduce transgene deletions associated with T-DNA integration by providing additional sequences next to the two T-DNA borders. Truncations of flanking MAR sequences were frequently observed in MAR-containing transgenic events, and the

association of the presence of MAR with DNA truncations was strongly significant. This might also explain the reported ability of MARs to enhance plant transformation efficiency in that selectable markers, usually placed close to the left border, will be protected from truncation during integration. MARs also significantly reduced the integration of inverted repeats at a single locus, which further supports the hypothesis that MARs reduce the occurrence of weakly expressing events.

4. The *GFP* and *BAR* transgenes, although assembled in the same T-DNA region, had weakly correlated expression among transgenic events in the absence of MARs ($r = 0.15$, $p = 0.11$). In the most extreme cases, one transgene was completely silenced, while the other was highly expressed. The presence of MARs increased coordinated expression between two transgenes ($r = 0.35$; $p = 0.0005$).
5. Our *Agrobacterium*-mediated plant transformation protocol generated a high percentage of single copy transformants (85%). The rest of transformants contained two to four copies. Transgene copy number was positively correlated with transgene expression levels ($r = 0.35 \sim 0.46$, $p < 0.001$). Thus, conventional wisdom that the presence of multiple copies increases the possibility of transgene silencing was not supported by our study.
6. We observed frequent formation of direct repeats when multiple copies of T-DNA were present (42%). The formation of inverted repeats or both types of repeats was relatively rare (5.5% and 7%). When the number of integrated copies was more than two, tandem repeats of at least two copies occurred in all transgenic events. The formation of direct repeats at a locus did not reduce expression. In contrast, expression levels were significantly elevated in direct vs other repeat types (inverted or a combination of inverted and direct).

7. The level of transgene expression was generally maintained after physiological stress imposed by a further round of organogenesis. Physical elimination of transgenes occurred at a frequency of 2.5% (4 of 160 subevents) during organogenesis. In two subevents regenerated from the same transgenic event, both had lost transgenes suggesting this event was prone to instability. This supports the hypothesis that the stress imposed by tissue culture conditions induces genomic instability at particular loci (Gould 1986; Romano et al., 2005), and may be a useful tool for early screening to remove unstable events.
8. Use of the poplar native *rbc*S promoter in our transgenic constructs did not adversely cause obvious silencing compared to the 35S promoter. Therefore, it appears that native promoters, at least those driving tissue-specific expression may be successfully deployed in transgenic poplars under vegetative propagation.

RNAi efficiency, stability, and correlates

1. dsRNA constructs using an IR directed at coding sequence induced suppression of the target *BAR* transgene at high frequency. The RNAi transgenic events tended to be either highly suppressed or near wild-type; 80% of transgenic events showed more than 90% suppression of *BAR* expression. RNAi using dsRNA directed at coding sequence may therefore be a powerful tool to study gene function as well as genetic engineering of sterility in poplar.
2. dsRNA constructs using IR directed at the *BAR* promoter sequence gave a much lower frequency of gene suppression; only 6% of the events showed more than 90% suppression.

3. A high degree of gene suppression was maintained during seasonal development as well as between two different years in the field. Ambient temperature and bud-break in spring had no effect on estimated RNAi suppression. Stability of RNAi directed at transgenes therefore appears to be very stable in poplar under vegetative propagation.
4. The structure of integrated dsRNA transgenes, such as the presence of flanking MARs and number of copies was unassociated with efficiency of suppression.
5. Promoter methylation was associated with high level suppression induced by IR directed at the promoter sequence. DNA methylation only occurred in the region homologous to the IR and not the other regions of the targeted promoter. DNA methylation occurred in all symmetric CpG and CAG sequence contexts, but not in non-symmetric sequences.
6. DNA methylation was observed in two of eight highly silenced events induced by dsRNA directed at coding sequences. Cytosine methylation occurred in almost all CpG sequences (95%). Symmetric CNG sequence contexts were also frequently methylated (58%), while methylation on non-symmetric CNN sequence contexts was rare (8%).

SUGGESTIONS FOR FUTURE RESEARCH

1. Further studies on the efficiency of RNAi induced by dsRNA directed at promoter sequence is needed to provide further insights into its value for gene function studies and genetic engineering of poplar. This can be done by using different regions of a specific promoter, targeting different genes whose promoters have different sequence properties such as CpG richness, and promoters that vary in strength as well as expression

pattern. These studies should focus on endogenes, and compare endogenes and transgenes as their silencing efficiency may vary widely.

2. Our study has demonstrated that both reporter genes and RNAi transgenes can achieve stable expression/suppression of transgenes over multiple years in the field. However, when RNAi technology is used against endogenes, its efficiency and stability may differ, for example, suppression of floral sterility genes to minimize dispersal of transgenes would need to be specifically demonstrated in juvenile and flowering trees. Due to the prolonged vegetative phase, initial selection of transgenic events with high suppression of floral genes might be based on vegetative tissues instead of floral tissues if the target genes are expressed there. This will require long-term field trial to study and may be of great commercial value if it identifies sterile trees years in advance of actual flowering.
3. Temperature-dependence of RNAi silencing in the field environment was not observed on the six sampling dates with varied temperatures, some of which were lower than reported temperature limit which affected RNAi in *Arabidopsis*, tobacco and potato (Szittya et al., 2003). However, in addition to low temperatures on the sampling days, the length of period of consecutive low temperatures before sampling times might be critical in inducing temperature-dependent RNAi. Therefore, further investigation using more events and more frequent sampling over a full seasonal cycle will provide valuable information about RNAi stability in natural field environments. This can be done using endogenous flowering genes to look at seasonal RNAi suppression changes in different tissues (if a constitutive promoter is used). If there is any temperature- or environment-dependent silencing observed, subsequent alteration, if there will be any, in phenotypes of modified

traits (sterility) can also be evaluated to see whether temporary loss of RNAi silencing is critical in the deployment of RNAi-related traits in the field.

4. Based on the present and previous study (Meilan et al., 2002a; Hawinks et al., 2003) complete gene silencing in vegetatively propagated poplar is very rare. The major issue remaining is what degree/amount of expression change is acceptable in a commercial application program. Based on the correlation of phenotypic values (herbicide damage) and protein expression levels, events with expression values in the top 50% (3 fold range) can confer the same high herbicide resistance (unpubl. data). Therefore, it seems that if expression changes are within a range that gives expected phenotypes, such change might not need to be concerned. Future study on the correlation in changes between observed phenotypes and measured mRNA/protein levels will provide relevant biological data for transgene risk and stability evaluation.
5. Although we did not observe elevated transgene expression with the presence of flanking MARs in the transgenic constructs, several intriguing roles of MARs implied in our results might be worth further exploration. One aspect to investigate is the frequency of T-DNA truncations in single-copy transformants containing MARs. If frequent truncations occur in the MAR sequence, then it will further support our hypothesis that MARs enhance transformation efficiency and reduce gene silencing by providing buffering sequence for DNA truncations during the integration process. In addition, this hypothesis can be further tested by replacing MAR sequences with other DNA sequence, ideally with sequences differing in GC or AT richness, to see if they also prevent T-DNA truncations, resulting in increased transformation efficiency and reduced gene silencing. Increased transformation might be

very critical to species which are recalcitrant to be transformed. In addition, the insignificant effect of MARs on elevating transgene expression observed in the present study might be due to the types of transgenes and promoters used. Strongly and weakly expressing promoters might interact differently with the presence of MARs (Mankin et al., 2003). Both the 35S promoter and poplar *rbc*S promoters used to drive the *GFP* and *BAR* genes can give strong expression. A more significant effect might be observed for the promoters with a weaker strength. Finally, we used the MAR fragment from tobacco root expressed gene RB7 in the present study. Whether other types of MARs from different species and genes act in a similar manner in poplar will need to be investigated by flanking the same transgenes with different MARs.

BIBLIOGRAPHY

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.

Allen, G.C., G.E. Hall, Jr., S. Michalowski, W. Newman, S. Spiker, A.K. Weissinger and W.F.Thompson. (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco. Plant Cell 8: 899-913.

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

Alonso, J.M., A.N. Stepanova, T.J. Leisse, C.J. Kim, H. Chen, P. Shinn, D.K. Stevenson, J. Zimmerman, P. Barajas, and R. Cheuk. (2003) Genome-wide insertional mutagenesis of *Arabidopsis* thaliana. Science 301: 653-657.

Amedeo, P., Y. Habu, K. Afsar, S.O. Mittelsten, and J. Paszkowski. (2000) Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. Nature 405: 203-206.

Araki, T. (2001) Transition from vegetative to reproductive phase. Curr. Opin. Plant Biol. 4: 63-68.

Aufsatz, W., M. F. Mette, J. van der Winden, A.J. Matzke, and M. Matzke. (2002) RNA-directed DNA methylation in *Arabidopsis*. Proc. Natl. Acad. Sci. USA. 99**:** 16499-16506.

Battey, N.H. and F. Tooke F. (2002) Molecular control and variation in the floral transition. Curr. Opin. Plant Biol. 5: 62-68.

Baucher, M., B. Monties, M.V. Montagu, and W. Boerjan. (1998) Biosynthesis and genetic engineering of lignin. Crit. Rev. Plant Sci. 17(2): 125-198.

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

Baulcombe, D. (2004) RNA silencing in plants. Nature 431: 356-363.

Baulcombe, D. (2004) RNA silencing in plants. Nature 431: 356-363.

Bender, J. (2001) A vicious cycle: RNA silencing and DNA methylation in plants. Cell 106: 129-132.

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

Bishop-Hurley, S.L., R.J. Zabkiewicz, L. Grace, R.C. Gardner, R.C. Wagner, and C. Walter. (2001) Conifer genetic engineering: transgenic Pinus radiata (D. Don) and Picea abies (Karst) plants are resistant to the herbicide Buster. Plant Cell Rep. 20: 235–243.

Bode, J., M. Stengert-Iber, T. Schlake, and A. Dietz-Pfeilstetter. (1996) Scaffold/matrix-attached regions: topological switches with multiple regulatory functions. Crit. Rev. Euc. Gene Exp. 6: 115-138.

Boerjan, W. (2005) Biotechnology and the domestication of forest trees. Curr. Opin. Biotechnol. 16: 159-166.

Borkowska, M., K. Kleczkowski, A. Pawelczak, and B. Wielgat. (1995) Transformation of diploid potato with an Agrobacterium tumefaciens binary vector system. II. Stability of transformation in tubers, micropropagated and greenhouse grown plants. Acta Phys. Plant. 17: 275-280.

Bradshaw, H.D., Jr, R. Ceulemans, J. Davis, R.F. Stettler. (2000) Emerging model systems: Poplar (*Populus*) as a model forest tree. J. Plant Growth Regulators 19: 306–313.

Broer, I. (1996) Stress inactivation of foreign genes in transgenic plants. Field Crop Res. 45: 19-25.

Brouwer, C., W. Bruce, S. Maddock, Z. Avramova, and B. Bowen (2002) Suppression of transgene silencing by matrix attachment regions in Maize: a dual role for the Maize 5' *ADH1* matrix attachment region. Plant Cell 14: 251 - 2264.

Brummell, D.A., P.J. Balint-Kurti, M.H. Harpster, J.M. Palys, P.W. Oeller, and N. Gutterson. (2003) Inverted repeat of a heterologous 3'-untranslated region for high-efficiency, high-throughput gene silencing. Plant J. 33: 793-800.

Brunner, A.M., W.H. Rottmann, L.A. Sheppard, K.V. Krutovskii, S.P. DiFazio, S. Leonardi, and S.H. Strauss. (2000) Structure and expression of duplicate AGAMOUS orthologs in poplar. Plant Mol. Biol. 44(5): 619-634.

Brunner, A.M., and O. Nilsson. (2004a) Revisiting tree maturation and floral initiation in the poplar functional genomics era. New Phytologist 164: 43-51.

Brunner, A.M., V.B. Busov, and S.H. Strauss. (2004b) Poplar genome sequence: functional genomics in an ecologically dominant plant species. Trends Plant Sci. 9: 49-56.

Brunner, A.M., I.A. Yakovlev, and S.H. Strauss (2004c) Validating internal controls for quantitative plant gene expression studies. BMC Plant Biol. 4: 14-20.

Campbell, M.M., A.M. Brunner, H.M. Jones, and S.H. Strauss. (2003) Forestry's Fertile Crescent: The application of biotechnology to forest trees. Plant Biotechnol. J. 1: 141-154.

Carlsbecker A, K. Tandre, U. Johanson, M. Englund, and P. Engstrom. (2004) The MADS-box gene DAL1 is a potential mediator of the juvenile-to-adult transition in Norway spruce (Picea abies). Plant J. 40: 546–557.

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, S., W. Jin, M. Wang, F. Zhang, J. Zhou, Q. Jia, Y. Wu, F. Liu, and P. Wu. (2003) Distibution and characterization of over 1000 T-DNA tags in rice genome. Plant J. 36: 105-113.

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.

Chuang, C.F. and E.M. Meyerowitz. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis* thaliana. Proc. Natl. Acid. Sci. USA, 97: 4985-4990.

Cigan, A.M., E. Unger-Wallace, and K. Haug-Collet. (2005) Transcriptional gene silencing as a tool for uncovering gene function in maize. Plant J. 43: 929-940.

Cluster, P.D., M. O'Dell, M. Metzlaff, and R.B. Flavell. (1996) Details of T-DNA structural organization from a transgene Petunia population exhibiting co-suppresssion. Plant Mol. Biol. 32: 1197-1203.

Conner, A.J., L. Mlynárová, W.J. Stiekema, and J. Nap. (1998) Meiotic stability of transgene expression is unaffected by flanking matrix-associated regions. Mol Breed. 4: 47-58.

Coruzzi, G., R. Broglie, C. Edwards, and N.H. Chua. (1984)Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase. EMBO J. 3: 1671-1679.

DiFazio, S.P., G.T. Slavov, J. Burczyk, S. Leonardi, and S.H. Strauss. (2004) Gene flow from tree plantations and implications for transgenic risk assessment. In C. Walter and M. Carson (eds.) Plantation Forest Biotechnology for the 21st Century. Research Signpost. Kerala, India. 405-422.

Dowda, P.F., L.M. Lagriminib, and D.A. Hermsc. (1998) Differential leaf resistance to insects of transgenic sweetgum (Liquidambar styraciflua) expressing tobacco anionic peroxidase. CMLS, Cell. Mol. Life Sci. 54 : 712-720.

Ebbs, M.L., L. *BAR*tee, and J. Bender. (2005) H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. Mol. and Cell Biol. 25(23): 10507-10515.

Elo, A, J. Lemmetyinen, M.L. Turunen, L. Tikka, and T. Sopanen. (2001) Three MADS-box genes similar to APETALA1 and FRUITFUL from silver birch (Betula pendula). Physiol. Plant. 112: 95-103 .

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

FAO. (2004) Preliminary review of biotechnology in forestry, including genetic modification. Forest Genetic Resources Working Paper FGR/59E. Forest Resources Development Service, Forest Resources Division, Rome, Italy.

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.

Fernandez-Cornejo, J. and C. Margriet. (2006) The first decade of genetically engineered crops in the United States U.S. Department of Agriculture, economic Research Service, Technical Bulletin No. 11. April. 2006.

Filichkin, S.A., R. Meilan, V.B. Busov, C. Ma, A.M. Brunner, and S.H. Strauss. (2006) Alcohol-inducible gene expression in transgenic *Populus*. Plant Cell Rep. (online version: DOI: 10.1007/s00299-005-0112-3).

Fire, A., S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, and C.C. Mello. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans.* Nature 391**:** 806-811.

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

Fort, E. and J.M. Belote. (2000) Temperature-dependent gene silencing by an expressed inverted repeat in Drosophila. Genesis 26: 240-244.

Francis, K.E. and S. Spiker. (2005) Identification of *Arabidopsis* thaliana transformants without selection reveals a high occurrence of silenced T-DNA integrations. Plant J. 41: 464-477.

Franke, R., C.M. McMichael, K. Meyer, A.M. Shirley, J.C. Cusumano, and C. Chapple. (2000) Modified lignin in tobacco and poplar plants over-expressing the *Arabidopsis* gene encoding ferulate 5-hydroxylase. Plant J. 22(3): 223-234.

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

Gelvin, S.B. (2003) Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol. Biol. Rev. 67(1): 16-37.

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

Gould, A.R. (1986) Factors controlling generation of variability *in vitro*. In: Cell Culture and Somatic Cell Genetics of Plants (Vasil, I.K., ed.). Vol. 3. Academic Press, Orlando, FL, USA, pp. 549-567.

Halfhill, M.D., R.J. Millwood, and 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.P. Ma, and S.H. Strauss. (1997) Matrix attachment regions (MARs) enhance transformation frequency and transgene expression in poplar. Transgen. Res. 6: 415-420.

Han, K.-H., M.P. Gordon, and S.H. Strauss. (1997) High frequency transformation of cottonwoods (*genus Populus*) by *Agrobacterium* rhizogenes. Can. J. For. Res. 27: 464-470.

Hansen, G. and M.D. Chilton. (1999) Lessons in gene transfer to plants by a gifted microbe. Curr. Top. Microbiol. Immunol. 240: 21-57.

Hansen, G., R.D. Shillito, and M.D. Chilton. (1997) T-strand integration in maize protoplasts after codelivery of an T-DNA substrate and virulence genes. Proc. Natl. Acad. Sci. USA. 94: 11726-11730.

Harcourt, R.L., J. Kyozuka1, R.B. Floyd, K.S. Bateman, H. Tanaka, V. Decroocq, D.J. Llewellyn,_X. Zhu, W.J. Peacock, and E.S. Dennis. (2000) Insect- and herbicide resistant transgenic eucalypts. Mol. Breed. 6: 307–315.

Harper, B.K. and C.N. Stewart Jr. (2000) Patterns of green fluorescent protein expression in transgenic plants. Plant Mol. Biol. Rep. 18: 141a-141i.

Harper, B.K., S.A. Mabon, S.M. Leffel, M.D. Halfhill, H.A. Richards, K.A. Moyer, C. Stewart, and J. Neal. (1999) Green fluorescent protein as a marker for expression of a second gene in transgenic plants. Nature Biotechnol. 17: 1125-1129.

Haseloff, J., K.R. Siemering, D.C. Prasher, and S. Hodge. (1997) Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly. Proc. Natl. Acad. Sci. 94: 2122-2127.

Hawkins, S., J. Leplé, D. Cornu, L. Jouanin, and G. Pilate. (2003) Stability of transgene expression in poplar: A model forest tree species. Ann. For. Sci. 60:

Hellens, R.P., E.A. Edwards, N.R. Leyland, S. Bean, and P.M. Mullineaux. (2000) pGreen: a versatile and flexible binary Ti Vector for *Agrobacteruim*-mediated plant transformation. Plant Mol. Biol. 42: 819-832.

Heuchelin, S.A., L. Jouanin, N.B. Klopenstein, and H.S. McNabb. (1997) Potential of proteinase inhibitors for enhanced resistance to Populus arthropod and pathogen pests. In Klopfenstein, N.B., Chun, Y.W., Kim, M.S. & Ahuja, M.R., eds. Micropropagation, genetic engineering, and molecular biology of Populus. Gen. Tech. Rep. RM-GTR-297, pp. 173–177. Fort Collins, CO, US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

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: Fladung, M., & D. Ewald (eds).Tree transgenesis - recent developments. Springer Berlin, Heidelberg, pp. 301-321.

Hu, W., S.A. Harding, J. Lung, J.L. Popko, J. Ralph, D.D. Stokke, C. Tsai, and V.L. Chiang. (1999) Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nature Biotechonol. 17: 808-812.

Huntley, S.K., D. Ellis, M. Gilbert, C. Chapple, and S.D. Mansfield. (2003) Significant Increases in Pulping Efficiency in C4H-F5H-Transformed Poplars: Improved Chemical Savings and Reduced Environmental Toxins. J. Agric. Food Chem. 51(21): 6178 -6183.

Iglesias, V.A., E.A. Moscone, I. Papp, F. Neuhuber, S. Michalowshi, 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.

James, C. (2005) Global Status of Commercialized Biotech/GM Crops: 2005. ISAAA Briefs No. 34. http://www.isaaa.org.

James, V. A., B. Worland, J. W. Snape, and P. Vain. (2004) Strategies for precise quantification of transgene expression levels over several generations in rice. J. Exp. Bot. 55(401): 1307-1313.

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.

Jouanin, L., T. Goujon, V. Nadaï, M. Martin, I. Mila, C. Vallet, B. Pollet, A. Yoshinaga, B. Chabbert, M. Petit-Conil, and C. Lapierre. (2000) Lignification in Transgenic Poplars with Extremely Reduced Caffeic Acid O-Methyltransferase Activity. Plant Physiol. 123: 1363-1374.

Kameda, T., K. Ikegami, Y. Liu, K. Terada, and T. Sugiyama. (2004) A hypothermic-temperature-sensitive gene silencing by the mammalian RNAi. Biochem. Biophys. Res. Commun. 315(3): 599-602.

Kass, S.U., N. Landsberger, and A.P. Wolffe. (1997) DNA methylation directs a time-dependent repression of transcription initiation. Curr. Biol. 7: 157-165.

Kawasaki H. and K. Taira. (2004) Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature 431: 211-216.

Keinonen, H.M. and T. Sopanen. (2004) Functional characterization of SEPALLATA3 and AGAMOUS orthologues in silver birch. Physiologia Plantarum 121: 149-162.

Kerschen, A., C.A.. Napoli, R.A. Jorgensen, and A.E. MÜller. (2004) Effectiveness of RNA interference in transgenic plants. FEBS Letters 566: 223-228.

Kilby, N.J., H.M.O. Leyser, and I.J. Furner. (1992) Promoter methylation and progressive transgene inactivation in *Arabidopsis*. Plant Mol. Biol. 20: 103-112.

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. Weger, E. Stoger, and P. Christou. (2003) Transgene integration, organization and interaction in plants. Plant Mol. Biol. 52: 247-258.

Komeda, Y. (2004) Genetic regulation of time to flower in Arabidopsis thallana. Annu. Rev. Plant Biol. 55: 521–35.

Kooter, J.M., M.A. Matzke, and P. Meyer. (1999) Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends in Plant Sci. 4(9): 340-347.

Krizkova, L and M. Hrouda. (1998) Direct repeats of T-DNA integrated in tobacco chromosome characterization of junction regions. Plant J. 16(6): 673-680.

Kumar, S. and M. Fladung. (2000a) Transgene repeats in aspen: molecular characterization 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. (2000b) Determination of transgene repeat formation and promoter methylation in transgenic plants. BioTechniques. 28: 1128-1137.

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.

Kurreck, K. (2006) siRNA Efficiency: Structure or Sequence—That Is the Question. J. Biomedicine Biotechnol. 2006: 1-7.

Kusaba, M. (2004) RNA interference in crop plants. Curr. Opin. Biotechnol. 15: 139-143.

Lapierre, C., B. Pollet, and M. Petit-Conil. (1999) Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid O-methyltransferase activity have an opposite impact on the efficiency of industrial kraft pulping. Plant Physiol. 119: 153-164.

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(2): 671- 677.

Leplé1, M.B, S. Augustin, G. Pilate, V.D.L. Tân, A. Delplanque, D. Cornu, and L. Jouanin. Toxicity toChrysomela tremulae (Coleoptera: Chrysomelidae) of transgenic poplars expressing a cysteine proteinase inhibitor. Mol. Breed. 1(4): 319-328.

Levee, V., E. Garin, K. Klimaszewska, and A. Seguin. (1999) Stable genetic transformation of white pine (Pinus strobus L.) after cocultivation of embryogenic tissues with *Agrobacterium* tumefaciens. Mol. Breed. 5: 429–440.

Li, L.C. and R. Dahiya. (2002) MethPrimer: designing primers for methylation PCRs. Bioinformatics. 8(11): 1427-31.

Linn, F., I. Heidmann, H. Saedler, and P. Meyer. (1990) Epigenetic changes in the expression of the maize A1 gene in Petunia hybrida: Role of numbers of integrated gene copies and state of methylation. Mol. Gen. Genet. 222: 329-336.

Lippman, Z. and R. Martienssen. (2004) The role of RNA interference in heterochromatic silencing. Nature 431: 364-370.

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.

Mankin, S.L., G.C. Allen, T. Phelan, S. Spiker, and W.F. Thompson. (2003) Elevation of transgene expression level by flanking matrix attachment regions (MAR) is promoter dependent: a study of the interactions of six promoters with the RB7 3' MAR. Transgenic Res. 12(1): 3-12.

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

Mathieu, O. and J. Bender. (2004) RNA-directed DNA methylation. J. Cell Sci. 117(21): 4881 - 4888.

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.

Maximova, S., C. Miller, G. Antfflnez de Mayolo, S. Pishak, A. Young, and M. J. Guiltinan. (2003) Stable transformation of Theobroma cacao L. and influenceof matrix attachment regions on *GFP* expression. Plant Cell Rep. 21: 872-883.

McCabe, M.S., U.B. Mohapatra, S.C. Debnath, J.B. 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., K. Han, C. Ma, R.R. James, J.A. Eaton, B.J. Stanton, E.A. Hoien, R.P. Crockett, and S.H. Strauss. (2000a) Development of glyphosate-tolerant hybrid cottonwoods. Tappi J. 83(1): 164-166.

Meilan, R., C. Ma, S. Cheng, J.A. Eaton, L.K. Miller, R.P. Crockett, S.P. DiFazio, and S.H. Strauss. (2000b) High levels of Roundup® and leaf-beetle resistance in genetically engineered hybrid cottonwoods. In K.A. Blatner, J.D. Johnson, and D.M. Baumgartner (Eds.), Hybrid Poplars in the Pacific Northwest: Culture, Commerce and Capability Washington State University Cooperative Extension Bulletin MISC0272, Pullman, WA, USA Pp 29-38.

Meilan, R., A.M. Brunner, J.S. Skinner, and S.H. Strauss. (2001) Modification of Flowering in Transgenic Trees. In N. Morohoshi and A. Komamine (Eds.) Elsevier Science B.V, Pp 247-256.

Meilan, R., D.J. Auerbach, C. Ma, S.P. DiFazio, and S.H. Strauss. (2002a) Stability of herbicide resistance and GUS expression in transgenic hybrid poplars (*Populus* sp.) during several years of field trials and vegetative propagation. HortScience 37(2): 277-280.

Meilan, R., K. Han, C. Ma, S.P. DiFazio, J.A. Eaton, E.A. Hoien, B.J. Stanton, R.P. Crockett, M.L. Taylor, R.R. James, J.S. Skinner, L. Jouanin, G. Pilate, and S.H. Strauss. (2002b) The *CP4* transgene provides high levels of tolerance to Roundup® herbicide in field-grown hybrid poplars. Can. J. For. Res. 32(6): 967-976.

Mette, M.F., J. van der Winden, M.A. Matzke, and A.J.M. Matzke. (1999) Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters in trans. EMBO J. 18(1): 241-248.

Mette, M.F., W. Aufsatz, J. van der Winden, M.A. Matzke, A.J.M. Matzke. (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J. 19: 5194-5201.

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

Meyer P, F. Linn, I. Heidmann, H. Meyer, I. Niedenhof, H. Saedler. (1992) Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. Mol. Gen. Genet. 231: 345-352.

Meyer, P. and I. Heidmann. (1994) Epigenetic variants of a transgenic petunia line show hypermethylation in transgene DNA: an indication for specific recognition of foreign DNA in transgenic plants. Mol. Gen. Genet. 243: 390-399.

Meyer, P. and H. Saedler. (1996) Homology-dependent gene silencing in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47: 23-48.

Meza, T.J., B. Stangeland, I.S. Mercy, M. Skårn, D.A. Nymoen, A. Berg, M.A. Butenko, A-M. Håkelien, C. Haslekås, and L.A. Meza-Zepeda. (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.

Meza, T.J., D. Kamfjord, A. Hakelien, 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. Transgen. Res. 10: 53-67.

Millwood, R.J., M.D. Halfhill, D. Harkins, R. Russotti, and C.N. Stewart, Jr. (2003) Instrumentation and methodology for quantifying *GFP* Fluorescence in intact plant organs. BioTechniques. 34: 638-643.

Mlynárová, L., A. Loonen, J. Heldens, R.C Jansen, P. Keizer, W.J. Stiekema, and J.P. Nap. (1994) Reduced position effectin mature transgenic plants conferred by the chicken lysozyme matrix-associated region. Plant Cell 6: 417-426.

Mlynárová, L., R.C. Jansen, A.J. Conner, W.J. Stiekema, and J.P. Nap. (1995) The MAR-mediated reduction in position effect can be uncoupled from copy number-dependent expression in transgenic plants. Plant Cell 7: 599–609.

Mlynárová, L., A. Loonen, E. Mietkiewska, R.C. Jansen, and J. Nap. (2002) Assembly of two transgenes in an artificial chromatin domain gives highly coordinated expression in tobacco. Genetics 160: 727–740.

Mlynárová, L., A. Hricová, A. Loonen, and J. Nap. (2003) The Presence of a chromatin boundary appears to shielda transgene in tobacco from RNA silencing. Plant Cell 15: 2203–2217.

Mohamed, R. (2006) Expression and Function of Populus Homologs to TERMINAL FLOWER 1 Genes: Roles in Onset of Flowering and Shoot Phenology. PhD Thesis, Oregon State University, Corvallis, Oregon USA. 136 pp.

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

Moon, J., S. Suh , H. Lee, K. Choi, C.B. Hong, N. Paek, S. Kim, and I. Lee. (2003) The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. Plant J. 35: 612-623.

Morina, K., O. 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.

Morris K.V., S.W. Chan, S.E. Jacobsen, and D.J. Looney. (2004) Small interfering RNA-induced transcriptional gene silencing in human cells. Science 305: 1289–1292.

Müller A.J., R.R. Mendel, J. Schiemann, C. Simoens, and D. Inzé. (1987) High meiotic stability of a foreign gene introduced into tobacco by *Agrobacterium*-mediated transformation. Mol. Gen. Genet. 207: 171-175.

Neumann, K., W. Droge-Laser, S. Köhne, and I. Broer. (19**97)** Heat Treatment Results in a Loss of Transgene-Encoded Activities in Several Tobacco Lines. Plant Physiol. 115: 939-947.

Ottaviani, M.P., C.H. Hanisch ten Cate, and L.V. Doting. (1992) Expression of introduced genes after tuber propagation of transgenic potato plants. Plant Breed. 109: 89-96.

Paszkowski, J. and S.A.Whitham.(2001) Gene silencing and DNA methylation processes. Curr. Opin. Plant Biol. 4: 123-129.

Pavingerová D, Dostál J, Bísková R, Benetka V (1994) Somatic embryogenesis and Agrobacterium-mediated transformation of chrysanthemum. Plant Sci 97:95-101.Peña, L. and A. Seguin. (2001) Recent advances in the genetic transformation of trees. Trends Biotechnol. 19: 500–506.

Pélissier, T., S. Thalmeir, D. Kempe, H.L. Sänger, and M. Wassenegger. (1999) Heavy *de novo* methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. Nucleic Acids. Res**.** 27**:** 1625–1634.

Peña, L., M. Martín-Trillo, J. Juárez, J.A. Pina, L. Navarro, and J.M. Martınez-Zapater. (2001) Constitutive expression of Arabidopsis LEAFY or APETALA1 genes in citrus reduces their generation time. Nat. Biotechnol. 19: 263-267.

Petersen, K., R. Leah, S. Knudsen, and V. Cameron-Mills. (2002) Matrix attachment regions (MARs) enhance transformation frequencies and reduce variance of transgene expression in *BAR*ley. Plant Mol. Biol. 49: 45-58.

Phillip, 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.

Pilate, G., E. Guiney, K. Holt, M. Petit-Conil, C. Lapierre, J. Lepl´e, B. Pollet, I. Mila, E.A. Webster, H.G. Marstorp, D.W. Hopkins, L. Jouanin, W. Boerjan, W. Schuch, D. Cornu, and C. Halpin. (2002) Field and pulping performances of transgenic trees with altered lignification. Nature Biotechnol. 20: 607-612.

Plasterk, R.H.A. and R.F. Ketting. (2000) The silencing of the genes. Curr. Opin. in Genet. and Dev. 10: 562-567.

Preuss, D. (1999) Chromatin silencing and Arabidopsis development: a role for polycomb proteins. Plant J. 11: 765-767.

Rathore, K.S. and G. Sunilkumar. (2005) Homologous Promoter Use in Genetic Modification. ISB News Report.

Risseeuw, E., M.E. Fanke-van Dijk, and P.J. Hooykaas. (1997) Gene targeting and instability of *Agrobacterium* T-DNA loci in the plant genome. Plant J. 11: 717-728.

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(2): 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.

Rozen, S. and H.J. Skaletsky. (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.

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

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

Schimizu, K., M. Takahashi, N. Goshima, S. Kawakami, K. Irifune, and H. Morikawa. (2001) Presence of an SAR-like sequence in junction regions between an introduced transgene and genomic DNA of cultured tobacco cells: its effect on transformation frequency. The Plant Journal. 26(4): 375-384.

Schmlling, T. and H. Rhrig. (1995). Gene silencing in transgenic tobacco hybrids: frequency of the event and visualization of somatic inactivation pattern. Mol. Gen. Genet. 249: 375-390.

Schöffl, F., G. Schroder, M. Kliem, and M. Rieping. (1993) An SAR sequence containing 395 bp DNA fragment mediates enhanced, gene-dosage-correlated expression of a chimaeric heat shock gene in transgenic tobacco plants. Transgen. Res. 2: 93-100.

Sedjo, R. (2004) Genetically engineered trees: promises and concerns. Resources for the Future, Washington, DC. 54 pp. (http://www.rff.org/rff/Documents/RFF-RPT-Forest.pdf, last accessed May 2, 2006).

Sidorenko, L., W. Bruce, S. Maddock, L. Tagliani, X. Li, M. Daniels, and T. Peterson1. (2003) Functional analysis of two matrix attachment region (MAR) elements in transgenic maize plants. Transgen. Res. 12: 137-154.

Sijen, T., I. Vijn, A. Rebocho, R. van Blokland, D. Roelofs, J.N. Mol, and J.M. Kooter. (2001) Transcriptional and post-transcriptional gene silencing are mechanistically related. Curr. Biol. 11: 436-440.

Skinner, J.S., R. Meilan, A.M. Brunner, and S.H. Strauss. (2000) Options for genetic engineering of floral sterility in forest trees. In S.M. Jain & S.C. Minocha (Eds.), Molecular Biology of Woody Plants, Kluwer Academic Publishers, Dordrecht, The Netherlands Pp 135-153.

Smith, N.A., S.P. Singh, M.B. Wang, P.A. Stoutjesdijk, A.G. Green, and P.M. Waterhouse. (2000) Total silencing by intron-spliced hairpin RNAs. Nature 407: 319-320.

Somers, D.A., and I. Makarevitch. (2004) Transgene integration in plants: poking or patching holes in promiscuous genomes. Curr. Opin. Biotechnol. 15(2): 126-131.

Sós-Hegedűs, A., A. Lovas, M. Kondrák, G. Kovács, and Z. Bánfalvi. (2005) Active RNA silencing at low temperature indicates distinct pathways for antisense-mediated gene-silencing in potato. Plant Mol. Biol. 59: 595-602.

Sousa-Majer, M.J.d., N.C. Turner, D.C. Hardie, R.L. Morton, B. Lamont, and T. J.V. Higgins. (2004) Response to water deficit and high temperature of transgenic peas (Pisum sativum L.) containing a seed-specific alpha-amylase inhibitor and the subsequent effects on pea weevil (Bruchus pisorum L.) survival. J. Exp. Bot. 55 (396): 497-505.

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

Stam, M., N.M. Joseph, and J.M. Kooter. (1997) The silencing of genes in transgenic plants. Ann. Bot. 79: 3-12.

Stoutjesdijk, P.A., S.P. Singh, Q. Liu, C.J. Hurlstone, P.A. Waterhouse, and A.G. Green. (2002) hpRNA-mediated targeting of the *Arabidopsis* FAD2 gene gives highly efficient and stable silencing. Plant Physiol. 129: 1723-1731.

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.

Strauss, S.H., S.P. DiFazio, and R. Meilan. (2001) Genetically modified poplars in context. Forestry Chronicle. 77(2): 271-279.

Strauss, S.H., A.M. Brunner, V.B. Busov, C. Ma, and R. Meilan. (2004) Ten lessons from 15 years of transgenic Populus research. Forestry. 77: 455-465.

Sugita, M. AND W. Gruissem. (1987) Developmental, organ-specific, and light-dependent expression of the tomato ribulose-1,5-bisphosphate carboxylase small subunit gene family. Proc. Natl. Acad. Sci. USA. 84: 7104-7108.

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.

Szittya, G., D. Silhavy, A. Molnár, Z. Havelda, A. Lovas, L. Lakatos, Z. Banfalvi, and J. Burgyán. (2003) Low temperature inhibits RNA silencing-mediated defence by the control of siRNA generation. EMBO J. 22(3): 633-640.

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.

Thompson, E.M., E. Christians, M.G. Stinnakre, J.P. Renard. (1994) Scaffold attachment regions stimulate HSP70.1 expression in mouse preimplantation embryos but not in differentiated tissues. Mol. Cell Biol. 14: 4694-4703.

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

Ülker, B., G.C. Allen, W.F. Thompson, S. Spiker, and A.K. Weissinger. (1999) A tobacco matrix attachment region reduces the loss of transgene expression in the progeny of transgenic tobacco plants. Plant J. 18: 253-263.

Vain, P., B. Worland, A. Kohli, J.W. Snape, P. Christou, G.C. Allen, and W.F. Thompson. (1999) Matrix attachment regions increase transgene expression levels and stability in transgenic rice plants and their progeny. Plant J. 18: 233-242.

Valenzuela, S., C. Balocchi, and J. Rodríguez. (2006) Transgenic trees and forestry biosafety. Electronic J. Biotechnol. 9(3): 335-339.

van Blokland, R., M. ten Lohuis, and P. Meyer. (1997) Condensation of chromatin in transcriptional regions of an inactivated plant transgene: evidence for an active role of transcription in gene silencing. Mol. Gen. Genet. 257: 1-13.

Van der Geest, A.H.M., G.E .Hall, Jr., S. Spiker, and T.C. Hall. (1994) The beta-phaseolin gene is flanked by matrix attachment regions. Plant J. 6: 413-423.

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

Van Doorsselaere, J., M. Baucher, E. Chognot, B. Chabbert, and M. Tollier. (1995) A novel lignin in poplar trees with reduced caffeic acid/5-hydroxyferulic acid O-methyltransferase activity. Plant J. 8: 855–64.

Vance, V. and H. Vaucheret. (2001) RNA silencing in plants - defense and counterdefense. Science 292: 2277-2280.

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

Vaucheret, H., C. Bclin, and M. Fagard. (2001b) Post-transcriptional gene silencing in plants. J. Cell Sci. 114: 3083-3091.

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

Wagner, A, L. Phillips, R.D. Narayan, J.D. Moody, and B. Geddes. (2005) Gene silencing studies in the gymnosperm species *Pinus radiata*. Plant Cell Rep. 24: 95-102.

Walter, C., I. Broer, D. Hillemann, and A. Pühler. (1992) High frequency, heat treatment-induced inactivation of the phosphinothricin resistance gene. Mol. Gen. Genet. 235: 189-196.

Wassenegger, M. (2000) RNA-directed DNA methylation. Plant Mol. Biol. 43: 203-220.

Waterhouse, P.M., H.W. Graham, and M.B. Wang. (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci. USA. 95: 13959-13964.

Weigel, D. and O. Nilsson. (1995) A developmental switch sufficient for flower initiation in diverse plants. Nature 377: 495–500.

Wesley, S.V., C.A. Helliwell, N.A Smith., M.B. Wang, D.T. Rouse, Q. Liu, P.S. Goodings, S.P. Singh, D. Abbott, P.A. Stoutjesdijk, S.P. Robinson, A.P. Gleave, A.G. Green, and P.M. Waterhouse. (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J. 6: 581–590.

Whetten, R.W., J.J. MacKay, and R.R. Sederoff. (1998) Recent advances in understanding lignin biosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49: 585–609.

Yan, L., A. Loukoianov,A. Blechl, G. Tranquilli, W. Ramakrishna, P. SanMiguel, J.L. Bennetzen, V. Echenique, and J. Dubcovsky. (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303: 1640-1644.

Zupan, J., T.R. Muth, O. Draper, and P. Zambryski. (2000) The transfer of DNA from Agrobacterium tumefaciens into plants: a feast of fundamental insights. Plant J. 23(1): 11-28.

APPENDICES

Appendix A. ANOVA of GFP expression among different areas within a single leaf and multiple comparisons. L1 ~ Ln represents different leaf areas delimited by the mid-vein (V) (see Appendix F2.2) Multiple comparisons used 95% simultaneous confidence intervals for specified linear combinations by the Tukey method. Two areas with significant difference in mean GFP expression are flagged by '****'. Abbreviations: SE: standard error; L.B.: low bound; U.B.: upper bound.

Event#1: 3MGB242

	Df	SS	MS	F Value	Pr(F)
Leaf area	6	40971.48	6828.579	8.229963	0.00004749856
Residuals	26	21572.77	829.722		

	Estimate	Std.Error	Lower Bound	Upper Bound	
L1-L2	-1.40	18.2	-59.5	56.70	
L1-L3	-11.70	24.1	-88.6	65.20	
L1-L4	8.55	19.3	-53.1	70.20	
L1-L5	4.55	19.3	-57.1	66.20	
L1-L6	3.13	21.0	-64.0	70.20	
L1-V	-74.70	15.8	-125.0	-24.40	****
L2-L3	-10.30	24.1	-87.2	66.60	
L2-L4	9.95	19.3	-51.7	71.60	
L2-L5	5.95	19.3	-55.7	67.60	
L2-L6	4.53	21.0	-62.6	71.60	
L2-V	-73.30	15.8	-124.0	-23.00	****
L3-L4	20.30	24.9	-59.3	99.80	
L3-L5	16.30	24.9	-63.3	95.80	
L3-L6	14.80	26.3	-69.1	98.70	
L3-V	-63.00	22.3	-134.0	8.18	
L4-L5	-4.00	20.4	-69.0	61.00	
L4-L6	-5.42	22.0	-75.6	64.80	
L4-V	-83.30	17.0	-138.0	-28.90	****
L5-L6	-1.42	22.0	-71.6	68.80	
L5-V	-79.30	17.0	-134.0	-24.90	****
L6-V	-77.80	19.0	-138.0	-17.30	****

Event #2: 3GB183

	Df	SS	MS	F Value	Pr(F)
Leaf area	6	27317.9	4553.0	4.52	0.0014
Residuals	40	40256.9	1006.4		

	Estimate	Std.Error	Lower Bound	Upper Bound
L1-L2	13.00	18.3	-43.8	69.80
L1-L3	24.40	20.5	-39.1	88.00
L1-L4	36.20	25.9	-44.2	117.00
L1-L5	8.04	17.1	-45.1	61.20

L1-L6	21.40	15.9	-27.8	70.60	
L1-V	-42.20	16.7	-94.1	9.67	
L2-L3	11.40	20.5	-52.1	75.00	
L2-L4	23.20	25.9	-57.2	104.00	
L2-L5	-4.96	17.1	-58.1	48.20	
L2-L6	8.42	15.9	-40.8	57.60	
L2-V	-55.20	16.7	-107.0	-3.33	****
L3-L4	11.70	27.5	-73.5	97.00	
L3-L5	-16.40	19.4	-76.7	43.90	
L3-L6	-3.00	18.3	-59.8	53.80	
L3-V	-66.60	19.1	-126.0	-7.47	****
L4-L5	-28.10	25.1	-106.0	49.70	
L4-L6	-14.70	24.2	-89.9	60.40	
L4-V	-78.40	24.8	-155.0	-1.42	****
L5-L6	13.40	14.5	-31.6	58.30	
L5-V	-50.30	15.4	-98.1	-2.42	****
L6-V	-63.60	14.0	-107.0	-20.20	****

Event #3: 3GB141

	Df	SS	MS	F Value	Pr(F)
Leaf area	6	39969.2	6661.5	6.34	0.00018
Residuals	32	33615.6	1050.487		

	Estimate	Std.Error	Lower Bound	Upper Bound	
L1-L2	-17.80	18.0	-74.5	38.90	
L1-L3	-10.20	20.9	-76.0	55.50	
L1-L5	-32.00	20.9	-97.8	33.80	
L1-L6	4.70	19.6	-57.0	66.40	
L1-L7	-4.50	22.9	-76.5	67.50	
L1-V	-79.20	16.7	-132.0	-26.60	****
L2-L3	7.54	20.3	-56.3	71.40	
L2-L5	-14.20	20.3	-78.1	49.60	
L2-L6	22.50	19.0	-37.2	82.10	
L2-L7	13.30	22.4	-57.0	83.60	
L2-V	-61.40	16.0	-112.0	-11.20	****
L3-L5	-21.80	22.9	-93.8	50.30	
L3-L6	14.90	21.7	-53.4	83.30	
L3-L7	5.75	24.8	-72.1	83.60	
L3-V	-69.00	19.2	-129.0	-8.68	****
L5-L6	36.70	21.7	-31.6	105.00	
L5-L7	27.50	24.8	-50.3	105.00	
L5-V	-47.20	19.2	-107.0	13.10	
L6-L7	-9.20	23.7	-83.6	65.20	
L6-V	-83.90	17.8	-140.0	-28.10	****
L7-V	-74.70	21.3	-142.0	-7.64	****

Event #4: 3MGB149

	Df	SS	MS	F Value	Pr(F)
Leaf area	5	20294.8	4058.951	2.69	0.043
Residuals	26	39222.7	1508.566		

	Estimate	Std.Error	Lower Bound	Upper Bound	
L1-L2	-0.0714	21.6	-66.5	66.30	
L1-L3	-2.5700	26.8	-84.9	79.80	
L1-L5	-22.8000	24.3	-97.6	52.00	
L1-L6	-17.3000	24.3	-92.1	57.50	
L1-V	-62.1000	20.1	-124.0	-0.31	****
L2-L3	-2.5000	27.5	-86.9	81.90	
L2-L5	-22.8000	25.1	-99.8	54.30	
L2-L6	-17.3000	25.1	-94.3	59.80	
L2-V	-62.0000	21.0	-126.0	2.45	
L3-L5	-20.3000	29.7	-111.0	70.90	
L3-L6	-14.8000	29.7	-106.0	76.40	
L3-V	-59.5000	26.3	-140.0	21.30	
L5-L6	5.5000	27.5	-78.9	89.90	
L5-V	-39.3000	23.8	-112.0	33.80	
L6-V	-44.8000	23.8	-118.0	28.30	

Event #5: 7MGB291

	Df	SS	MS	F Value	Pr(F)
Leaf area	6	11470.1	1911.7	5.50	0.0011
Residuals	24	8347.3	347.8		

	Estimate	Std.Error	Lower Bound	Upper Bound	
L1-L2	11.20	13.2	-31.2	53.50	
L1-L3	13.80	14.2	-31.9	59.60	
L1-L6	4.33	15.2	-44.6	53.20	
L1-L7	10.00	13.2	-32.3	52.30	
L1-L8	-1.00	15.2	-49.9	47.90	
L1-V	-39.50	13.2	-81.8	2.84	
L2-L3	2.67	12.0	-36.0	41.30	
L2-L6	-6.83	13.2	-49.2	35.50	
L2-L7	-1.17	10.8	-35.7	33.40	
L2-L8	-12.20	13.2	-54.5	30.20	
L2-V	-50.70	10.8	-85.2	-16.10	****
L3-L6	-9.50	14.2	-55.2	36.20	
L3-L7	-3.83	12.0	-42.5	34.80	
L3-L8	-14.80	14.2	-60.6	30.90	
L3-V	-53.30	12.0	-92.0	-14.70	****
L6-L7	5.67	13.2	-36.7	48.00	
L6-L8	-5.33	15.2	-54.2	43.60	
L6-V	-43.80	13.2	-86.2	-1.49	****
L7-L8	-11.00	13.2	-53.3	31.30	
L7-V	-49.50	10.8	-84.1	-14.90	****
L8-V	-38.50	13.2	-80.8	3.84	

Event #6: 7MGB565

	Df	SS	MS	F Value	Pr(F)
Leaf area	8	23525.5	2940.7	4.49	0.00051
Residuals	43	28137.98	654.372		

	Estimate	Std.Error	Lower Bound	Upper Bound	
L1-L2	-14.700	15.5	-65.30	35.800	
L1-L3	-10.400	16.2	-63.20	42.400	
L1-L4	-28.400	17.2	-84.40	27.600	
L1-L5	-4.000	16.2	-56.80	48.800	
L1-L6	24.000	16.2	-28.80	76.800	
L1-L7	-15.400	16.2	-68.20	37.400	
L1-L8	-14.600	16.2	-67.40	38.200	
L1-V	-48.700	13.6	-93.10	-4.200	****
L2-L3	4.330	15.5	-46.20	54.900	
L2-L4	-13.700	16.5	-67.60	40.200	
L2-L5	10.700	15.5	-39.80	61.300	
L2-L6	38.700	15.5	-11.80	89.300	
L2-L7	-0.667	15.5	-51.20	49.900	
L2-L8	0.133	15.5	-50.40	50.700	
L2-V	-33.900	12.8	-75.70	7.840	
L3-L4	-18.000	17.2	-74.00	38.000	
L3-L5	6.400	16.2	-46.40	59.200	
L3-L6	34.400	16.2	-18.40	87.200	
L3-L7	-5.000	16.2	-57.80	47.800	
L3-L8	-4.200	16.2	-57.00	48.600	
L3-V	-38.300	13.6	-82.70	6.200	
L4-L5	24.400	17.2	-31.60	80.400	
L4-L6	52.400	17.2	-3.62	108.000	
L4-L7	13.000	17.2	-43.00	69.000	
L4-L8	13.800	17.2	-42.20	69.800	
L4-V	-20.300	14.8	-68.50	28.000	
L5-L6	28.000	16.2	-24.80	80.800	
L5-L7	-11.400	16.2	-64.20	41.400	
L5-L8	-10.600	16.2	-63.40	42.200	
L5-V	-44.700	13.6	-89.10	-0.199	****
L6-L7	-39.400	16.2	-92.20	13.400	
L6-L8	-38.600	16.2	-91.40	14.200	
L6-V	-72.700	13.6	-117.00	-28.200	****
L7-L8	0.800	16.2	-52.00	53.600	
L7-V	-33.200	13.6	-77.70	11.200	
L8-V	-34.000	13.6	-78.50	10.400	

Event #7: 7GB381

	Df	SS	MS	F Value	Pr(F)
Leaf area	8	3692.1	461.5	2.042319	0.06502535
Residuals	41	9264.9	225.9733		

Event #8: 7GB77

	Df	SS	MS	F Value	Pr(F)
Leaf area	8	2433.7	304.2	0.80	0.61
Residuals	40	15199.6	379.9		

Appendix B. *BAR* coding sequence plus the transit peptide (5' to 3').
CpG sequences are in red. Underlined sequences indicate the regions studied for methylation analysis. Sequences in bold were used to made double-stranded RNAi constructs.

```
  1 GAATTCGAGC TCGGTACCCG GGGATCTACC ATGAGCCCAG AACGACGCCC GGCCGACATC
 61 CGCCGTGCCA CCGAGGCGGA CATGCCGGCG GTCTGCACCA TCGTCAACCA CTACATCCA
121 ACAAGCACGG TCAACTTCCG TACCGAGCCG CAGGAACCGC AGGAGTGGAC GGACGACCTC
181 GTCCGTCTGC GGGAGCGCTA TCCCTGGCTC GTCGCCGAGG TGGACGGCGA GGTCGCCGGC
241 ATCGCCTACG CGGGCCCCTG GAAGGCACGC AACGCCTACG ACTGGACGGC CGAGTCGACC
301 GTGTACGTCT CCCCCCGCCA CCAGCGGACG GGACTGGGCT CCACGCTCTA CACCCACCTG
361 CTGAAGTCCC TGGAGGCACA GGGCTTCAAG AGCGTGGTCG CTGTCATCGG GCTGCCCAAC
421 GACCCGAGCG TGCGCATGCA CGAGGCGCTC GGATATGCCC CCGCGGCAT GCTGCGGGCG
481 GCCGGCTTCA AGCACGGAA CTGGCATGAC GTGGGTTTCT GGCAGCTGGA CTTCAGCCTG
541 CCGGTACCGC CCCGTCCGGT CCTGCCCGTC ACCGAGATCT GATGACCCGG GGGATCCCTG
601 CAGGCATGCA AGCTT
```

Appendix C. *Arabidopsis rbc*S promoter sequence (5' to 3'). CpG sequences are in red. Underlined sequences indicate the regions studied for methylation analysis. Sequences in bold were used to make double-stranded RNAi constructs.

1 GAATTCAAAT TTATTATGTG TTTTTTTTCC GTGGTCGAGA TTGTGTATTA TTCTTTAGTT
61 ATTACAAGAC TTTTAGCTAA AATTTGAAAG AATTTACTTT AAGAAAATCT TAACATCTGA
121 GATAATTTCA GCAATAGATT ATATTTTTCA TTACTCTAGC AGTATTTTTG CAGATCAATC
181 GCAACATATA TGGTTGTTAG AAAAAATGCA CTATATATAT ATATATTATT TTTTCAATTA
241 AAAGAGCATG ATATATAATA TATATATATA TATATATATG TGTGTGTGTA TATGGTCAAA
301 GAAATTCTTA TACAAATATA CACGAACACA TATATTTGAC AAAATCAAAG TATTACACTA
361 AACAATGAGT TGGTGCATGG CCAAAACAAA TATGTAGATT AAAAATTCCA GCCTCCAAAA
421 AAAAATCCAA GTGTTGTAAA GCATTATATA TATATAGTAG ATCCCAAATT TTTGTACAAT
481 TCCACACTGA TCGAATTTTT AAAGTTGAAT ATCTGACGTA GGATTTTTTT AATGTCTTAC
541 CTGACCATTT ACTAATAACA TTCATACGTT TTCATTTGAA ATATCCTCTA TAATTATATT
601 GAATTTGGCA CATAATAAGA AACCTAATTG GTGATTTATT TTACTAGTAA ATTTCTGGTG
661 ATGGGCTTTC TACTAGAAAG CTCTCGGAAA ATCTTGGACC AAATCCATAT TCCATGACTT
721 CGATTGTTAA CCCTATTAGT TTTCACAAAC ATACTATCAA TATCATTGCA ACGGAAAAGG
781 TACAAGTAAA ACATTCAATC CGATAGGGAA GTGATGTAGG AGGTTGGGAA GACAGGCCCA
841 GAAAGAGATT TATCTGACTT GTTTTGTGTA TAGTTTTCAA TGTTCATAAA GGAAGATGGA
901 GACTTGAGAA GTTTTTTTTG GACTTTGTTT AGCTTTGTTG GGCGTTTTTT TTTTTTGATC
961 AATAACTTTG TTGGGCTTAT GATTTGTAAT ATTTTCGTGG ACTCTTTAGT TTATTTAGAC
1021 GTGCTAACTT TGTTGGGCTT ATGACTTGTT GTAACATATT GTAACAGATG ACTTGATGTG
1081 CGACTAATCT TTACACATTA AACATAGTTC TGTTTTTTGA AAGTTCTTAT TTTCATTTTT
1141 ATTTGAATGT TATATATTTT TCTATATTTA TAATTCTAGT AAAAGGCAAA TTTTGCTTTT
1201 AAATGAAAAA AAT**ATATATT CCACAGTTTC ACCTAATCTT ATGCATTTAG CAGTACAAAT**
1261 **TCAAAAATTT CCCATTTTTA TTCATGAATC ATACCATTAT ATATTAACTA AATCCAAGGT**
1321 **AAAAAAAAGG TATGAAAGCT CTATAGTAAG TAAAATATAA ATTCCCCATA AGGAAAGGGC**
1381 **CAAGTCCACC AGGCAAGTAA AATGAGCAAG CACCACTCCA CCATCACACA ATTTCACTCA**
1441 **TAGATAACGA TAAGATTCAT GGAATTATCT TCCACGTGGC ATTATTCCAG CGGTTCAAGC**
1501 **CGATAAGGGT CTCAACACCT CTCCTTAGGC CTTTGTGGCC GTTACCAAGT AAAATTAACC**
1561 **TCACACATAT CCACACTCAA AATCCAACGG TGTAGATCCT AGTCCACTTG AATCTCATGT**
1621 **ATCCTAGACC CTCCGATCAC TCCAAAGCTT GTTCTCATTG TTGTTATCAT TATATATAGA**
1681 **TGACCAAAGC ACTAGACCAA** ACCTCAGTCA CACAAAGAGT AAAGAAGAAC A

APPENDIX TABLES

Appendix T2. 1. Number of transgenic events produced for the four different clone x construct combinations.

Construct	Clone	
	353-53	717-1B4
pG3KGB	99	98
pG3MKGB	108	99

Appendix T2. 2. Field design for reporter gene stability field trial.
The plants were planted in four blocks with a random split-block design. The field site was divided into four blocks. Transgenic events from two poplar clones 353-53 and 717-1B4 were separated and randomly planted in one of two subplots (A & B).

Block	A	B
I	353-53	-1B4
II	717-1B4	353-53
III	717-1B4	353-53
IV	353-53	717-1B4

Appendix T2. 3. *GFP* expression analysis with PROC MIXED to text event, block, clone, MAR, and year effect.

Analysis was performed according to randomized split block design using *GFP* field expression data (year 2004 and 2005). EVENT, YEAR, and BLOCK*CLONE were treated as random effect. CLONE, MAR, and BLOCK were considered as fixed effect.

.

SAS CODE:

```
proc mixed data=loggfp method=REML;
class event block clone MAR;
model loggfp=block clone MAR clone*MAR;
random event(clone MAR) block*clone year;
run;
```

Abbreviated output

The Mixed Procedure

Covariance Parameter Estimates

Cov Parm	Estimate
EVENT(CLONE*MAR)	0.1404
BLOCK*CLONE	0.005542
YEAR	0.000710
RESIDUAL	0.08597

Type 3 Tests of Fixed Effects

Effect	NDF	DDF	F Value	Pr > F
BLOCK	3	3	2.42	0.2435
CLONE	1	3	0.09	0.7827
MAR	1	400	0.60	0.4388
CLONE*MAR	1	400	0.25	0.6149

Appendix T2. 4. *BAR* expression analysis with PROC MIXED to test MAR, event, and year effect. Analysis was performed using *BAR* expression data from three different years (2003, 2004, and 2005). EVENT and YEAR were treated as random effect, and MAR was considered as fixed effect.

SAS CODE:

```
proc mixed method=type3;
class event year mar;
model sqrtbar=mar year;
random event(MAR);
run;
```

```
                    **Abbreviated output**
                      The Mixed Procedure
                Covariance Parameter Estimates
                    Cov Parm       Estimate
                    EVENT(MAR)       0.3928
                    YEAR             0.002852
                   RESIDUAL          0.08235
                 Type 3 Tests of Fixed Effect
Effect          NDF     DDF     F Value     Pr > F
MAR              1      166       4.41      0.0372
```

Appendix T2. 5. Levene's test of MAR, COPY, and DR (direct repeats) on expression variance among or within events. The statistics in the table are tests of absolute residuals from the first model (T2.2 for statistical models used).

Levene's test of MAR, COPY, T-DNA structure (DR), MAR*COPY on expression variance among events						
	Sources	DF	SS	MS	F-Value	Pr > F
GFP	MAR	1	0.23	0.23	4.02	0.0457
	COPY	1	0.55	0.55	11.68	0.0007
	MAR*COPY	3	1.1049	0.368	7.99	0.0001
	REPEATS	1	0.07	0.07	1.69	0.2015
BAR	MAR	1	1.16	1.16	6.71	0.0104
	COPY	1	0.017	0.017	0.6	0.4398
	MAR*COPY	3	0.89	0.297	1.88	0.1359
	REPEATS	1	0.00037	0.00037	0.00	0.9484
Levene's test of MAR, COPY, REPEATS, MAR*COPY on expression variance within events						
GFP	MAR	1	0.0005	0.0005	0.04	0.8439
	COPY	1	0.128	0.128	10.26	0.0014
	MAR*COPY	3	0.1319	0.14397	3.61	0.0147

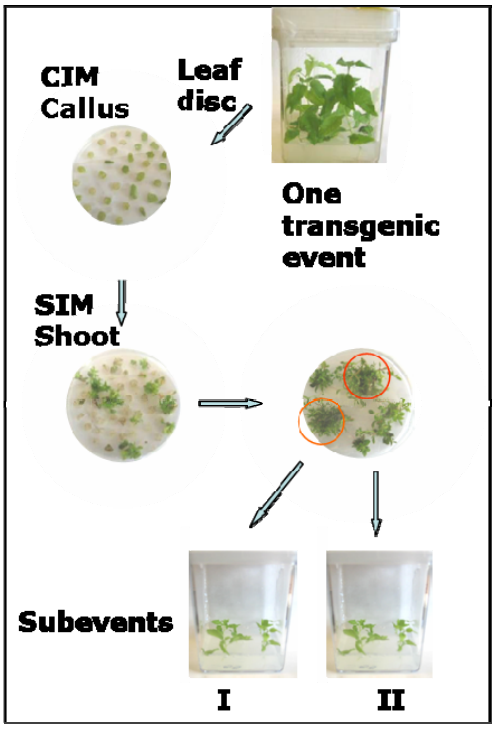
Appendix T3. 1. Transformation efficiency (TE) of RNAi constructs with different parental events.

Parent Event	Pi M-		Bi M-		Pi M+		Bi M+	
	Co-cultivated Explants	Produced Transgenics (TE)	Co-cultivated Explants	Verified Transgenics (TE)	Co-cultivated Explants	Verified Transgenics (TE)	Co-cultivated Explants	Verified Transgenics (TE)
717-C	818	2 (0.2%)	1156	2(0.2%)	813	8(1.0%)	1168	2(0.2%)
717-A	832	8(1.0%)	1741	7(0.4%)	1028	3(0.3%)	1658	5(0.3%)
353-29	2336	7(0.3%)	1781	7(0.4%)	759	2(0.3%)	1839	2(0.1%)
353-38	1735	0(0.0%)	1793	1(0.1%)	1262	2(0.2%)	1967	0(0.0%)

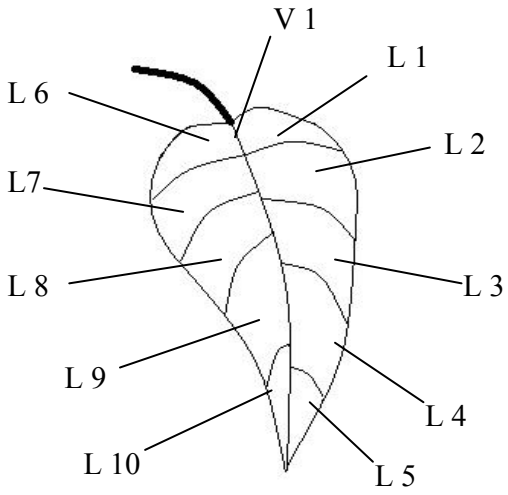
Appendix T3. 2. Number and RNAi suppression efficiency for each of event and construct combinations. Abbreviations: Pi: IR of promoter sequence; Bi: IR of coding sequence; M-: without flanking MARs; M+: with flanking MARs.

Original Line	Construct	No of transgenic Events	Suppression > 90% (% of Total)[a]	Suppression > 50% (% of Total)	Suppression > 0% (% of Total)
717-A	PiM-	8	0 (0%)	4 (50%)	6 (75%)
	PiM+	3	0 (0%)	0 (0%)	3 (100%)
	BiM-	7	6 (86%)	7 (100%)	7 (100%)
	BiM+	5	4 (80%)	4 (80%)	4 (80%)
717-C	PiM-	1	0 (0%)	0 (0%)	0 (0%)
	PiM+	8	0 (0%)	1 (13%)	3 (38%)
	BiM-	1	1(100%)	1(100%)	1(100%)
	BiM+	2	2(100%)	2(100%)	2(100%)
353-29	PiM-	7	0 (0%)	0 (0%)	2 (100%)
	PiM+	2	0 (0%)	0 (0%)	0 (0%)
	BiM-	7	5 (71%)	5 (71%)	5 (71%)
	BiM+	2	1 (50%)	1(50%)	1(50%)
353-38	PiM-	-	-	-	-
	PiM+	2	2(100%)	2(100%)	2(100%)
	BiM-	1	1(100%)	1(100%)	1(100%)
	BiM+	-	-	-	-
Lines Pooled	PiM-	16	0(0%)	4(25%)	10(63%)
	PiM+	15	2(13%)	3(2%)	8 (53%)
	BiM-	16	13 (76%)	14(88%)	14(88%)
	BiM+	9	7 (78%)	8 ((89%)	8 (89%)
Line & MAR pooled	Pi	31	2 (6%)	7 (23%)	18 (58%)
	Bi	25	20(80%)	22(88%)	22(88%)
Line and IR pooled	M-	32	13 (41%)	18(56%)	24(75%)
	M+	24	9(38%)	11(46%)	16(67%)

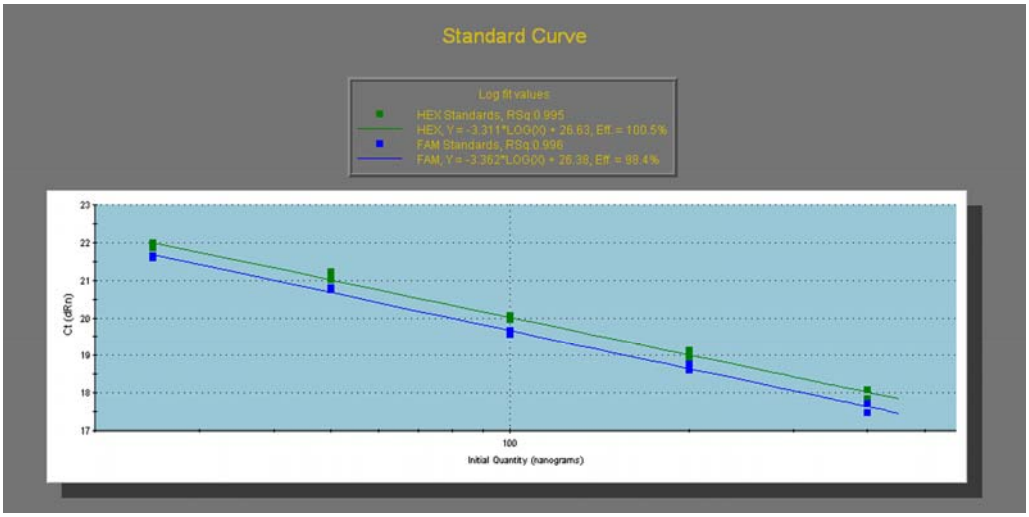
APPENDIX FIGURES



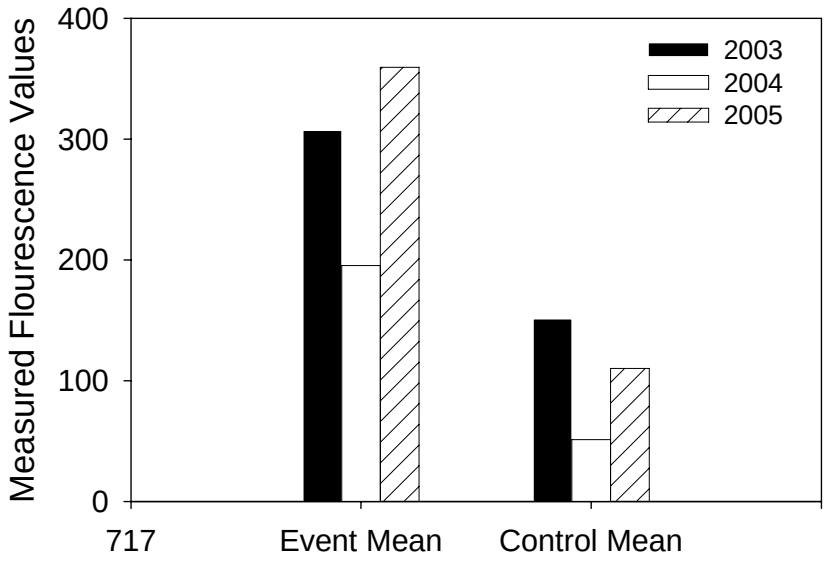
Appendix F2. 1. Organogenesis treatment used to produce subevents.
The diagram show how two subevents were derived from one transgenic event. Leaf explants were taken from plantlets of one transgenic event, and subjected to callus and shoot induction. Shoots from two different leaf discs were rooted in two different magenta boxes, and a single rooted plant from each box was chosen as subevent I and subevent II, respectively. CIM: callus induction medium; SIM: shoot induction medium.



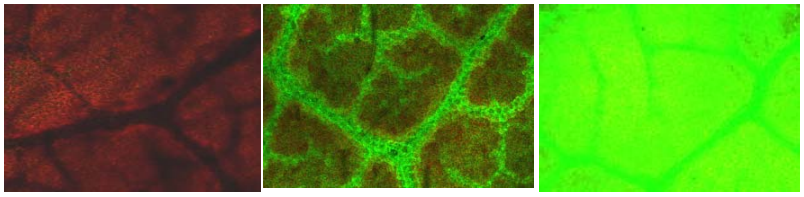
Appendix F2. 2. Measurements of GFP expression variation within a leaf. Measurements were made on the fourth leaf of approximately two-month old plants grown in the greenhouse. Different areas measured are shown: L1 ~ L8: leaf areas delimited by main and secondary veins; V1: main vein area. Leaf is not drawn to scale.



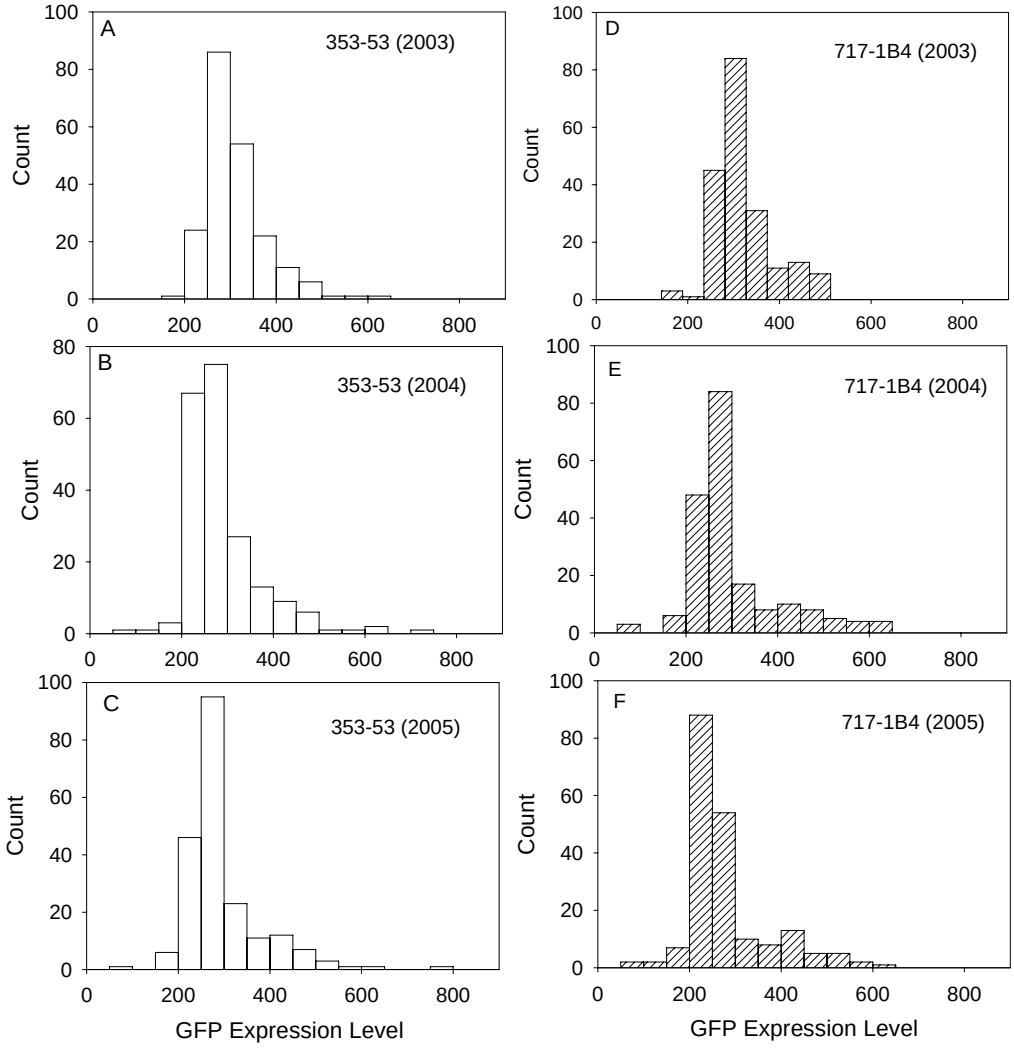
Appendix F2. 3. Standard curve of *GFP* and *PTLF* amplifications with TaqMan® probes. Two-fold serial dilutions of the genomic DNA from one transgenic event were used as PCR templates. Blue (bottom) and green (top) lines represent *PTLF* and *GFP* amplifications, respectively.



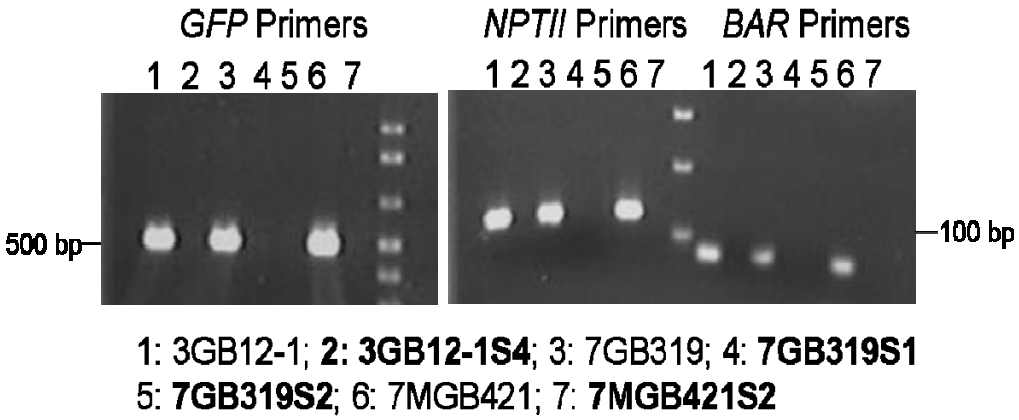
Appendix F2. 4. Mean unnormalized fluorescence levels of 404 transgenic events and 64 non-transgenic controls in the three different years. Measurements were made with a hand-held GFP-meter.



Appendix F2. 5. Confocal microscope images of *GFP* and quantified *GFP* values with a hand-held GFP-meter. Leaves were taken from the field-grown transgenic plants in 2005.

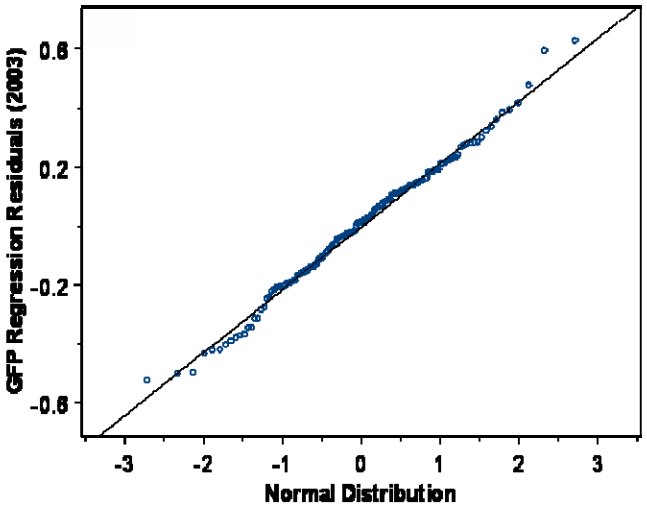


Appendix F2. 6. Normalized *GFP* expression distributions in the two poplar clones in three different years. *GFP* expression was normalized to medium expression value of each of the three years.



Appendix F2. 7. PCR amplification of four subevents and their corresponding parent events with transgene-specific primers. T-DNA sections of the four subevents that were physically lost are indicated in bold.

A



B

One sample Kolmogorov-Smirnov Test of normality of regression residuals.

Data: regression residuals of *GFP* expression of subevents on corresponding parent events.

ks = 0.0419, p-value = 0.709

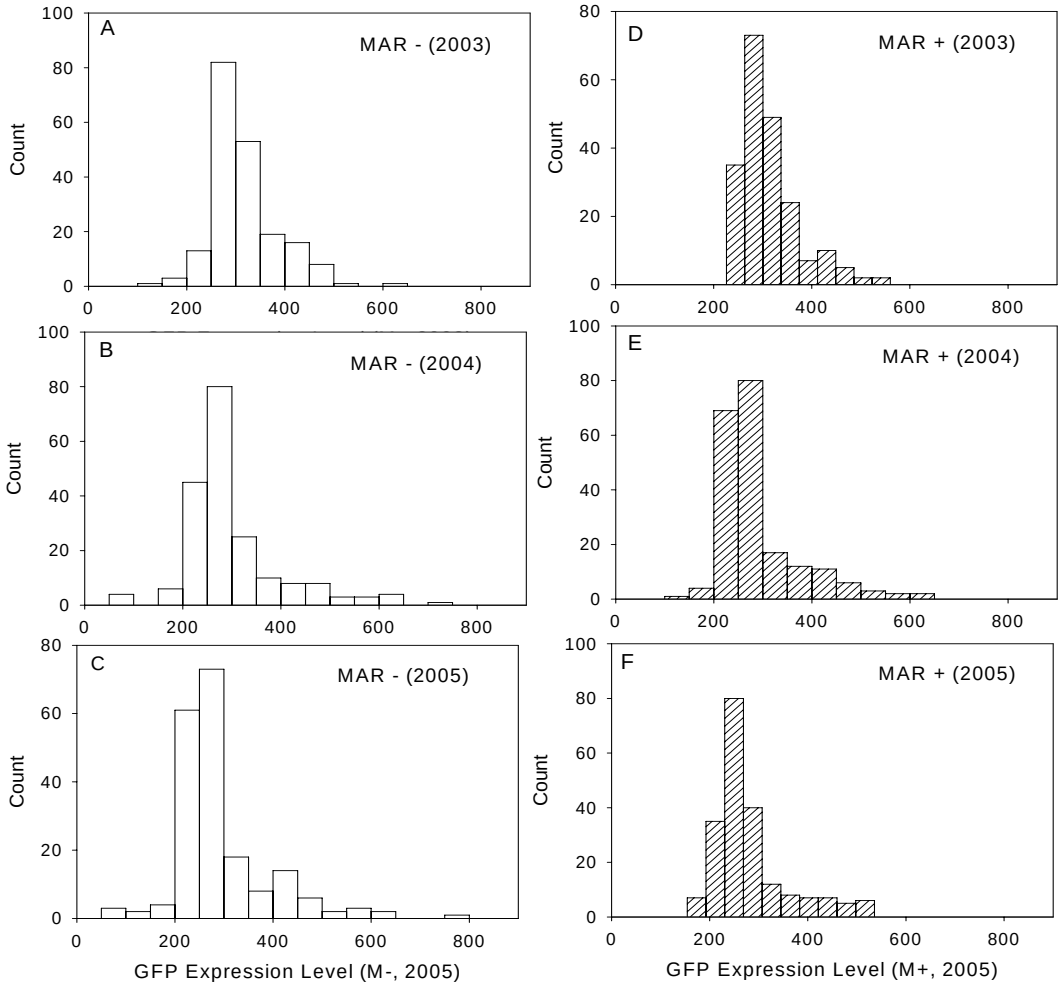
Alternative hypothesis: regression residuals are not normally distributed with estimated parameters.

Sample estimates:

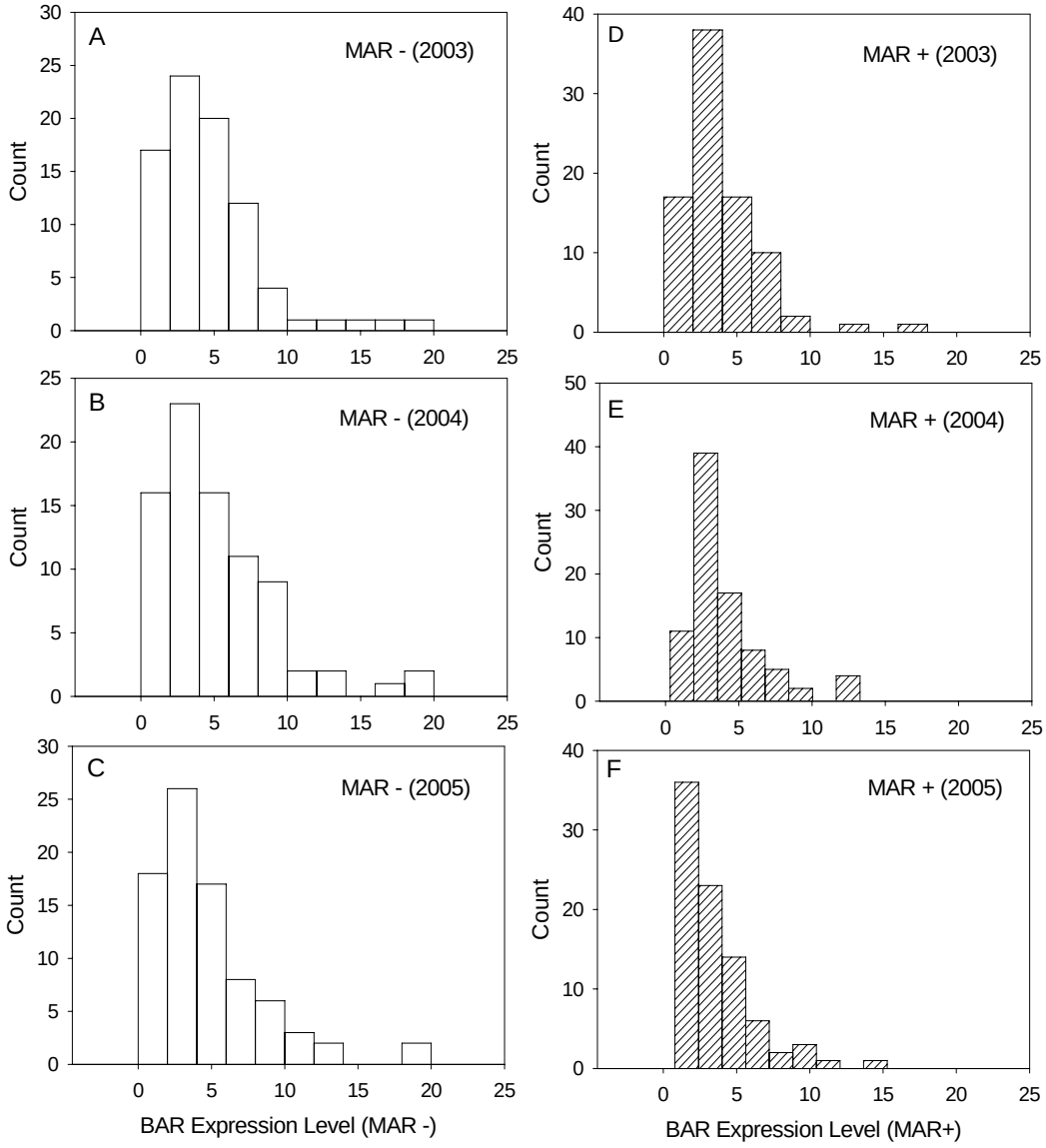
mean: 2.567152e-006

standard deviation: 0.217675

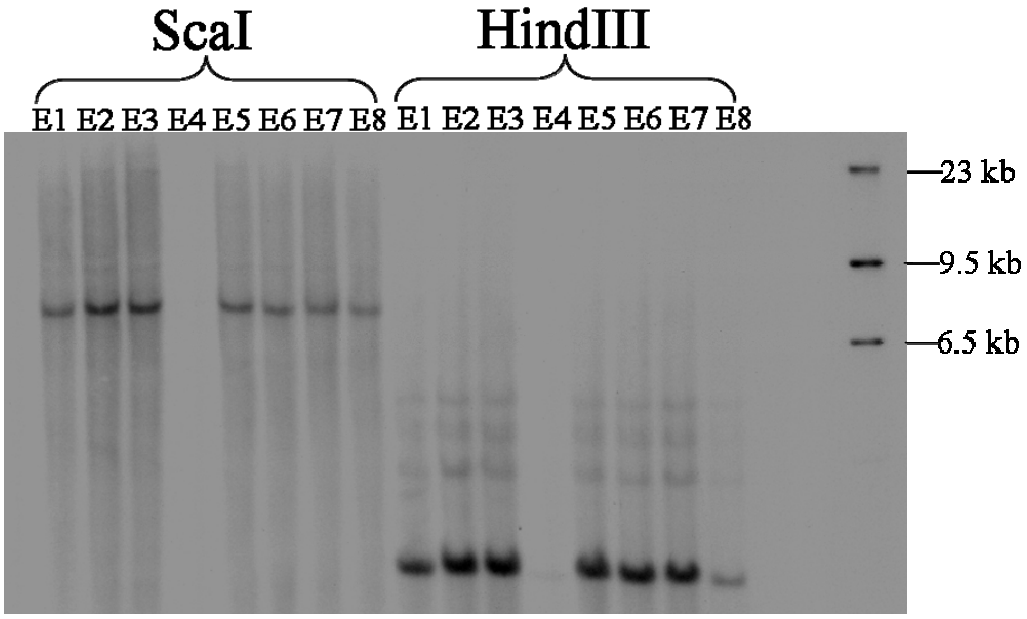
Appendix F2. 8. Normality test of regression residuals. Regression analysis was performed on *GFP* expression levels of 160 subevents and their corresponding parental events in three different years. Normality of residuals was checked with QQ normal plot (A) and Kolmogorov-Smirnov test at a significant level of 0.05 (B).



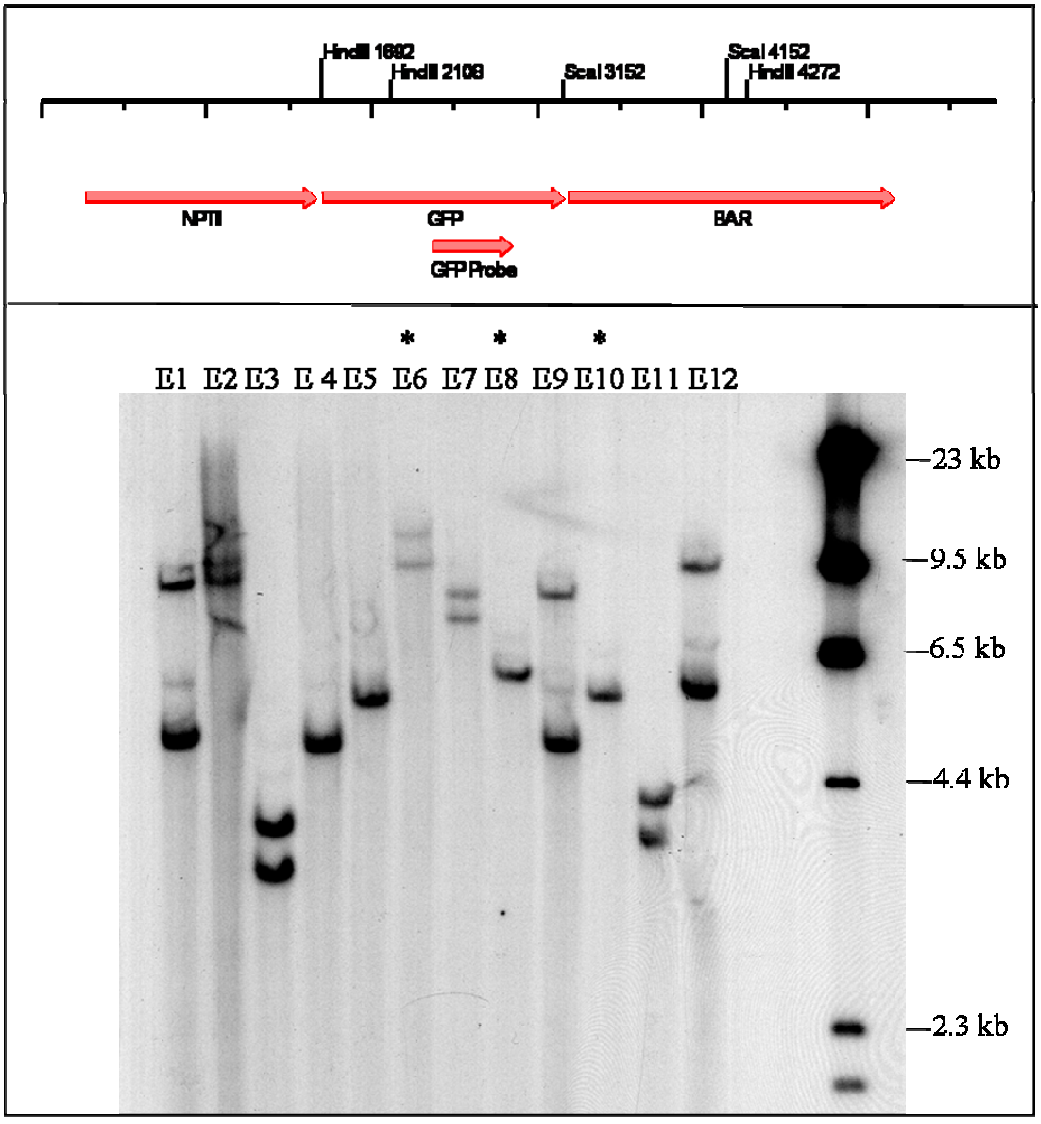
Appendix F2. 9. Normalized *GFP* expression distribution of transgenic events with and without MARs in three different years. *GFP* expression was normalized to the median expression values for each of the three years.A-C: *GFP* expression without MARs. D-F: *GFP* expression with MARs.



Appendix F2. 10. Normalized *BAR* expression distribution of transgenic events with and without MARs in three different years. *BAR* expression was normalized to the median expression values of each of the three years. A-C: *BAR* expression distributions without MARs. D-F: *BAR* expression distributions with MARs.

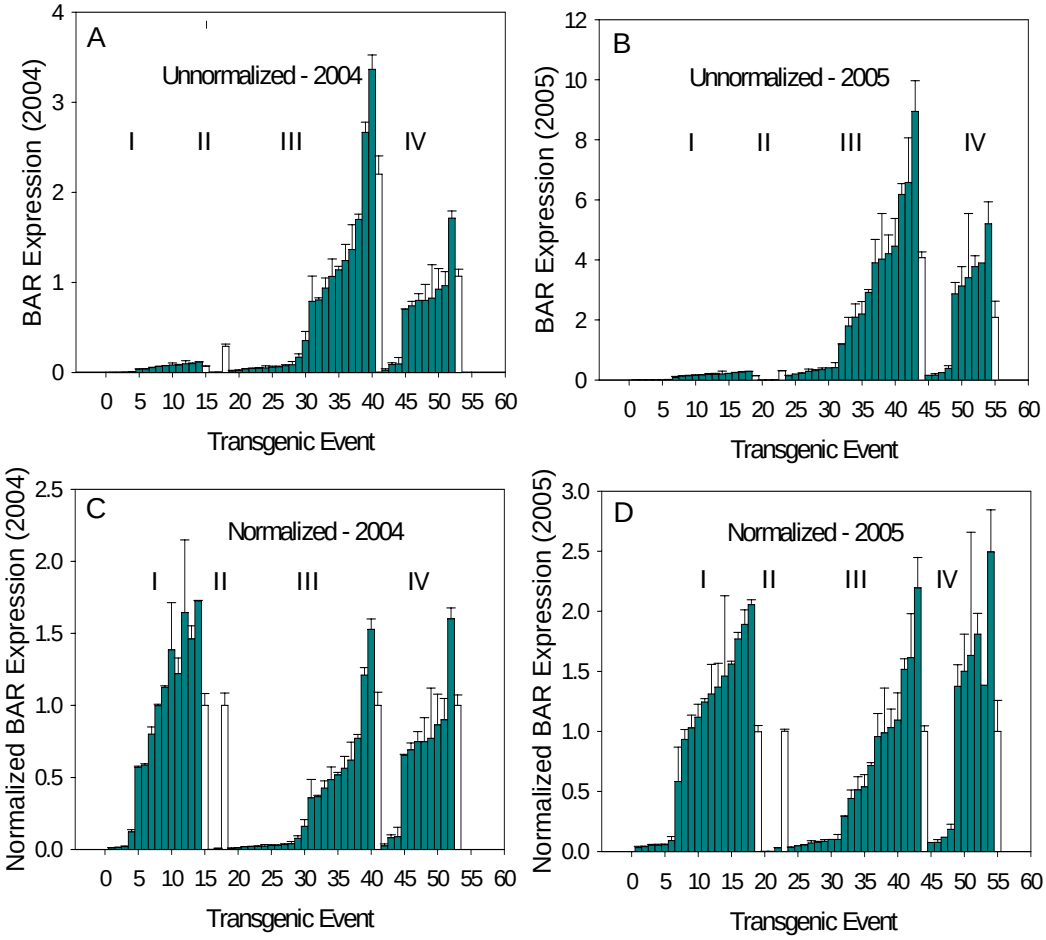


Appendix F2. 11. Southern blot of eight transgenic events (E1 ~ E8) with the *PTLF* probe. Genomic DNAs were digested with *Sca*I and *Hind*III, respectively, and hybridized with DIG-labeled 570 bp *PTLF* probe. No signal was detected for E4 due to low amount of genomic DNA loaded.

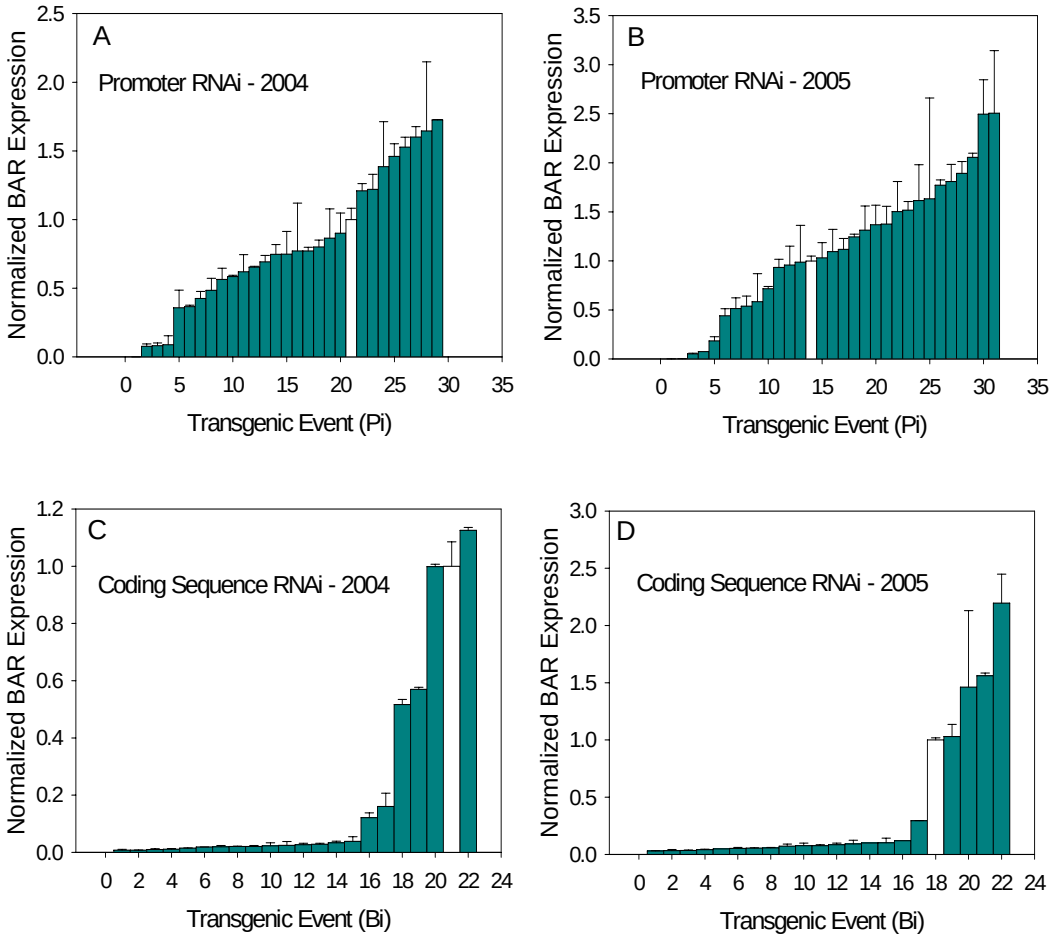


Appendix F2. 12. Southern blot of transgenic events estimated to have one or two copies by quantitative real-time PCR. All events had quantitative values (QV) of 1.4 to 1.6 from real-time PCR. Genomics DNAs were digested by *ScaI* and hybridized with DIG-labeled *GFP* probe. The relative locations of ScaI and the *GFP* probe in transgene construct pG3KGB are shown. DIG-labeled DNA molecular weight marker is shown at right. * indicates three events with inconsistent results between real-time PCR and Southern blot analysis.

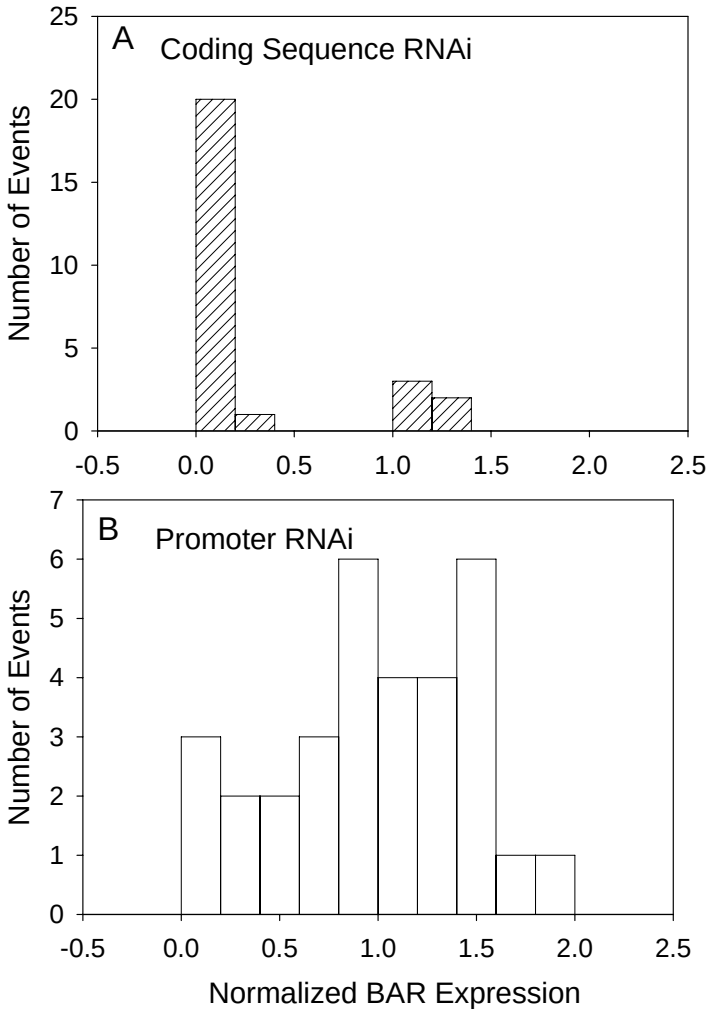
Appendix F3. 1. Schematic diagram of the T-DNA region of the plasmid vector pTTM8 used to generate the parent events. OCSt: the 3' untranslated region from the octopine synthase gene; NEO: neomycin phosphotransferase II; pNOSs: the promoter from the nopaline synthase gene; pTA29: the promoter from a tobacco anther-specific gene TA-29; NOSt: the 3' untranslated end of the nopaline synthase gene; pSSUARA-TP: the promoter from the *atS1*A ribulose-1,5-biphosphate carboxylase small subunit gene from *Arabidopsis* thaliana; G7t: the 3' untranslated fragment from the TL-DNA gene 7; RB: right border; LB: left border. T-DNA regions are not drawn to scale.



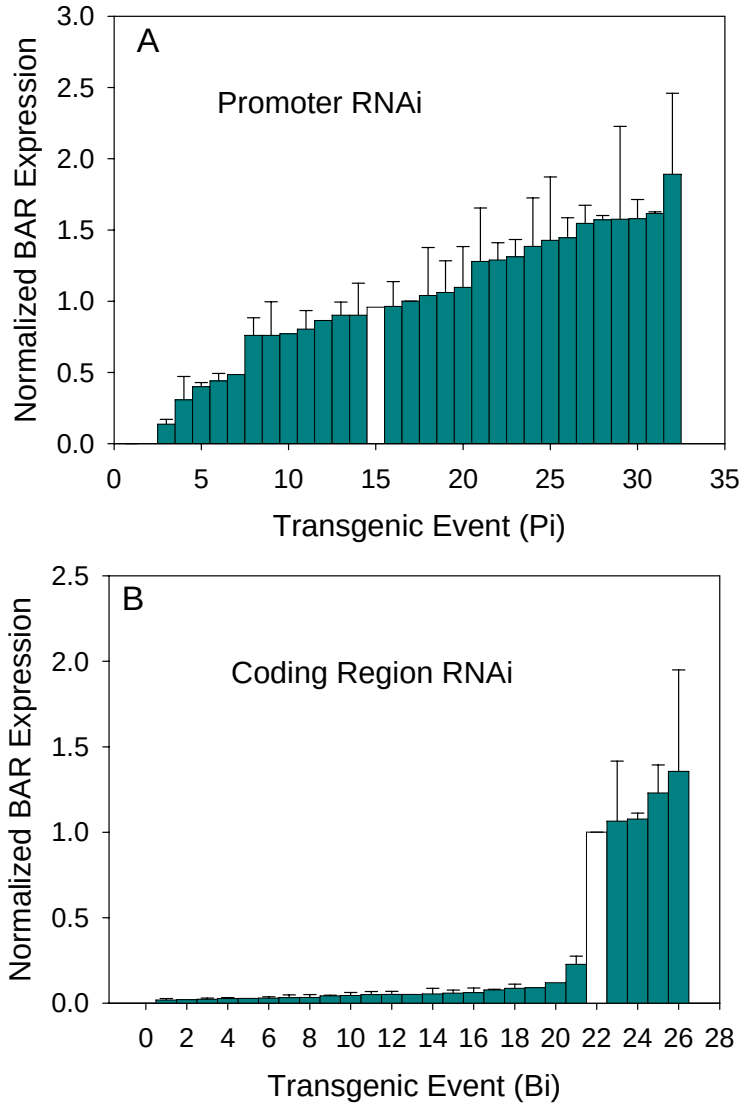
Appendix F3. 2. Normalized and non-normalized *BAR* expression levels in RNAi transgenics in two different years. Expression levels were determined by real-time RT-PCR with a *UBQ* gene as a reference gene. A, B: expression of *BAR* in 2004 and 2005. C, D: *BAR* expression normalized to expression level of parent events. White bars represent the parent events (I: 353-29, II: 353-38, III: 717-A, and IV: 717-C).



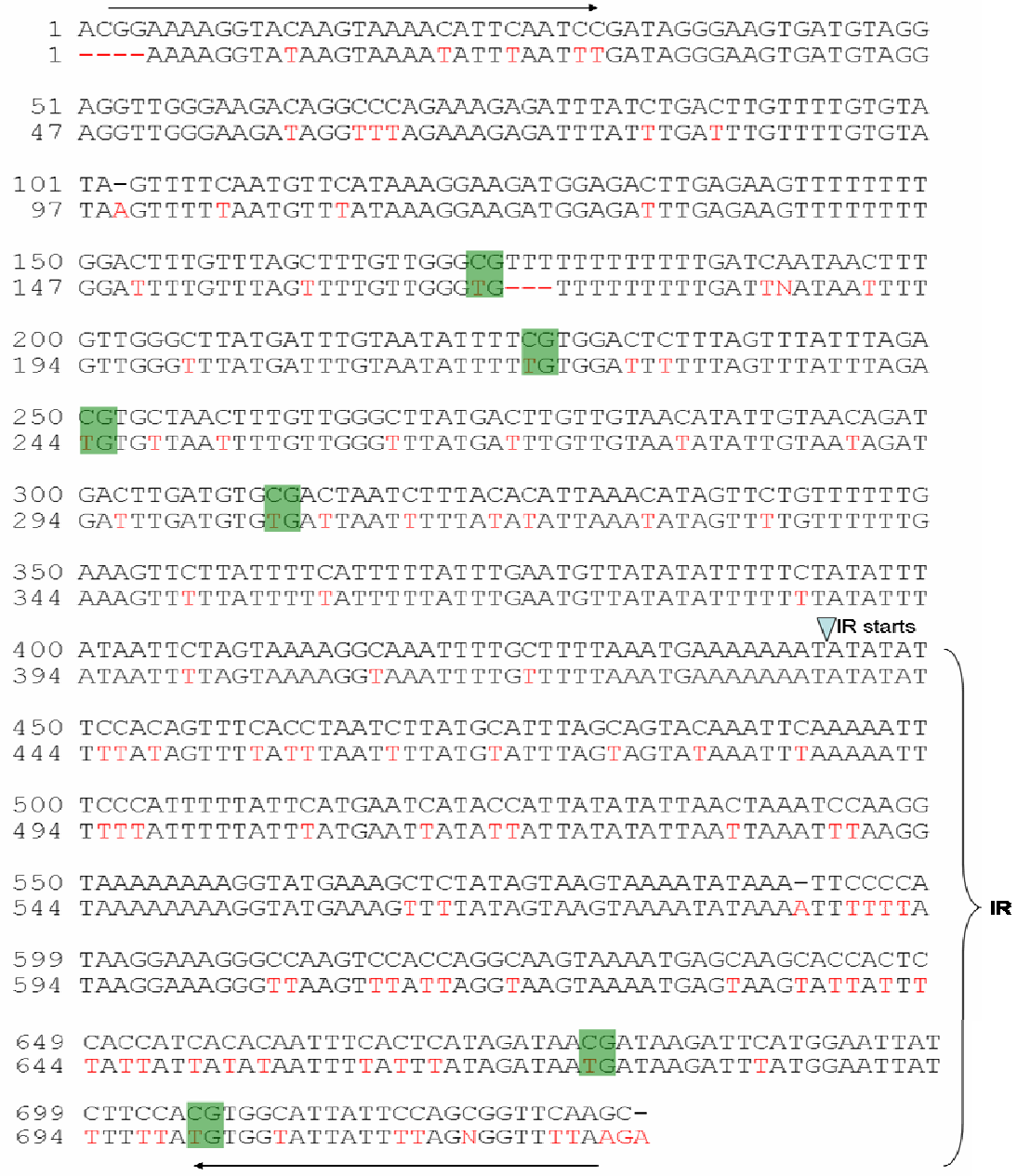
Appendix F3. 3. Normalized *BAR* expression in RNAi transgenics grouped by IR type in two different years. Expression of the *BAR* gene in RNAi transgenic events determined by real-time PCR, and normalized to their corresponding parent events. White bars represent parent event, set to a value of 1. Error bars represent one standard error of the mean.



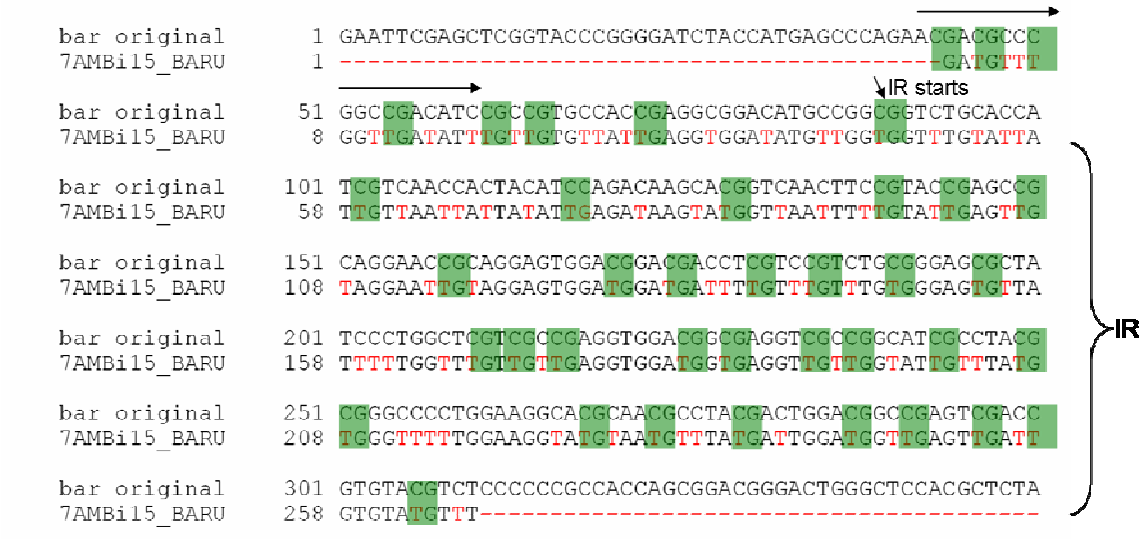
Appendix F3. 4. Distribution of normalized *BAR* expression of transgenic events transformed with A) promoter directed (Pi) or B) coding region directed (Bi) constructs. Expression was normalized to that of parent events, which was set to 1.0.



Appendix F3. 5. Normalized *BAR* expression in RNAi transgenic events averaged over two years. Expression level of the *BAR* gene was determined by real-time PCR using the *UBQ* gene as a reference, normalized to *BAR* expression levels of their corresponding parent events, and averaged over two years. White bars represent the parent events, set to a value of 1.0. Error bars represent one standard error of the mean over two years.



Appendix F3. 6. Methylation analysis of *rbc*S promoter of a PiM+ RNAi transgenic event showing no suppression (7C-MPi-16). The top strand represents untreated *rbc*S promoter genomic sequence. The bottom strand represents PCR amplified sequence from bisulfite-treated sequence. CpG sequences are shaded. Arrows indicate MSP primer sites. Sequence region included in IR is indicated.



Appendix F3. 7. Methylation analysis of *BAR* coding sequence of a BiM+ RNAi transgenic event showing no suppression (7A-MBi-15). The top strand represents untreated *BAR* genomic sequence. The bottom strand is PCR amplified sequence from bisulfite-treated genomic sequence. Unmethylated cytosines in CpG sequences were shaded. All cytosines in *BAR* genomic sequence were converted to uraciles and amplified as thymines.

ABBREVIATIONS

GE: genetically engineered

TGS: transcriptional gene silencing

PTGS: post-transcriptional gene silencing

RNAi: RNA interference

MAR: matrix attachment region from tobacco root specific gene RB7

*rbc*S: the small subunit of ribulose bisphosphate caboxylase

UBQ: ubiquitin gene

PAT: phosphinothricin acetyltransferase

Pi: RNAi transgenic constructs containing inverted repeats of the *Arabidopsis rbc*S promoter sequence (475 bp)

Bi: RNAi transgenic constructs containing inverted repeats of coding sequence of herbicide resistance *BAR* gene

PiM-: RNAi transgene construct containing inverted repeats of the *rbc*S promoter sequence without flanking MARs

PiM+: RNAi transgene construct containing inverted repeats of the *rbc*S promoter sequence with flanking MARs

BiM-: RNAi transgene construct containing inverted repeats of the BAR coding sequence without flanking MARs

BiM+: RNAi transgene construct containing inverted repeats of the BAR coding sequence with flanking MARs

IR: inverted repeats

IR-HH: inverted repeats arranged as head to head

IR-TT: inverted repeats arranged as tail to tail

DR: directed repeats

MSP: methylation-specific PCR

QV: quantitative value of copy number with real-time PCR