Tree Genetics & Genomes
DOI 10.1007/s11295-010-0317-7

ORIGINAL PAPER

Activation tagging is an effective gene tagging system in *Populus*

Victor Busov · Yordan Yordanov · Jiqing Gou · Richard Meilan · Cathleen Ma · Sharon Regan · Steven Strauss

Received: 23 March 2010 / Revised: 24 May 2010 / Accepted: 8 June 2010

Abstract Knowledge of the functional relationship between genes and organismal phenotypes in perennial plants is extremely limited. Using a population of 627 independent events, we assessed the feasibility of activation tagging as a forward genetics tool for *Populus*. Mutant identification after 2 years of field testing was nearly sevenfold (6.5%) higher than in greenhouse studies that employed *Arabidopsis* and identical transformation vectors. Approximately two thirds of all mutant phenotypes were not seen in vitro and in the greenhouse; they were discovered only after the second year of field assessment. The trees' large size (5-10 m in height), perennial growth, and interactions with the natural environment are factors that are thought to have contributed to the high rate of observable phenotypes in the field. The mutant phenotypes affected a variety of morphological and physiological traits, including leaf size and morphology, crown architecture, stature, vegetative dormancy, and tropic responses. Characterization of the insertion in more than 100 events with and without mutant phenotypes showed that tags predominantly (70%) inserted in a 13-Kbp region up- and downstream of the genes' coding regions with approximately even distribution among the 19 chromosomes. Transcriptional activation was observed in many proximal genes studied. Successful phenotype recapitulation was observed in 10 of 12 retransformed genes tested, indicating true tagging and a functional relationship between the genes and observed phenotypes for most activation lines. Our studies indicate that in addition to associating mapping and QTL approaches, activation tagging can be used successfully as an effective forward gene discovery tool in *Populus*.

Keywords Mutagenesis · Gene function · Adaptation · Woody perennials

Communicated by W. Boerjan

Electronic supplementary material The online version of this article (doi:10.1007/s11295-010-0317-7) contains supplementary material, which is available to authorized users.

V. Busov (✉) · Y. Yordanov · J. Gou
School of Forest Resources and Environmental Science,
Michigan Technological University,
Houghton, MI 49931-1295, USA
e-mail: vbusov@mtu.edu

R. Meilan
Department of Forestry and Natural Resources, Purdue University,
West Lafayette, IN 47907-2061, USA

C. Ma · S. Strauss
Department of Forest Ecosystems and Society,
Oregon State University,
Corvallis, OR 97331-5752, USA

S. Regan
Department of Biology, Queen's University,
Kingston, ON K7L 3N6, Canada

Introduction

Insertional mutagenesis using transposons or T-DNA has become an extremely valuable research tool for model plant systems (Alonso et al. 2003; Jeong et al. 2002; Krysan et al. 1999; Martienssen 1998; Raizada et al. 2001). Activation tagging is a T-DNA-based approach in which a series of strong enhancers are positioned next to the right- or left-hand T-DNA border. Insertion of the vector near a gene causes up-regulation and dominant gain-of-function mutations. Several attributes of activation tagging make it an attractive tool for functional genomics studies and leading to development of large collections predominantly in *Arabidopsis* (Ichikawa et al. 2003) and

Published online: 08 July 2010

rice (Jeong et al. 2006). Activation tagging produces dominant or semi-dominant mutations that can be identified in primary transformants. This is important for plant species, such as most perennials, that are predominantly out-crossing. Selfing or mating with close relatives causes severe inbreeding depression and the generation of a great deal of phenotypic variability, making the recognition and recovery of homozygous plants (needed to reveal the recessive, loss-of-function phenotypes) difficult. In addition, juvenility can last from several years to decades in many perennial plants, most prominently trees, making inbreeding impractical.

Activation tagging has a number of advantages over other tagging systems. Activation vectors show insertion preference for gene-rich sectors, insertions are uniformly distributed among chromosomes, and phenotypic alterations are observed more frequently than are knockout phenotypes. Ichikawa et al. (2003) found that the number of activation vector insertions on each chromosome is nearly the same, ranging from 8.7 to 10.0/Mb. Because there is a bias for T-DNA integration outside transcribed regions (Sessions et al. 2002), activation tagging (which can act over several Kbp) should be more efficient at generating mutations than T-DNA insertions. In a population of 1,172 mutant events transformed with an activation tagging vector, mutant phenotypes were caused by insertion outside the gene in 855 cases, whereas 317 mutations were caused by insertion within an open reading frame (Ichikawa et al. 2003), an approximately threefold higher frequency.

Many types of genes are present in multiple copies that have partially redundant functions (Byrne et al. 2002; Hiratsu et al. 2003). This makes functional dissection via loss-of-function mutations problematic. Because activation tagging generates gain-of-function mutations, it provides an alternative for functional characterization of gene families. For example, using activation tagging, Nakazawa et al. (2003) were able to functionally characterize three new members of the large *ASSYMETRIC LEAVES 2* (*AS2*)/*LATERAL ORGAN BOUNDARIES* (*LOB*) gene family (42 members in *Arabidopsis*). Activation tagging was used for the identification of gene family members that are so divergent in sequence that homology searches fail to predict orthology. Schomburg et al. (2002) isolated two novel members of the *GA 2-oxidase* gene family with poor sequence relatedness to the other family members (Hedden and Phillips 2000). Activation tagging was also successfully used to discover non-protein coding genes like microRNAs (Palatnik et al. 2003).

The only prerequisite for using activation tagging in any species is the need for at least one genotype to be efficiently transformed. Hence, activation tagging has been growing in popularity in diverse non-model plant species like resurrection plant (*Craterostigma plantagineum*) (Furini et al. 1997), *Petunia hybrida* (Zubko et al. 2002), tomato (Mathews et al. 2003), barley (Ayliffe et al. 2007) and poplar (genus *Populus*; Busov et al. 2003; Harrison et al. 2007).

Advances in genomics technology have allowed the sequencing and annotation of numerous plant genomes. *Populus* was the third plant species for which a genome sequence was produced (Tuskan et al. 2006). However, for all plant genomes scientists are faced with the enormous task of functional annotation. Genes with functions that are important to perennial plant biology or biotechnology are often difficult to recognize based on inferences from annual plants or biochemistry alone. A system for mutagenesis and gene tagging would enable the discovery of new genes, or genes whose functions in perennial plants were not fully appreciated from studies of annual models (Fladung et al. 2004).

In this paper, we describe the results of more than a decade of study of a pilot activation tagging population produced in poplar trees (genus *Populus*). We report that activation tagging is an effective method for mutagenesis in poplar.

Materials and methods

Generation and screening of poplar activation tagging population

Binary vectors pSKI015 and pSKI074 were used for transformation of a hybrid aspen clone INRA 717-1B4 via an *Agrobacterium*-mediated procedure (Han et al., 2000). Before being transferred to a greenhouse or the field, all putative transformants were PCR-verified for the presence of the T-DNA insert using the following primer pair: pSK-F: 5′-CTCACTATAGGGCGAATTG-3′ and pSK-R: 5′-AACACTGATAGTTTCGGATC-3′.

A field study was established during the fall of 2002 in the vicinity of Corvallis, Oregon. Four ramets of each event (i.e., independent gene insertion) were planted in pairs; each pair was randomly assigned to a position in the field. All trees were planted at a spacing of 2.7×2.0 m (between rows×within rows). A border of non-experimental trees, derived from the same clone, was planted at the same spacing around the periphery of the plantation. Trees were irrigated and weeds intensively controlled throughout the first growing season.

Four phenotypic screens were performed; three during the second and one at the beginning of the third growing season. An event was classified as a putative mutant if all four ramets showed the same phenotypic abnormality. Some of the events used in this study (e.g., IDs starting with ‘E’ and ‘D’) were generated in a different study using

the same poplar genotype, activation tagging vectors, and transformation method (Harrison et al. 2007).

TAIL-PCR and plasmid rescue

Genomic DNA was extracted using Qiagen (Valencia, CA, USA) a DNAeasy kit following the manufacturer's instructions. Plasmid rescue was performed as previously described (Busov et al. 2003). TAIL-PCR was performed using vector-specific primers (LBr1, 5′-AAGCCCCCATTTGGACGT GAATGTAGACAC-3′; LBr2, 5′-TTGCTTTCGCCTA TAAATACGACGGATCG-3′, LBr3, 5′-TAACGCTGCG GACATCTAC-3′) and arbitrary degenerate (AD) primers (AD21, 5′-NGTCGASWGANAW GAA-3′; AD22, 5′-NGTCGASWGANAWGTT-3′; AD23, 5′-NGTCGASWGA NAWGAC-3′; AD24, 5′-NGTCGASWGANAWCAA-3′; AD25, 5′-NGTCGASW GANAWCTT-3′; AD3, 5′-WCAGNTGWTNGTNCTG-3′; AD4, 5′-NGTA WAASGTNTSCAA-3′). PCR reaction and conditions were essentially same as those described by Liu et al. (1995). PCR amplification products were separated by electrophoresis and purified using a Gel Extraction Kit (Qiagen; Valencia, CA, USA); and sequenced using the LBr3 primers.

Sequence analysis

Recovered genomic sequences flanking the activation tagging insertion sites were used to perform BLAST searches of the poplar genome. Approximately 10 Kb up- and downstream of the insertion sites were analyzed for presence of annotated gene models using the *Populus* genome browser (http://genome.jgi-psf.org/Poptr1/Poptr1.home.html). Sequence homology searches and analyses were performed using the NCBI BLAST server. Sequence alignments were carried out by the CLUSTAL W method and using the EMBL server (www.ebi.ac.uk/clustalw/).

RT-PCR analyses

RNA was extracted as previously described (Busov et al. 2003) and DNAse treated using DNA-free TM kit (Ambion; Foster City, CA, USA). cDNA was synthesized using SuperScriptII (Invitrogen; Carlsbad, CA, USA) and 1-5 μg of total RNA. Reverse transcription polymerase chain reaction (RT-PCR) was performed using gene-specific primers. Equal amounts of cDNA were used in each reaction and ubiquitin-like gene expression was used as a loading control as previously described (Busov et al. 2003). Gel images were acquired by a UVP Gel Doc-It (UVP; Upland, CA, USA) documentation system and quantified using ImageJ (http://rsb.info.nih.gov/ij/index.html).

Results

Generating the tagged population

To generate a population of activation tagged events, we used binary vector pSKI074 (kanamycin selection) and pSKI015 (Basta selection) that were previously tested in *Arabidopsis* (Weigel et al. 2000). Both vectors contain a tetramer of the cauliflower mosaic virus (CaMV) 35S enhancer near the right-hand T-DNA border. The transformation efficiency obtained using pSKI074 was approximately twice that of pSKI015 ($P<0.001$; Table 1). A total of 627 independent events were produced. In addition, we used a subset of events generated in a different study (Harrison et al. 2007).

Rate of mutant phenotype recovery

Transformed plants were transplanted into a 2.5-acre field study in west-central Oregon. Two pairs of each event (a total of four ramets) were planted in a completely randomized design. All trees were phenotypically evaluated repeatedly during the 2003 growing season. In the spring of 2004, we rescreened the trees and identified additional mutants with an early bud-break phenology (Fig. 1).

In total, 41 events showed scorable phenotypes that were consistently identified in all four field-grown ramets. The frequency of mutant phenotypes was 6.5% or approximately 6.5-fold higher than what have been observed in *Arabidopsis* with the same vectors (Weigel et al. 2000). We observed a slightly higher mutant phenotype rate with pSKI015 than with pSKI074 (Table 1), but the difference was not statistically significant. A variety of traits were affected, including leaf size and morphology, branch angle, height, wood formation, entry to and release from dormancy, and crown architecture (Table 2, Figs. 2, 3).

Seasonal and perennial field screens aided mutant isolation

The rate of putative mutant identification prior to field exposure was 1.6% (10 of 627 transgenic events), similar to what has been obtained with *Arabidopsis* (Weigel et al. 2000). However, after a single growing season in the field, where plants grew rapidly and interacted with a changing environment, the rate of mutant phenotype identification increased dramatically (Fig. 1). We isolated putative mutants with altered morphological, phenological, and physiological traits (Figs. 2, 3). We have identified numerous alterations in leaf size, serration, curvature, texture, venation (Fig. 2). For example, one event that we called 'Big Leaf' displayed extreme leaf size, almost fourfold larger than a wild-type (WT) leaf (Fig. 2e). Another event was affected with respect to leaf curvature

Table 1 Transformation and frequency of mutant recovery with two activation-tagging vectors

Vector	Selection	Transgenic events[a]	Transformation efficiency	Phenotypic events	Mutant (%)
pSKI074	Kanamycin	445	26.63 (3.49)*	28	6.3
pSKI015	Basta	182	11.33 (1.83)	13	7.1
Total		627	19.13	41	6.5

Transformation efficiency estimates are based on an average of five (pSKI074) and six (pSKI015) independent co-cultivation experiments

[a] All transgenic plants were PCR-verified for the presence of the activation-tagging vector

*$P<0.01$, significant as determined by Student's t test

and indentation (Fig. 2c). Six events had altered timing of bud break (Figs. 3a, b) and/or leaf senescence (Fig. 3c), key aspects of adaptation of perennial plants to seasonal changes. Other trait variations identified included plagiotropic growth (Fig. 3d) with the bend always south oriented. Several events were affected with respect to branch size, number, and crown length and structure (Fig. 3e, f). Many events were affected with respect to their stature (Fig. 3g-i). Interestingly, abnormal height growth of some events (e.g., Fig. 3g, h) was only detected after 2 years of field growth, but was not obvious under in vitro and greenhouse conditions.

Genome insertion of the tag

We used thermal asymmetric interlaced PCR (TAIL-PCR) and plasmid rescue to recover genomic sequences flanking the insertion sites. TAIL-PCR is a higher throughput method for recovering of adjacent DNA; hence, most insertions were positioned using this method. We optimized the originally described method for poplar (see "Materials and methods" section). The majority (93%) of flanking sequences was amplified by four primers AD21, AD23, AD24, and AD3 (Table S1). We had the highest (32%) success rate with AD21.

To date, we have recovered a total of 136 insertions in 109 events (both with and without a phenotype). Recovered sequences were used to position the insertions in the poplar genome (Table S1). More than half of the sequences (56%) produced hits in the genome with E (expected probability of homology by chance)=0 and significant part (91%) with $E>0.00001$ (Table S1). Therefore sequence homology between the sequenced genome and the INRA 717-1B4 (*Populus tremula* x *P. alba*) genotype is sufficient to allow efficient positioning of the tag in the heterologous genome. Although 24 sequences mapped to unassembled scaffolds, the scaffold size in all cases was of sufficient length to allow identification of the putative candidate gene neighboring the insertion site. Using the *Populus* genome browser, we have identified the distance and orientation of the closest annotated gene. We found the majority (70.1%) of the inserts to be in the proximal 10 Kbp of the 5′/3′ intergenic regions of the genes, while only 17 (13.4%) were in introns, and 20 (15.7%) in exons (Fig. 4a). The average distance from the translation start site of the closest gene to the 35S enhancers was 5.0 Kbp. To date, we have recovered insertions in all 19 *Populus* linkage groups, suggesting an approximate uniform distribution of the tag in the genome (Fig. 4b).

Transcriptional activation

To determine if there was transcriptional activation, we selected six different candidate genes on which to perform expression analysis. In all cases, we detected strong activation of the candidate gene (e.g., Fig. 5a), whereas the ubiquitin control showed approximately equal levels of expression in

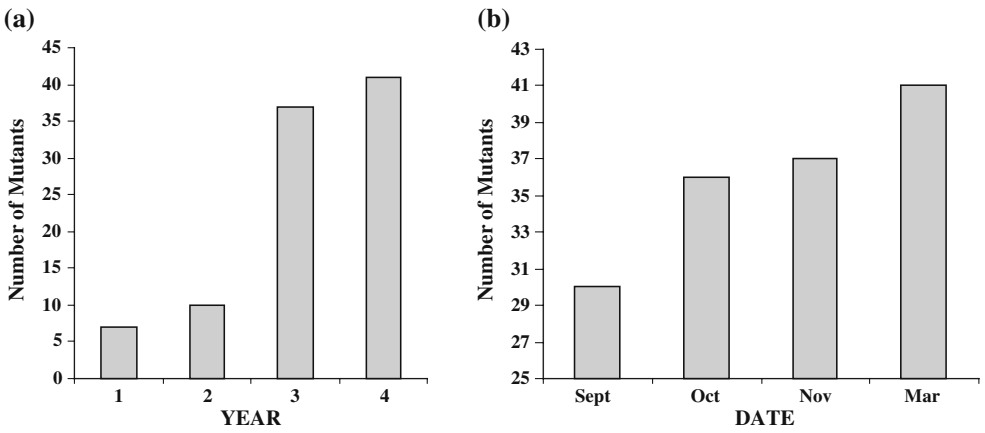

Fig. 1 Rate of mutant phenotype identification: **a** during laboratory and greenhouse (years 1-2) and field (years 3-4) phases of study. **b** During different seasons within a single year in the field

Table 2 Mutant description and identification

Event ID	Trait(s) affected	Identification
B120	Reduced stature. Leaves smaller with different shape. Advanced bud set	F
B144	Reduced stature	I/G
B147	Fine branches, small leaves with less dentation	F
B154	Smaller and less dentated leaves. Smaller and finer branches. Mites danmage	F
B164	Early bud break in field 2004	F
B17	Reduced stature	I/G
B18	Proliferation of many branches	I/G
B204	Reduced stature. Fewer syleptic branches, smaller leaves, less indented	F
B245	Reduced stature, wavy stem, sylleptic branches	F
B50	Forking and extensive yellow spots like mite damage	F
B67	Reduced stature and leaf size	F
B73	No apical dominance, forking, small leaves	F
B89	Reduced stature and leaf size. No sylleptic branches	F
K1011	Many fine syleptic branches. Leaves smaller birchy. Avanaced senescence	F
K1055	Advanced senescence. Reduced stature and leaf size. Tiny sylepitic branches	I/G
K1082	Extensive frost damage.	F
K123	Early bud break in field 2004	F
K133	Early bud break in field 2004	F
K144	Reduced overall growth	I/G
K147	Fine branches, small leaves with less dentation	F
K155	Narrow leaves, lethal under field conditions	I/G
K163	No apical dominance. Leaves smaller than WT and leathery	F
K190	Curvy leaves, early bud break in the field 2004	I/G
K244	Early senescence	F
K263	Reduced stature and leaf size. Advanced leaf senescence	F
K299	Reduced stature and leaf size	F
K345	Abundance of syleptic branches appearing almost to the top of the plant. Leaves smaller than WT	F
K352	Numerous, fine syleptic branches in the lower crown	F
K368	Round leaves. Fewer syleptic branches than WT	F
K390	Plagiotropic	F
K584	Reduced stature and leaf size.	F
K638	Small/narrow leaves	I/G
K645	Big leaves, early bud break in the in the field 2004	F
K694	Green leaves after several frosts, no apical dominance	I/G
K70	Reduced apical dominance. Leaves wedged at the bottom and margins more serrated than WT	F
K714	No apical dominance. Leaves smaller than WT and round	F
K771	Sparse crown, fine syleptic branches, premature yellowing	F
K850	Smaller stature	F
K895	Reduced stature. Leaves slightly lobed with curvy margin	F
K902	Reduced stature	F
K916	Deep margin and curved leaves	I/G

F field, *I/G* in vitro/greenhouse

both mutant and WT trees. In one case, we performed RT-PCR on both genes flanking the insertion site (Fig. 5b), and expression of the proximal gene was high while the distal gene was unaffected (Fig. 5b). We found transcription activation in genes as far as ~13 Kbp from the tag.

High recapitulation success indicates true gene tagging

To assess the rate of true tagging, we attempted to recapitulate 12 events showing diversity of mutant phenotypes by fusing the cDNA protein-coding region of each

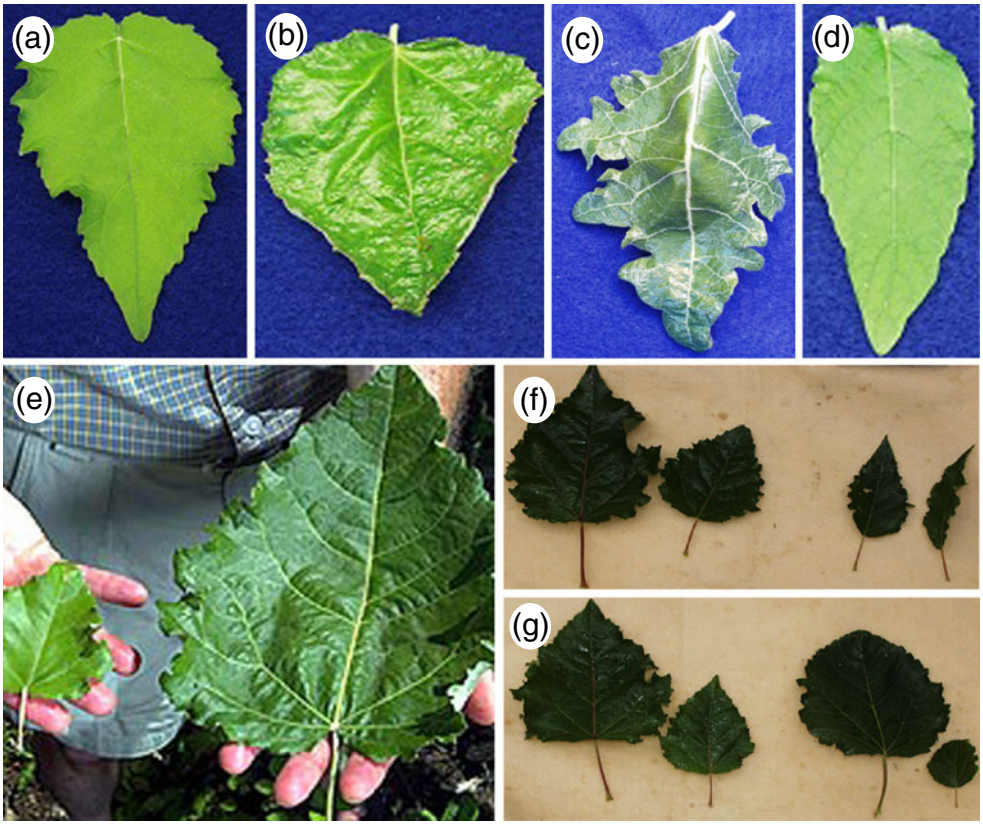

Fig. 2 Mutant phenotypes affecting leaf shape. **a** WT leaf; **b-d** leaf with altered shape and curvature identified during greenhouse growth; **e-g** leaf with altered size and shape identified under field conditions. Leaf in **e** to the left is WT and the leaf to the right from a plant displaying a mutant phenotype. Leaves in **f-g** are arranged in pairs where WT leaf's pair is to the left and the leaf pair from the mutant phenotype plants is to the right. For each pair of leaves, the leaf to the left in each pair is taken from the main stem and the leaf to the right is from the syleptic branch. All leaves shown are representative of at least four ramets of the same event

putatively tagged gene to the strong CaMV 35S promoter and re-transforming the same genotype, 717-1B4, with the construct. We observed clear recapitulation of the mutant phenotype in 10 of 12 cases, corresponding to a success rate of about 83% (Table 3). We successfully recapitulated genes modulating diverse characteristics in trees like tropic response, leaf size and shape, phenology, and stature. Full reports characterizing these genes in detail will be published elsewhere.

Discussion

The perennial habit can expand mutant detection

Perennial plants have many developmental differences from annuals, including vegetative dormancy, delayed onset of flowering, extended periods of secondary (woody) growth, and gradual vegetative maturation. Activation tagging and other forward-genetic approaches in perennial plants therefore have the potential to uncover many new types or functions of regulatory genes. We identified phenotypes that are difficult to observe in annual plants, including modified phenology, shoot orientation, and crown architecture. Indeed, the perennial growth habit of trees provide a number of advantages for mutant detection. First, trees are bigger and, hence, subtle mutant phenotypes can be more easily discerned and measured than in small annual plants. Under our experimental conditions, the fastest growing trees reached 4 m in height in a single year. Second, because of their perennial growth habit, differences in growth rate of trees are compounded over time, further helping to reveal subtle differences. Third, woody tissues record the process of development for all non-deciduous tissues, providing a 'developmental memory' that is 'saved' each year. This enables developmental differences to be detected with infrequent observations. Fourth, in annual plants, growth is often highly dependent on initial cultural conditions. Thus, it is difficult to separate the effect of poor in vitro and greenhouse conditions from subsequent growth (Ichikawa et al. 2003). The perennial growth of trees allows full recovery after the initial selection and nursery phase, so that genetic potential and reaction to newly imposed environments are fully expressed. Fifth, trees contain multiple repeating developmental units (e.g., branches), and species such as poplars can be easily cloned, enabling many observations on the same individual or genetically identical propagules. Sixth, the diversity and nature of stresses that occur under field conditions cannot be fully duplicated in the laboratory. By employing a woody perennial, we can study cumulative responses to numerous and repeated seasonal and annual environmental variations.

Fig. 3 Mutant phenotypes affecting tree form and phenology. **a**, **b** Tree from two different transgenic events showing early bud break. Photos were taken on March 30, 2004. c Aerial photo of event showing late senescence in the fall (November 2003). This event also showed early bud-break. d Events showing plagiotropic growth, always leaning southward. **e** Event with extreme proliferation of syleptic branching covering the main stem almost to the top. A WT plant growing on a neighboring row is shown to the right with the main stem completely devoid of syleptic branching. **f** Event with a conspicuous absence of any syleptic branching to the bottom of the plant. *Arrows* point to the plants with the mutant phenotype; **g-i** plants from different transgenic events showing a severe reduction of height. All plants shown are representative of at least four ramets of the same event

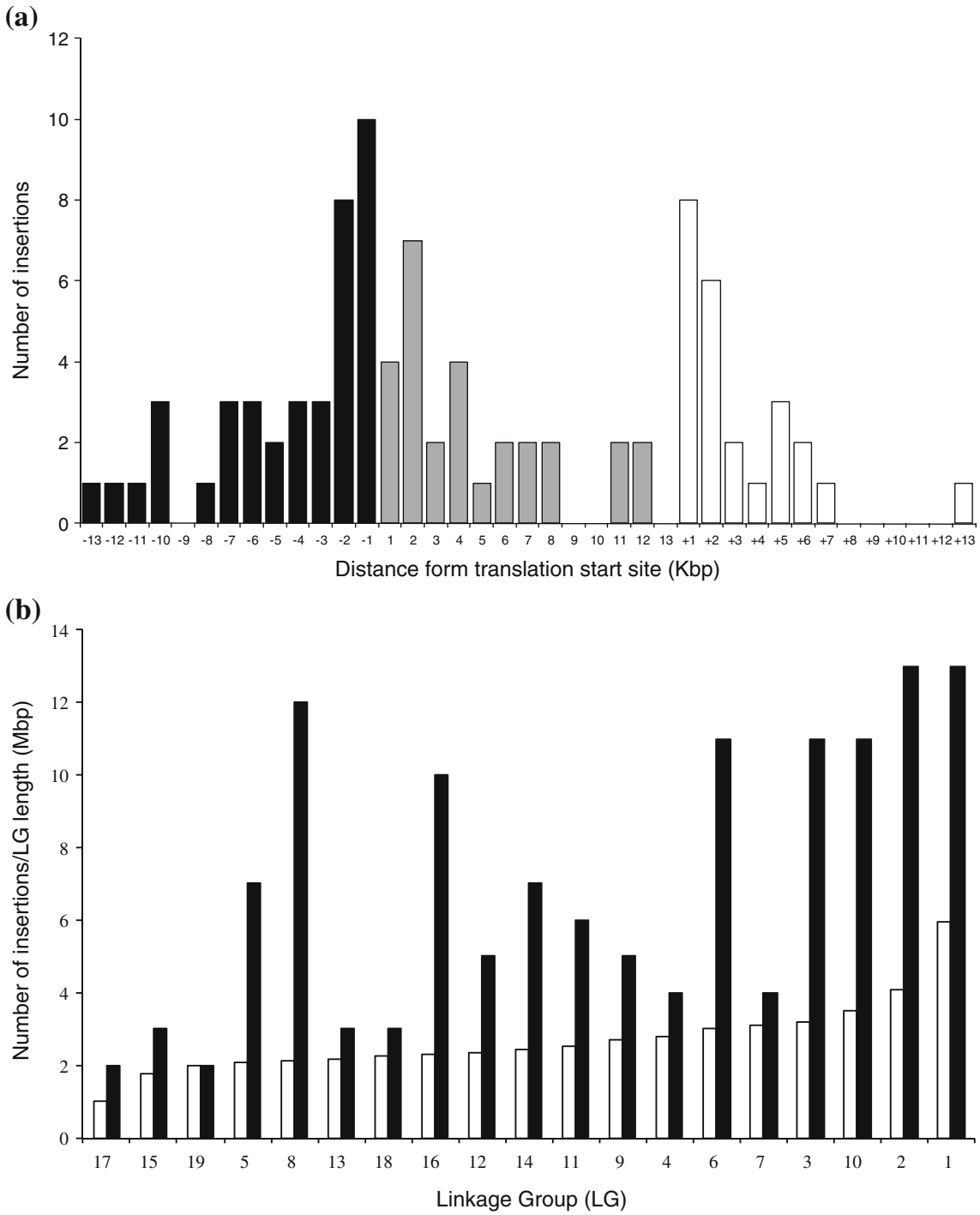

Fig. 4 Insertion of the tag in the *Populus* genome. **a** Distribution of insertions relative to gene-coding regions most proximal to the tag. The 5′ region is in black, coding region (exons+introns) in *gray*, and 3′ region in *white*. **b** Insertion number is by linkage group. *Black bars* represent number of insertions and the white shows chromosome size in Mbp. Insertions which map on unassembled scaffolds are not depicted in **b**

Insertion of the tag in the poplar genome

Characterization of 136 insertion sites indicated that, in most cases, the tags were inserted in proximity of gene-coding regions. Approximately 70% of all insertions were within 13 Kbp up- and downstream of the translation start site. Although the average distance of the insertion was 5 Kbp from the start codon of the closest gene, about one

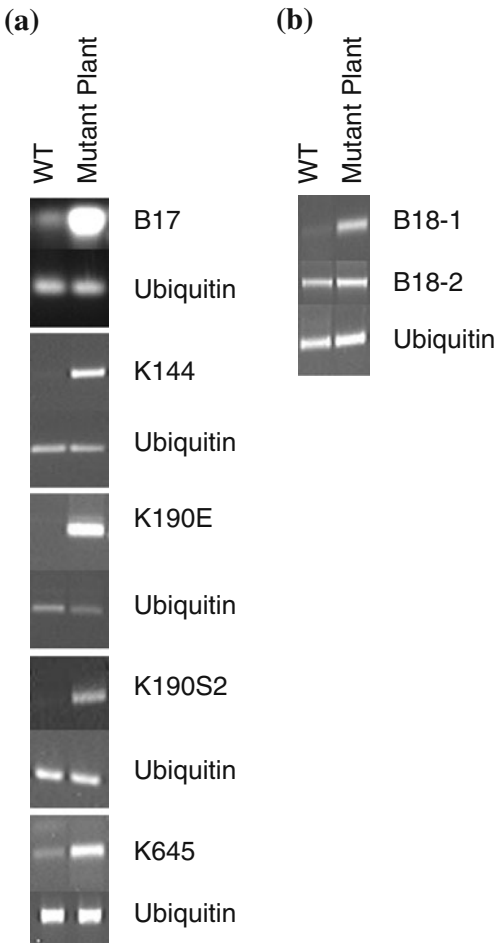

Fig. 5 Transcriptional activation of six candidate genes. **a** Five genes with high transcription activation positioned close to the enhancer elements. **b** Transcriptional activation of a gene (B18-1) closest to the enhancer trap and lacking activation in the next closest gene (B18-2) (~10 Kbp) to the enhancer elements

third of all insertions were within 2.0 kb. We also observed clear transcriptional activation of genes proximal to the insertion site. In *Arabidopsis*, the largest distance over which activation was reported is 3.6 Kbp (Weigel et al. 2000). In poplar, we found that activation may occur from enhancers situated far away as 13 Kbp. In summary, high rates of insertion near genes and activation over large genetic distances make this an efficient method for tagging genes.

Multiple insertions may make it difficult to associate genes with phenotypes. We have recovered sequences that map to different locations in the genome. In one case, we demonstrated activation of two different genes flanking the enhancers at two different genomic loci. We consider these sequences to have been recovered from separate insertion events. Based on the 136 sequences being recovered from 109 independent events, we estimate the average number of insertions per event to be 1.25. This value probably underestimates the real number of insertions per event because TAIL-PCR or plasmid rescue may not have recovered sequence adjacent to every insertion. Our estimate of 1.25 insertions per event is slightly lower than the average of 1.6 insertions per event found by Groover et al. (2004) via Southern blots in a study of enhancer and gene traps employing the same poplar genotype and the same transformation protocol. We believe that multiple insertions do not present significant problems for this and other insertional mutagenesis methods in poplar, for several reasons. First, the relative proportion of multiple insertions using *Agrobacterium*-mediated transformation is low; in poplar and other species, it results in average of one to three transgene loci, consistent with our results and those of Groover et al. (2004). Second, even in the more complicated case, selection of the genes and recapitulation of the phenotype is a relatively minor undertaking (e.g., compared to dozens of genes that might need to be screened in a QTL interval). There is much interest in saturating a genome

Table 3 Recapitulation in 12 activation-tagged events that showed distinct phenotypes

Event	Phenotype	Recapitulation
D17-72[a]	Increased adventitious rooting	Yes
K133[a]	Early bud break in spring	Yes
B120	Decreased stature	Yes
B17	Dwarfism	Yes
E7-11[a]	Change in bark texture	Yes
K190[a]	Early bud break in spring	Yes
K368	Round leaf shape and cup-like curvature	Yes
K390[a]	Plagiotropic south-oriented stem growth	Yes
K638	Reduced stature	No
K645[a]	Increased leaf size and early bud break	Yes
K694	Increased adventitious rooting	Yes
K70	Reduced apical dominance	No

[a] Events for which gene characterization will be published in separate reports

with insertions near genes, a goal for which multiple insertions would be an advantage—by reducing the size of the transgenic population required. Once a phenotype of interest is produced, it is generally worth the effort needed to test a few candidate genes. Third, simple, high-throughput PCR assays could be used to allow rapid screening and early culling of very complex insertion events.

High rate of successful recapitulation

In addition to many other experimental advantages, poplars can be transformed with high efficiency (Busov et al. 2005; Cseke et al. 2007; Filichkin et al. 2006). Therefore, verifying functionality of genes identified via activation tagging is feasible, allowing strong gene-phenotype inferences to be made in reasonable time frames. We have achieved a high recapitulation success in this study: 10 of 12 (>83%) of tested genes. Recapitulation is the ultimate evidence for true tagging, and our data strongly suggests true tagging occurs most of the time. The true recapitulation success may be even higher because, in one case, we observed a phenotype in the original event only when it was grown under field conditions, and have not yet been able to test recapitulation phenotypes in the field. Thus, in some cases, unique interactions between the affected gene and the environment may not have been replicated under greenhouse conditions. A much larger number of retransformations and phenotypic assessments are needed to make general conclusions about the efficiency of tagging. However, our data strongly suggests that in most cases phenotypes result from tagging.

Conclusions

We have shown that a dominant gene-tagging approach, combined with phenotypic analysis in a field environment, appears to greatly enhance the ability to identify loci controlling development of perennial plants. We demonstrate the power of this approach in poplar, a tree taxon with rich genomic resources (Jansson and Douglas 2007), and which is ecologically dominant in several types of terrestrial ecosystems (Brunner et al. 2004). However, this approach could be taken in any plant species for which there is as least one highly transformable genotype. Moreover, the growing capacity for low-cost genome-scale sequencing will greatly facilitate gene identification and confirmation (Emrich et al. 2007). Mutants and gene tagging no longer need to be restricted to model plant species; activation tagging allows effective forward genetics in virtually any species of interest. Because of the high efficiency of the method which we report here, activation tagging populations of various sizes can be successfully used for discovery of genes regulating important woody perennial traits and complement or enhance the power of other forward approaches like associating and QTL mapping.

Acknowledgments This work was supported in part by grants from the US Department of Energy Poplar Genome Based Research for Carbon Sequestration in Terrestrial Ecosystems Program (DE-FG02-06ER64185, DE-FG02-05ER64113); the Consortium for Plant Biotechnology Research, Inc. (GO12026-203A); USDA-CSREES, Biotechnology Risk Assessment Research Grants Program (2004-35300-14687); and Genome Canada. We thank Chamini Illangasinghe and Rama Joshi for production of binary constructs and RT-PCR analyses.

References

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, guilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301:653–657

Ayliffe MA, Pallotta M, Langridge P, Pryor AJ (2007) A barley activation tagging system. Plant Mol Biol 64:329–347

Brunner AM, Busov VB, Strauss SH (2004) The poplar genome sequence: functional genomics in a keystone plant species. Trends Plant Sci 9:49–56

Busov VB, Meilan R, Pearce DW, Ma C, Rood SB, Strauss SH (2003) Activation tagging of a dominant gibberellin catabolism gene (*GA 2-oxidase*) from poplar that regulates tree stature. Plant Physiol 132:1283–1291

Busov VB, Brunner AM, Meilan R, Filichkin S, Ganio L, Gandhi S, Strauss SH (2005) Genetic transformation: a powerful tool for dissection of adaptive traits in trees. New Phytol 167:9–18

Byrne ME, Simorowski J, Martienssen RA (2002) ASYMMETRIC LEAVES1 reveals knox gene redundancy in *Arabidopsis*. Development 129:1957–1965

Cseke LJ, Cseke SB, Podila GK (2007) High efficiency poplar transformation. Plant Cell Rep 26:1529–1538

Emrich SJ, Barbazuk WB, Li L, Schnable PS (2007) Gene discovery and annotation using LCM-454 transcriptome sequencing. Genome Res 17:69–73

Filichkin SA, Meilan R, Busov VB, Ma C, Brunner AM, Strauss SH (2006) Alcohol-inducible gene expression in transgenic *Populus*. Plant Cell Rep 25:660–667

Fladung M, Deutsch F, Honicka H, Kumar S (2004) T-DNA and transposon tagging in Aspen. Plant Biol (Stuttg) 6:5–11

Furini A, Koncz C, Salamini F, Bartels D (1997) High level transcription of a member of a repeated gene family confers dehydration tolerance to callus tissue of *Craterostigma plantagineum*. EMBO J 16:3599–3608

Groover A, Fontana JR, Dupper G, Ma CP, Martienssen R, Strauss S, Meilan R (2004) Gene and enhancer trap tagging of vascular-expressed genes in poplar trees. Plant Physiol 134:1742–1751

Han KH, Meilan R, Ma C, Strauss SH (2000) An *Agrobacterium tumefaciens* transformation protocol effective on a variety of cottonwood hybrids (genus *Populus*). Plant Cell Rep 19:315–320

Harrison EJ, Bush M, Plett JM, Mcphee DP, Vitez R, O'Malley B, Sharma V, Bosnich W, Seguin A, MacKay J, Regan S (2007) Diverse developmental mutants revealed in an activation-tagged population of poplar. Can J Bot 85:1071–1081

Hedden P, Phillips AL (2000) Gibberellin metabolism: new insights revealed by the genes. Trends Plant Sci 5:523–530

Hiratsu K, Matsui K, Koyama T, Ohme-Takagi M (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. Plant J 34:733–739

Ichikawa T, Nakazawa M, Kawashima M, Muto S, Gohda K, Suzuki K, Ishikawa A, Kobayashi H, Yoshizumi T, Tsumoto Y, Tsuhara Y, Iizumi H, Goto Y, Matsui M (2003) Sequence database of 1172T-DNA insertion sites in *Arabidopsis* activation-tagging lines that showed phenotypes in T1 generation. Plant J 36:421–429

Jansson S, Douglas CJ (2007) *Populus*: a model system for plant biology. Annu Rev Plant Biol 58(435–58):435–458

Jeong DH, An S, Kang HG, Moon S, Han JJ, Park S, Lee HS, An K, An G (2002) T-DNA insertional mutagenesis for activation tagging in rice. Plant Physiol 130:1636–1644

Jeong DH, An S, Park S, Kang HG, Park GG, Kim SR, Sim J, Kim YO, Kim MK, Kim SR, Kim J, Shin M, Jung M, An G (2006) Generation of a flanking sequence-tag database for activation-tagging lines in Japonica rice. Plant J 45:123–132

Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in *Arabidopsis*. Plant Cell 11:2283–2290

Liu YG, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J 8:457–463

Martienssen R (1998) Functional genomics: probing plant gene function and expression with transposons. Proc Natl Acad Sci USA 95:2021–2026

Mathews H, Clendennen SK, Caldwell CG, Liu XL, Connors K, Matheis N, Schuster DK, Menasco DJ, Wagoner W, Lightner J, Wagner DR (2003) Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport. Plant Cell 15:1689–1703

Nakazawa M, Ichikawa T, Ishikawa A, Kobayashi H, Tsuhara Y, Kawashima M, Suzuki K, Muto S, Matsui M (2003) Activation tagging, a novel tool to dissect the functions of a gene family. Plant J 34:741–750

Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D (2003) Control of leaf morphogenesis by microRNAs. Nature 425:257–263

Raizada MN, Nan GL, Walbot V (2001) Somatic and germinal mobility of the RescueMu transposon in transgenic maize. Plant Cell 13:1587–1608

Schomburg FM, Bizzell CM, Lee DJ, Zeevaart JAD, Amasino RM (2002) Overexpression of a novel class of gibberellin 2-oxidases decreases gibberellin levels and creates dwarf plants. Plant Cell 15:1–13

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke JD, Cotton D, Bullis D, Snell J, Miguel T, Hutchison D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff SA (2002) A high-throughput *Arabidopsis* reverse genetics system. Plant Cell 14:2985–2994

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjarvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leple JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Sci 313:1596–1604

Weigel D, Ahn JH, Blàzquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferràndiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia Y, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J (2000) Activation tagging in *Arabidopsis*. Plant Physiol 122:1003–1013

Zubko E, Adams CJ, Machaekova I, Malbeck J, Scollan C, Meyer P (2002) Activation tagging identifies a gene from *Petunia hybrida* responsible for the production of active cytokinins in plants. Plant J 29:797–808