Plant Biotechnology Journal

Plant Biotechnology Journal (2010), pp. 1–17 doi: 10.1111/j.1467-7652.2010.00537.x

Gibberellin-associated cisgenes modify growth, stature and wood properties in *Populus*

Katherine M. Han[1], Palitha Dharmawardhana[1], Reneé S. Arias[1,†], Cathleen Ma[1], Victor Busov[2] and Steven H. Strauss[1,*]

[1]*Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR, USA*
[2]*Michigan Technological University, School of Forest and Environmental Science, Houghton, MI, USA*

Received 12 October 2009;
revised 23 April 2010;
accepted 27 April 2010.
**Correspondence* (fax 1 541 737 1393; e-mail Steve.Strauss@oregonstate.edu)
†Present address: Genomics and Bioinformatics Research Unit, USDA-ARS, 141 Experiment Station Rd., Stoneville, MS 38776-0038, USA.
Accession numbers for EMBL sequence database: NC_008467 REGION: 31167983–31169547 (GA2ox2), NC_008468 REGION: 21539604–21541466 (RGL1_2), NC_008474 REGION: 8596643–8598445 (GAI1), NC_008470 REGION: 6458253–6460076 (RGL1_1), NC_008480 REGION: 1725399–1726822 (GA20ox7), NC_008477 REGION: 13135920–13136837 (*ubiquitin*).

Keywords: Intragenic, *Populus*, transformation, plant architecture, genetic engineering, forest biotechnology.

Summary

We studied the effects on plant growth from insertion of five cisgenes that encode proteins involved in gibberellin metabolism or signalling. Intact genomic copies of Pt*GA20ox7, PtGA2ox2,Pt RGL1_1, PtRGL1_2* and *PtGAI1* genes from the genome-sequenced *Populus trichocarpa* clone Nisqually-1 were transformed into *Populus tremula* × *alba* (clone INRA 717-1B4), and growth, morphology and xylem cell size characterized in the greenhouse. Each cisgene encompassed 1–2 kb of 5′ and 1 kb of 3′ flanking DNA, as well as all native exons and introns. Large numbers of independent insertion events per cisgene (19–38), including empty vector controls, were studied. Three of the cisgenic modifications had significant effects on plant growth rate, morphology or wood properties. The *PtGA20ox7* cisgene increased rate of shoot regeneration *in vitro*, accelerated early growth, and variation in growth rate was correlated with *PtGA20ox7* gene expression. *PtRGL1_1* and *PtGA2ox2* caused reduced growth, while *PtRGL1_2* gave rise to plants that grew normally but had significantly longer xylem fibres. RT-PCR studies suggested that the lack of growth inhibition observed in *PtRGL1_2* cisgenic plants was a result of co-suppression. *PtGAI1* slowed regeneration rate and both *PtGAI1* and *PtGA20ox7* gave rise to increased variance among events for early diameter and volume index, respectively. Our work suggests that cisgenic insertion of additional copies of native genes involved in growth regulation may provide tools to help modify plant architecture, expand the genetic variance in plant architecture available to breeders and accelerate transfer of alleles between difficult-to-cross species.

Introduction

Transgenic methods to aid breeding have largely been used to incorporate traits that are not otherwise available in native gene pools, such as herbicide resistance or novel mechanisms of pest tolerance as seen in the widely grown transgenic crops (ISAAA, 2010). However, with the increasing knowledge of plant genomes and the increased power of genetic modification (GM) to insert genes into a variety of species and genotypes, GM approaches might also be able to aid conventional breeding by delivering useful alleles based on modifications to the function, number or expression of native genes. These GM products are typically called 'cisgenic' when intact native genes are used, including all regulatory sequences, exons and introns in the sense direction and are derived from sexually compatible relatives. They are called 'intragenic' when modified versions of native genes also derived from sexually compatible relatives are used (Jacobsen and Nataraja, 2008; Jacobsen and Schouten, 2008; Poupin and Arce-Johnson, 2005; Schouten *et al.*, 2006). These kinds of GM approaches would be especially important to forest trees because of their long generation times, very limited degree of domestication, and thus inability to rapidly integrate useful alleles into breeding populations (Jacobsen and Nataraja, 2008; Poupin and Arce-Johnson, 2005;

Strauss and Brunner, 2004). GM approaches, often employing intragenic coding regions, have been predominantly used to modify woody biomass quality of trees (e.g., lignin quantity and quality: (Chapple *et al.*, 2007; Li *et al.*, 2003; Talukder, 2006; Tzfira *et al.*, 1998). Apart from basic physiological studies, by comparison, little attention has been devoted to modifications of growth rate, form and most plant physiological processes (exceptions include Eriksson *et al.*, 2000; Jing *et al.*, 2004). Nevertheless, growth and form characteristics are considered major goals for domestication of woody plants, particularly for bioenergy plantations (Ragauskas *et al.*, 2006).

A major reason that cisgenic and intragenic modifications of growth traits have been rarely used in plants may be because of the high frequency of co-suppression often seen there. However, similar approaches have been widely used in animals, particularly fish, to improve growth rate (Nam *et al.*, 2001). In vegetatively propagated plants such as poplar, however, the rate of co-suppression and unstable gene expression appears to be very modest (Brunner *et al.*, 2007a; Li *et al.*, 2008, 2009), making such approaches potentially feasible. Insertion of native genes for modification of growth traits takes advantage of the effects of a second functional locus containing all regulatory sequences and thus presumably functioning in a similar manner to the native endogene. However, the insertion in new genomic loci can modify the expression of the gene. This has been long known in transgenic research as 'position effect' (Matzke and Matzke, 1998). This produces a modified intensity and pattern of expression for the second locus because of its unique chromosomal position and epigenetic marks associated with insertion, and the possibility of new interactions between the regulatory factors, RNAs, and protein and biosynthetic products of the native locus. It is well known that gene and genome duplication is a major source of novelty for molecular evolution (Zhang, 2003). Variation in gene copy number is common in animal and plant populations (Morgante *et al.*, 2005; Stranger *et al.*, 2007), and natural processes such as stresses can stimulate the movement of transposon-associated genes to different chromosomal positions (Feschotte *et al.*, 2002). Thus, although the mutational processes are not identical, cisgenic insertions can be viewed as a directed means to accelerate the formation of duplicative transposition-like changes in genomes and thus amplify genetic variation in specific genes and their associated phenotypic traits.

Genes encoding components of gibberellin (GA) metabolic and signalling pathways are ideal targets for genetic modification of growth and form characteristics through cisgenic modifications (reviewed in Busov *et al.*, 2008). GAs are important hormonal regulators of shoot elongation, internode length, leaf shape and size, flowering, seed germination and fruit development, where they affect both cell division and cell elongation (Sponsel and Hedden, 2007). Elevated bioactive GA concentrations increase shoot elongation while deficiencies cause dwarfism. Two dwarfing forms of GA-modifying genes have been in wide commercial use in agriculture: *Rht* in wheat and *ga20ox* in rice (Hedden, 2003). GA metabolic and signalling pathways in Arabidopsis have been intensively dissected, and many of the genes involved have been identified and characterized (Yamaguchi, 2008). The poplar genome sequence allows easy cloning of full genome copies of homologous genes, including flanking and/or intervening sequences.

We chose for study the negative GA regulators *RGL*, *GAI* and *GA2*ox and the positive regulator *GA20ox*, because their effects on plant growth are reasonably well understood. Most of the key enzymes that control flux of bioactive GAs, or are central GA signal regulators, have been identified (Yamaguchi, 2008). The level of bioactive GAs is controlled by three enzymes—GA20-oxidase, GA3-oxidase and GA2-oxidase. GA20-oxidase (GA20ox) and GA3-oxidase (GA3ox) catalyse the last two steps in production of bioactive GAs, while GA2-oxidase is the main GA deactivation enzyme (Biemelt *et al.*, 2004; Hedden *et al.*, 2002; Olszewski *et al.*, 2002). Overexpression of *GA2ox* transgenes have been observed to cause dwarfism in tobacco and poplar (Biemelt *et al.*, 2004; Busov *et al.*, 2003). In contrast, overexpression of *GA20ox* caused elongated internodes in rice, citrus and aspen (Eriksson *et al.*, 2000; Fagoaga *et al.*, 2007; Oikawa *et al.*, 2004). In poplar, *GA20ox* overexpression caused increased growth, biomass production and xylem fibre length, but decreased root initiation (Eriksson *et al.*, 2000) and impeded cessation of shoot growth under short days (Eriksson and Moritz, 2002).

DELLA domain proteins like GA-insensitive (GAI), repressor of GA-1 (RGA) and RGA-like (RGL) are transcription factors from the GRAS family that contain a conserved 17-amino acid DELLA domain (Hussain and Peng, 2003; Pysh *et al.*, 1999). They play central, repressive roles in the GA signal transduction pathway. GAI, RGA and RGL proteins are all inhibitory to GA-induced gene expression effects that promote growth, such as flower initiation, seed germination and stem elongation. The deletion of the DELLA domain causes a dominant negative effect on

GA signalling because the protein becomes resistant to degradation and therefore constitutively represses GA signalling (Busov *et al.*, 2006; Hussain and Peng, 2003; Wen and Chang, 2002). Overexpression of *gai* and *rgl1* in poplar caused severe dwarfism and increased root proliferation (Busov *et al.*, 2006).

Poplar is widely considered the model woody plant for genomics and biotechnology (Bradshaw *et al.*, 2000; Brunner *et al.*, 2004). Because of its available genome sequence (Tuskan *et al.*, 2006), facile transformation, stable transgene expression (Strauss *et al.*, 2004), ease of clonal propagation and rapid growth, poplar provides an ideal system for studying the possible value of cis- and intragenic approaches in a forestry and woody biofuel crop. We show that several GA cisgenes provide substantial and diverse kinds of variation in poplar growth and form, suggesting that cisgenic approaches may be useful tools for acceleration of some forms of woody plant breeding.

Results

Selection of genes and generation of cisgenic plants

We used microarray expression data to identify genes for cisgenic transformation. Because we were interested in modifying stem growth and form, we selected genes that were preferentially expressed in xylem and phloem tissues (Table S1). The cisgenes included at least 2 kb upstream and 1 kb of downstream sequence, as well as all exon and intron sequences (Figure S1). Using the genome sequence, we designed primers, amplified the selected genes and sequenced each of the recovered fragments. We then compared the sequence of amplified fragments to the genome sequence. Protein alignments indicated 100% identity for *GA20*, *GA2* and *RG1* (data not shown). However, *RG2* and *GAI* displayed two and three amino acid changes, respectively (Figures S2 and S3). Inspection of the changes suggested that they should not substantially affect protein function. For RG2, the sequences have one deletion in an area where deletions were seen in other species, and the amino acid change created a biochemically similar substitution (both hydrophobic: I, isoleucine to V, valine). The same two kinds of changes were seen for GAI, including conserved hydrophobic amino acid changes (from L, leucine to V). The identified polymorphisms were expected, given the highly heterozygous nature of the wild poplar tree that was sequenced (Tuskan *et al.*, 2006) and the possibility of PCR errors. All amplified fragments were inserted into the genome using a standard transformation protocol (Filichkin *et al.*, 2006). Multiple cisgenic events were recovered (Table 1), and the successful insertion of the cisgene was PCR verified using primers that were directed at flanking T-DNA sequence that was not present in WT plants (Table S10).

Rate of regeneration during transformation

We noticed that plants transformed with *GA20* gave rise to visible shoots considerably earlier and in greater numbers than controls and other constructs. A nonparametric analysis of data on time of shoot regeneration showed that *GA20* increased the rate of regeneration (Wilcoxin Rank Sum Test, $P = 0.006$), while *RG2* decreased regeneration ($P = 0.0001$; Figure 1). After 9 months of transformation and regeneration, there were 46 confirmed transgenic events for *GA20*, and these took an average of 143 days until they began producing shoots. In contrast, the controls had 21 confirmed events that took an average of 219 days, the 18 *RG2* events took 168 days and the four *GAI* events took 286 days.

Table 1 Characteristics of genes and cisgenic populations studied

Gene	Abbreviation	Gene model	Independent events per construct	GA gene function
GA20ox7	*GA20*	fgenesh4_pm.C_LG_XIV000079	38	Biosynthesis of active forms of GA
GA2ox2	*GA2*	eugene3.00012757	26	GA degradation
GAI1	*GAI*	fgenesh4_pm.C_LG_VIII000571	20	DELLA protein, negative regulator of GA effect
RGL1_1	*RG1*	estExt_fgenesh4_pm.C_LG_IV0241	19	DELLA protein
RGL1_2	*RG2*	gw1.II.2542.1	19	DELLA protein
Empty Vector Control	Control	NA	23	NA

NA: Not applicable.

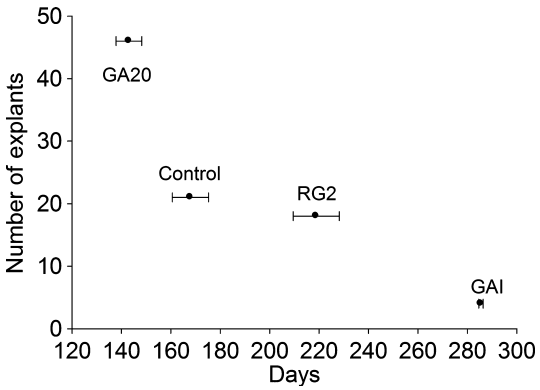

Figure 1 Cisgenic constructs affect rate of regeneration of transgenic shoots. Number of PCR-confirmed transgenic events for each construct vs. days in tissue culture before shoots that grew larger than 0.5 cm were scored. Error bars represent standard errors based on individual explants.

Cisgene-associated variation in plant growth and form

We grew cisgenic plants under greenhouse conditions and measured growth, morphology and wood fibre/cell properties (Table S3). We found that cisgenic manipulations affected plant size and architecture (Tables 2 and S2). We also used principal component analysis to reduce dimensionality of the data. The first principal component (PC1) captured majority of the growth traits (PCs and their interpretations are further discussed later).

Growth responses

Several of the cisgenic plants showed strong and statistically significant effects on one or more aspects of growth (Figures 2 and 3 and Tables 2 and S2).The *GA20* cisgene appeared to impart visible growth-promoting effect (Figure 2e). For example, at week 6, the volume index of *GA20* cisgenic plants was 40% larger than controls ($P < 0.05$; Tables 2 and S2). Consistent with expectation, *GA2* and *RG1* cisgenics displayed significant decreases in biomass ranging between half and three-quarters that of the controls (Tables 2 and S2). However, the most dramatic growth-inhibiting effects were seen in *RG1* cisgenics, which were significantly smaller for the majority of measured and estimated parameters (Figure 3). A growth-inhibitory effect of GAI was visible towards the end of the experiment, but did not reach statistical significance (Tables 2 and S2, Figure 3).

Morphology/wood responses

The different cisgenes had highly variable and complex effects on plant morphology. Plants transformed with *RG1* and *GA2* were more chlorotic (less green) than control

Table 2 Effect of cisgenic modification on growth and form characteristics. Values in one row that share at least one of the same letter in their subscripts indicate that the construct means are not significantly different according to a Tukey's multiple comparisons test for pairwise differences between means at a confidence level of 0.05

Characteristics	Control	*GA20*	*RG2*	*GAI*	*GA2*	*RG1*
Cell area (mm^2)¶‡	0.012_{a}	0.009_{a}	0.010_{a}	0.011_{a}	0.012_{a}	0.011_{a}
Chlorosis Score†	0.12_{c}	0.07_{d}*	0.05_{cd}	0.16_{cd}*	0.33_{b}	0.60_{a}*
Diameter week 6 (cm)§	0.32_{a}	0.34_{a}*	0.33_{a}*	0.30_{b}*	0.27_{c}*	0.26_{c}
Diameter week 10 (cm)	0.894_{abc}	0.912_{ab}*	0.909_{abc}*	0.855_{bcd}	0.820_{cd}	0.767_{e}*
Fibre length (mm)‡	0.53_{a}	0.53_{a}	0.64_{b}	0.56_{ab}	0.53_{a}	0.59_{ab}
Height Week 6 (cm)§	21.4_{a} *	24.9_{a}*	22.4_{a}*	22.6_{a}*	16.6_{b}	15.2_{b}*
Height Week 10 (cm)	111_{ab}	110_{ab}*	114_{ab}	107_{abc}*	102_{bcd}*	97_{cd}
Ratio root/total mass¶§	0.258_{a}*	0.256_{a}*	0.249_{a}	0.265_{a}	0.253_{a}*	0.271_{a}
SLA (cm^3)	145.3_{abc}*	149.9_{a}	143.7_{abcd}*	124.4_{bcde}*	128.7_{bcde}	115.1_{e}*
Stem Taper Ratio Week 10	140_{bc}	133_{c}	138_{bc}	147_{b}	153_{ab}*	168_{a}
Total mass (g)	36.5_{abc}*	38.1_{ab}*	39.9_{ab}*	32.6_{bcd}*	27.9_{cde}	23.0_{de}
Volume Index Week 6 (cm^3)§	2.4_{b}	3.3_{a}*	2.5_{ab}*	2.4_{b}*	1.4_{c}	1.2_{c}
Volume Index Week 10 (cm^3)	90.8_{abc}	94.4_{ab}*	95.6_{ab}*	80.9_{bcd}*	71.3_{cde}	58.9_{de}*
PC1§	0.377_{ab}	0.637_{a}*	0.580_{a}*	0.004_{b}*	-0.639_{c}	-0.996_{c}

Subscripts for diameter, height and volume index at week 6 and PC1 were determined by two-sided *t*-tests on the log-transformed data at a confidence level of $P < 0.05$. Asterisks indicate that the events within the construct have significantly different means from each other (one-way ANOVA, $P < 0.05$). See, Table S1 for data for all response variables.

†Differences determined by pairwise, two-sided chi-square tests on the data before averaging.

‡ANOVA between events within a construct not performed because of the lack of replication ability at that level.

§Significant differences were found between variances of the constructs, so comparisons of construct means were made by first taking the logarithm of the data and then comparing the means using a two-sided *t*-test for each combination of two constructs.

¶No significant differences between construct means.

Figure 2 Plant growth and morphology at different stages of development during the greenhouse experiment. The top row shows the plant size at (a) week 6, (b) week 8 and (c) week 10;(d) shows a chlorotic plant at week 8; (e) shows eight average-sized *GA20* plants (left) and eight control plants (right) at week 6, showing the rapid early growth imparted by this construct; (f) shows representative leaves from each cisgenic construct.

plants (chi-square test, $P < 0.05$; Figures 2d, 4 and S4). *RG1*-transformed plants exhibited a reduction in leaf size, while *RG2* plants had statistically significantly longer fibre lengths and a smaller PC2 value than controls (Table 2). The faster growing *GA20* plants displayed a smaller stem taper ratio than *GAI*, *GA2 and RG1* cisgenics. In contrast, the slower growing *RG1* and *GA2* plants had significantly larger taper ratios at weeks 6 and 10, respectively (Table S2). Internode length, root/total mass, ln(root/total mass), PC3 and cell area did not show any statistically significant differences between constructs or controls.

PCA summarizes correlations among growth traits

A principal component analysis was performed to reduce the dimensionality of our large data set and to summarize the correlations between the growth and morphology traits studied. The loadings (Table 3) suggest that PC1 was a measure of the overall growth rate of the plants, whereas PC2 mainly reflected a larger root/total mass ratio as well as a negative correlation with single leaf area (SLA) and leaf mass. PC3 primarily reflected internode length. PC1 was closely correlated with volume index, a measure of plant stem size, as indicated by their nearly identical linear regressions with gene expression for *GA20* (discussed in the following text). Graphs of PC2 vs. PC1 and PC3 vs. PC2 illustrate that the PCs did differentiate constructs, but mainly on PC1 (Figure S5), which explained a majority (77%) of the total variance. Variance component analysis of PC1 showed that the cisgene explained 24% of the variance in growth, insertion in an independent position in the genome (i.e., event) explained 32%

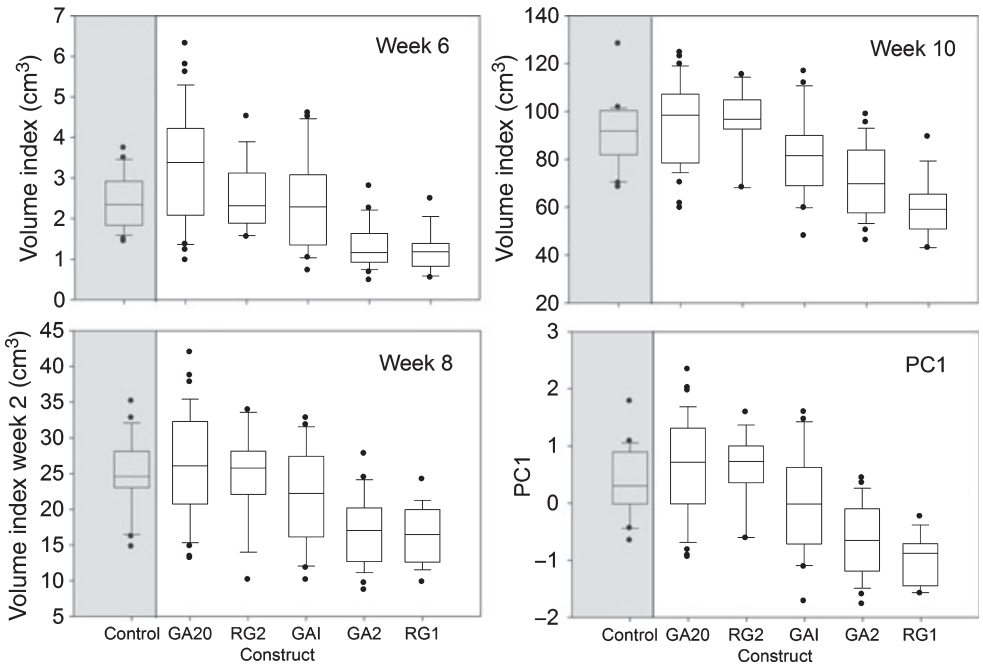

Figure 3 Variation in growth characteristics among different cisgenic modifications. Box and whisker plots of volume index for each construct are shown at weeks 6, 8 and 10, as well as the for first principal component that summarizes variation in overall growth rate. The limits of the boxes indicate the 25th and 75th percentiles, the line in the middle of the box is the construct median, and the whiskers end at the most extreme data point above the 10th percentile and below the 90th. Note differences in scale between the three volume index graphs.

Table 3 Principal component analysis loadings. Values represent PC loadings multiplied by 100. The first three principal components combined describe 99% of the variation in the data. Bold values show the major loadings for each PC. All data except for ln(root/total mass) are untransformed

Growth response	PC1	PC2	PC3
Diameter week 6	**27**	8	−10
Diameter week 8	**28**	2	−2
Diameter week 10	**27**	−12	−2
Height week 6	**26**	13	−10
Height week 8	**28**	6	7
Height week 10	**27**	−5	13
Internode length	−2	−13	**94**
Leaf mass	**25**	**−24**	−2
ln(root/total mass)	9	**57**	15
Root mass	**27**	14	6
Root/total mass	8	**58**	16
Single leaf area	16	**−37**	11
Stem mass	**28**	−11	0
Total mass	**28**	−10	1
Volume Index Week 6	**26**	12	−11
Volume Index Week 8	**28**	4	0
Volume Index Week 10	**28**	−12	1
Variance (%)	77	20	1

and the remaining 44% was because of variance among ramets within events. Thus, more than 50% of the growth variation was explained by the effect of the cisgene and insertion-related variation.

Heterogeneity of variances among events

Cisgene-imparted increases in trait variances can be of significant value in breeding. We therefore analysed the effect of the cisgenes on trait variability among events using Levene's test. Indeed, variances for several of the responses were significantly affected by the cisgenes. These included diameter, height and volume index at week 6, the composite response PC1 and the ratio of root/total mass (Tables S5 and S6 and Figure 5). Variance patterns were complex and tended to be statistically significant mainly at the earlier stages of the experiment. However, to our surprise, we found only minor statistically significant effects when the cisgene-imparted variances were contrasted to that among the empty vector control events. For *GA20* and *GAI* cisgenes, we observed significantly increased variance of volume index and diameter at week 6, respectively

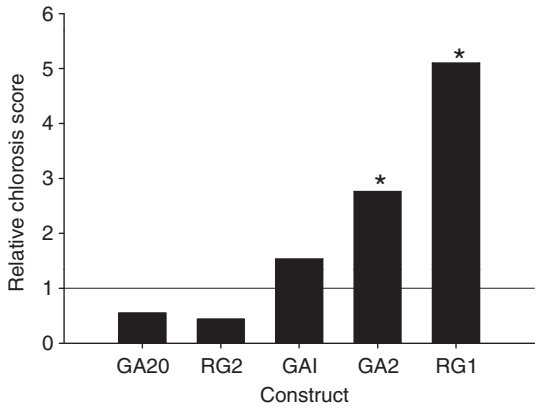

Figure 4 GA2 and RG1 cisgenes cause significant increase in leaf chlorosis. Construct means are given relative to the empty vector control, which was set to 1. Asterisks indicate significant differences from control ($P < 0.05$) based on chi-square tests.

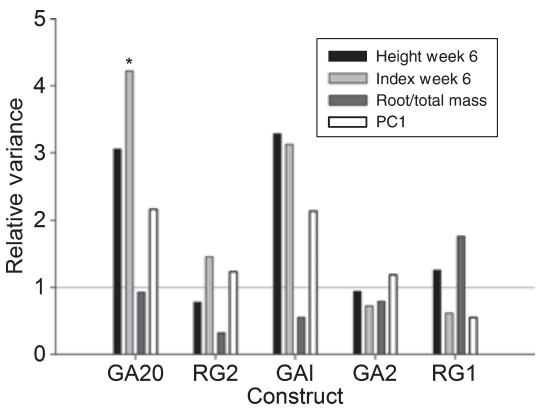

Figure 5 Effect of cisgenic modification on phenotypic variance. Variance among events within cisgenic constructs was standardized to that of the empty vector control events, which was set to 1. The asterisk indicates significantly higher variance than controls. Note that all of the traits graphed had nonhomogeneous variances in pairwise comparisons between at least two constructs based on Levene's test ($P < 0.05$).

(Table S5). For all other traits and cisgenes analysed, the effects on variance were either very small or not statistically significant (Figure 5).

Variation in trait distributions caused by cisgenes

We have recovered a sufficiently large number of independent insertions (events) for each of the studied cisgenes to study how they affected overall trait distributions. For cisgenes to be of most use to breeders, they should shift the entire distributions upwards or downwards. Events within cisgenes were a statistically significant source of variation,

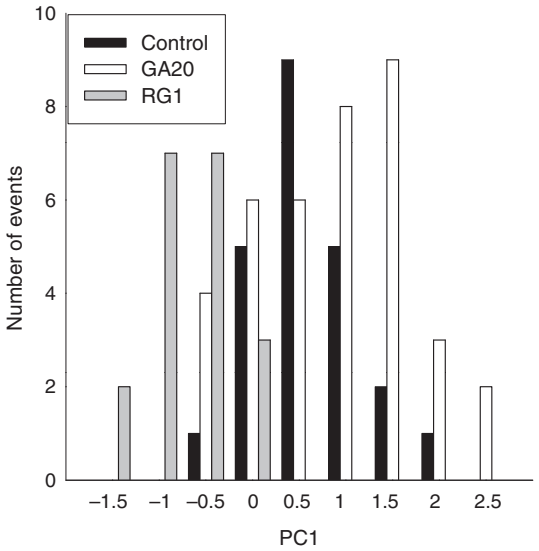

Figure 6 Histogram showing altered distribution of growth rate as summarized by PC1 for the fast-growing *GA20*, control and slow-growing *RG1* transgenic populations.

including the empty vector controls (Table S5). The upward and downward shifts in distributions of overall plant growth from the constructs are illustrated in Figure 6, as summarized by PC1.

Among the controls, many responses showed statistically significant event-based variation (e.g., biomass, root/total mass ratio and its log transformation, SLA, height at week 6, and the composite responses PC2 and PC3 (Table S5). Although most of the mass and mass ratio responses were statistically significant, the strength of the effects and the *P*-values for most of the traits in the *GA20*, *RG2* and *RG1* constructs were considerably stronger than in the controls. After logarithmic transformation, the only statistically significant differences between events in the control construct were for height at week 6, internode length (but not PC3), root mass and SLA.

Associations between gene expression and phenotype

We tested if the insertion of cisgenes led to increased mRNA abundance of the targeted genes. We found a statistically significant increase in gene expression for three (*GA20*, *GAI*, and *RG1*) of the four studied genes (Figure 7). Upon addition of the cisgenes, the expression of *GA20*, *GAI* and *RG1* increased compared to nontransformed controls by factors of 6, 4 and 2, respectively (*t*-test, $P < 0.05$). Expression of *GA2* was not statistically different from the nontransgenic controls.

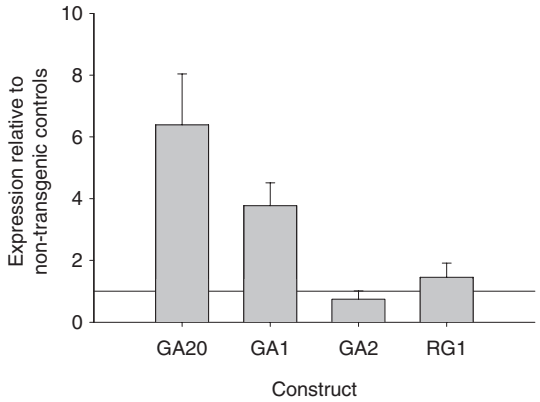

Figure 7 Average cumulative expression of cisgene and endogene relative to nontransgenic controls. Error bars are standard errors of event means.

We performed a linear regression analysis with gene expression as the explanatory variable and event means for each growth trait as the response variables. Based on previous knowledge about the function of similar genes, we hypothesized that plants with higher levels of *GA20* gene expression would show increased plant growth rate, a lower stem taper ratio (tall, narrow-stemmed plants) and a decreased root/total mass ratio. We expected the opposite effects from *GAI*, *RG1*, *RG2* and *GA2*. We therefore conducted one-sided statistical tests where we analysed the association of gene expression with these traits. Because visual inspection suggested that the relationship of gene expression to phenotype appeared to be nonlinear, we also used the logarithm of gene expression values in the analysis. We found that *GA20* was the only construct with a strong correlation between expression and many growth traits and that the correlations for all constructs tended to have higher statistical significance when log-transformed expression data were examined (cf. Tables S7 and S8, Figure 8). The logarithm of *GA20* gene expression was statistically significantly correlated with all heights, diameters, volume indexes, masses and PC1 (Table S8). We did not find statistically significant trait correlations with expression for the other cisgenic modifications.

Evidence of co-suppression in *RG2* cisgenics

Because the *RG2* cisgenics showed a lack of the growth suppression we had expected, we hypothesized that this may be a result of co-suppression. However, despite testing a large number of primers, we were unable to detect the expression of *RG2* gene in both WT and transgenic plants. Therefore, we were able to study only the expression of its close paralog *RG1*. Indeed, nearly all of the *RG2* cisgenic events studied showed decreased *RG1* expression compared to an empty vector control transgenic plant (Figure 9).

Discussion

The goal of this study was to examine the feasibility of a cisgenic approach for modification of biomass growth and architecture in poplar using endogenous genes encoding GA metabolic enzymes and response regulators. Most transgenic studies of gene function or biotechnology employ modifications that cause substantial perturbations in expression, usually via the use of strong promoters to achieve ectopic expression; RNA interference to down-regulate gene's transcript abundance; or tissue-specific promoters to reinforce or redirect gene expression in specific organ/tissue. Here, we show that transformation with native copies of genes produces a range of statistically significant phenotypic changes with the majority of the

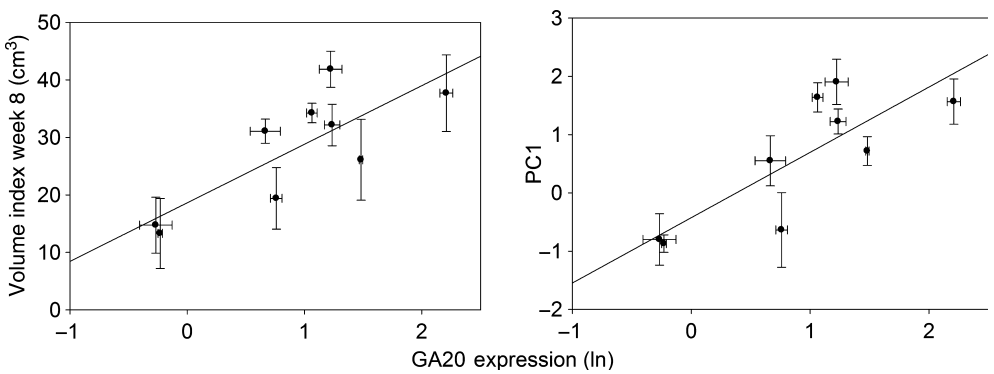

Figure 8 Association of growth characteristics with gene expression for *GA20*. Bars are standard errors for the triplicate reaction RT-PCR expression results (horizontal) and because of variation among ramets within events (vertical). Both correlations are statistically significant ($P < 0.05$), as were the nonlogarithm correlations (Table S6).

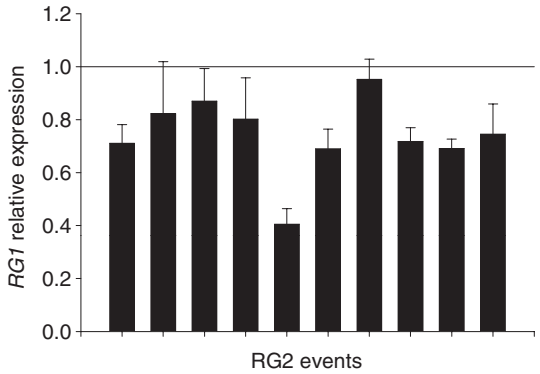

Figure 9 Evidence of co-suppression of the *RG1* gene in *RG2*-transformed plants. *RG1* expression in transgenic plants was reduced ($P < 0.01$) compared to controls.

cisgenes studied. In addition, we found significant event-based variation within the transgenic control plants, supporting the need for studying a number of transgenic control events when cisgenic growth effects are analysed.

We cloned native genes from *P. trichocarpa* including 2 kb upstream and 1 kb downstream of the coding region. We chose a 2-kb upstream region because most plant regulatory sequences contain proximal promoters, the majority of which are located within the few hundred base pairs 5′ from the gene (Shahmuradov *et al.*, 2003). In a study more directly relevant to our own work, Lee *et al.* (2002a) cloned a genomic fragment with the *Arabidopsis RGL2* ORF including 3.3 kb 5′ and 2 kb 3′ of the coding region that had an inserted Ds (dissociator) transposon containing the *GUS* gene. They found that the regulatory region was sufficient to express *RGL2* reliably based on GUS assays and affected plant growth. Kaneko *et al.* (2003) found that 1.8 kb of DNA 5′ to the *GA20* ORF in rice was sufficient to drive consistent expression of GUS. These data suggest that the flanking regions chosen were reasonable choices for obtaining normal control of the corresponding cisgenes studied.

The effect of some cisgenes could be observed very early on in the experiment, during plant transformation and *in vitro* regeneration. For example, cisgenic plants carrying GA20ox were much easier to regenerate when compared to the other cisgenics and/or controls. The *GA20* accelerated while *RG2* and *GAI* delayed the production of differentiated shoots. Thus, it appears that *GA20* may be useful for stimulating transformation and regeneration rate, of possible value in recalcitrant species (Arias *et al.*, 2006). To our knowledge, this beneficial effect of *GA20* to *in vitro* has not been previously reported. Gibberellins are selectively excluded from shoot apical meristems because they have a strong promoting effect on tissue differentiation and growth (Hay *et al.*, 2002; Sakamoto *et al.*, 2001a,b). Therefore, we hypothesize that the increased shoot regeneration is associated with increased GA concentrations because of the insertion of the GA20ox cisgene.

In addition to promoting *in vitro* shoot regeneration, the *GA20* cisgene had a strong growth-promoting effect in young plants; they were visually striking and statistically significant. Similar growth-promoting effects were reported in transgenic plants overexpressing *Arabidopsis GA20ox*. Our results are consistent with the key role of this enzyme in regulating bioactive GA levels in poplar and other species (Israelsson *et al.*, 2005).

We observed that the growth-enhancement effect in GA20ox cisgenics gradually diminished over time during the experiment. This may be a result of limited pot size, resulting in rapid depletion of soil nutrients and water by the faster growing plants. In support of this, we found that the cisgenics with retarded growth showed a higher relative growth rate during the final weeks of the experiment than the ones that displayed an early growth-promoting effect (data not shown).

Our study is likely to provide a lower estimate of the potential strength of GA cisgenes expression : phenotype associations. We determined expression in a modest number of cisgenic events and tissue samples for any single construct (8–10). The association between RNA levels and active GA concentration is likely to be modest because of post-translational modifications and interactions with other regulatory factors, and the majority of effect from GA perturbations may be limited to specific tissues and development times. In addition, the effects of cisgenes are expected to be much more subtle than have been observed in overexpression types of transgenic studies. A much larger sample of events, tissues and biological replications is needed to precisely estimate cisgenic effects. Moreover, for breeding applications, it would also be highly desirable to study a number of environments, including field environments, as the growth-modifying effects of the cisgenes are likely to show strong genotype by environment interaction.

The cisgene effects on growth, where observed, largely agreed with biological expectations. The *GA20ox* cisgenic plants showed increased growth while *GA2ox*,*GAI* and *RG1* cisgenics displayed semi-dwarf characteristics. DELLA domain proteins are strong repressors of GA-signalling pathway and when overexpressed cause severe dwarfism, particularly when carrying truncation or mutation in the

conserved DELLA domain. We have previously showed that *rgl1* from Arabidopsis elicited the most severe dwarfism in poplar (Busov *et al.*, 2006). Therefore, we hypothesized that all of the poplar cisgene homologues containing DELLA domains would have significant growth-retardation effects. Surprisingly, only one of the two *RGL* genes tested caused a change in growth. *RG1* cisgenic plants exhibited semi-dwarf characteristics, while *RG2* plants were as large as or larger than the control plants for many growth traits. The difference in the effects of the two genes could result from the drastically higher expression of *RG1* compared to *RG2*; *RG1* has three- to sevenfold higher expression in the phloem and xylem than does *RG2* (Table S1). Therefore, the insertion of *RG1* may have led to overall higher expression and thus the stronger effect.

The lack of a retarding effect of *RG2* on growth and the increase in fibre length could also be a result of co-suppression. Compared to empty vector controls, the *RG2* plants did have lower levels of *RG1* expression (Figure 9), indicating that the *RG2* transformation may have triggered co-suppression and thus may have been a contributing factor to the generally wild-type phenotype of these plants. Furthermore, as expected for a repressor of GA response, the event with the lowest *RG1* expression in *RG2* cisgenic plants (Figure 9) was found to have the faster growth rate among all of the *RG2* transgenic events (data not shown).

In addition to changes in trait means, we anticipated that GA cisgenes could lead to changes in the level of overall variance among independent insertion events. The changes were largest in the early growth traits, which could be because of early-acting genetic effects of the constructs as discussed earlier, as well as a greater effect of tissue culture and transplanting-induced variation in physiology. The variance among events of *RG1* was nearly double that of the controls, particularly for root biomass, which suggests that this construct might be of particular value where trees with very strong relative root growth are desired, such as for soil carbon sequestration, wind firmness, drought tolerance or biofiltration/remediation. This finding is consistent with our previous observations that overexpression of heterologous *rgl1* gene has a strong positive effect on lateral root formation and root biomass (Busov *et al.*, 2006).

All constructs had traits that showed significant variance among events, including to our surprise the empty vector control (Figures 10 and S3). Nevertheless, as expected, the GA constructs had directed and larger impacts on growth traits, whereas event-based variation in the controls primarily affected various aspects of plant morphology. However, none of the purely transformation-related effects were obvious to the eye at the laboratory or greenhouse level; large sample sizes and statistical analysis were required for their detection. This suggests caution when only one or a few transgenic control events are studied, particularly with subtle cis- or transgenic effects.

Overexpression of a *GA20ox* gene had been previously reported to increase wood fibre length in poplar (Eriksson *et al.*, 2000). We therefore measured fibre lengths in our *GA20* cisgenic plants after the 4-week greenhouse experiment. However, we did not find a significant change in fibre lengths. There could be a number of reasons for this. First, our cisgenic modifications are likely to have elicited subtle and localized changes in GA concentrations and/or response compared to the much stronger and generalized effects expected from the 35S promoter used in the Eriksson *et al.*'s (2000) study. The effects may be insufficient to elicit changes in fibre cell characteristics. Furthermore, GA concentrations are result of complex interplay between synthesis, catabolism and transport of bioactive forms or precursors (Björklund *et al.*, 2007; Israelsson *et al.*, 2005). Therefore, the outcome of our manipulations within this complex network on localized cell types is largely unpredictable. In addition to elongation growth, GAs may also affect the rate of cambial cell proliferation (Björklund *et al.*, 2007). Therefore, complex effects on wood-forming processes could be elicited as a result of the tissue/cellular expression specificity of the particular cisgene employed.

The definition of a cisgene is a gene from a crossable species that includes flanking regulatory regions, exons and introns (Schouten and Jacobsen, 2008). In this study, we have inserted genes from a cottonwood, *P. trichocarpa,* into a hybrid white poplar (*P. alba*) × aspen (*P. tremula*); both of these are species with which *P. trichocarpa* does not normally cross. Because of conserved chromosome structure and slow rate of molecular evolution in poplars and other tree species (Tuskan *et al.*, 2006), the genetic similarity among all of the poplars is very high and thus genes are likely to be very similar in function. Thus, we believe that this study is a good approximation of a formally cisgenic breeding experiment in poplar. For a commercial application, it would be desirable to employ cisgenic selectable markers and associated DNA elements, or to excise exogenous sequences after transformation, depending on regulatory and market considerations.

Our results suggest that cisgenic modification of trees can successfully be used to speed improvement for growth and form traits that are difficult to achieve through

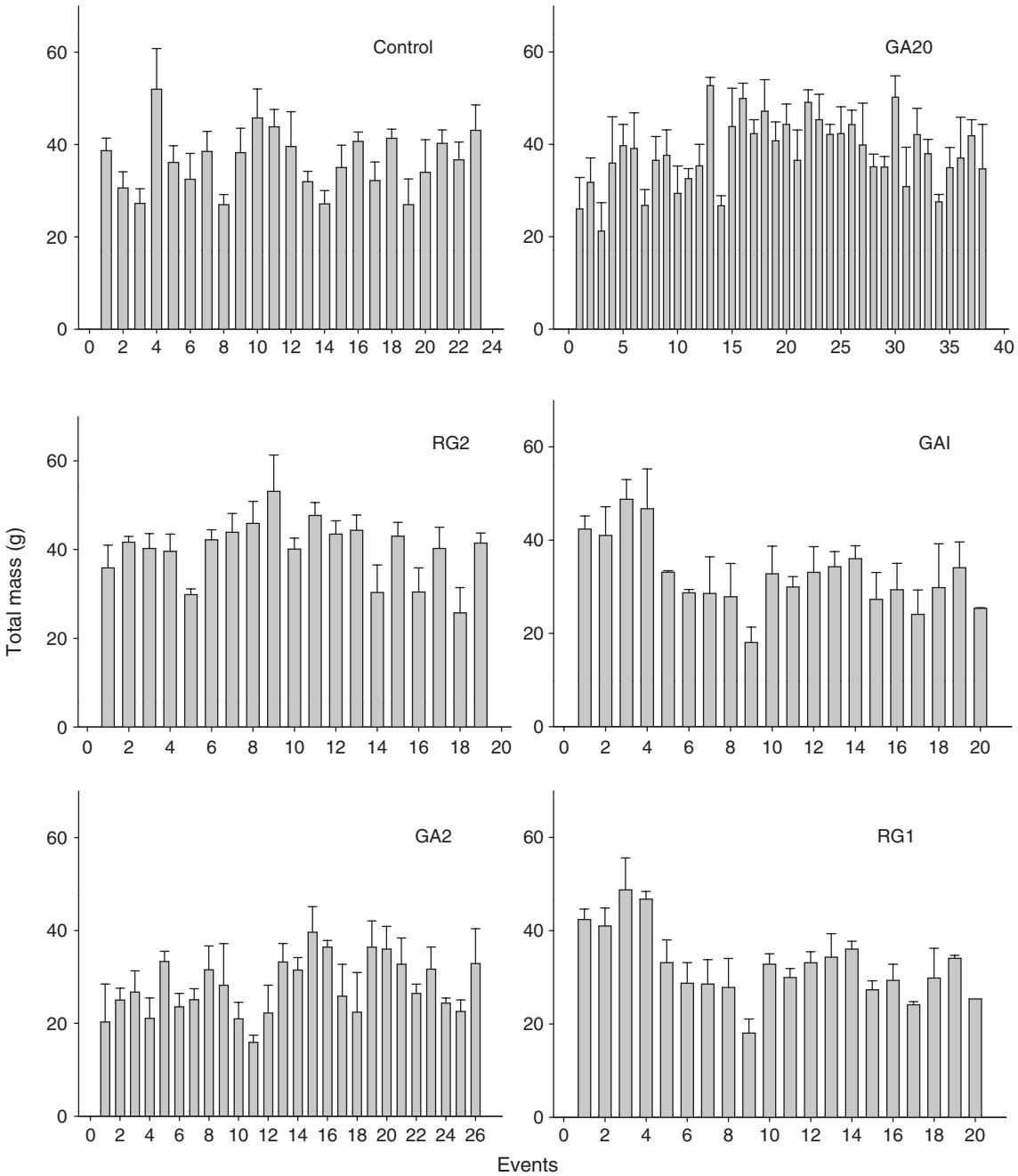

Figure 10 Strong variance among event means within cisgenic and empty vector constructs for total plant biomass. Error bars are standard errors over ramets.

traditional breeding. Specifically, we found that transformation with *GA20*, a biosynthetic enzyme gene, increased plant growth rate and ease of *in vitro* regeneration, while transformation with a GA catabolic enzyme and negative GA signal regulator genes (*GA2*, *RG1*) decreased plant growth rate. There were also changes to a number of wood and morphological traits that varied widely among constructs and events, suggesting that the effects of cisgenic perturbations of GA signalling are complex and require careful evaluation of their effects, including field studies, prior to incorporation in breeding.

Experimental procedures

Bioinformatics

We blasted gene sequences for *GA20ox*, *GA2ox*, *GAI* and *RGL* (Biemelt *et al.*, 2004; Gazzarrini and McCourt, 2003; Lee *et al.*,

2002b; Oikawa *et al.*, 2004; Remington and Purugganan, 2002; Thomas and Sun, 2004) that were obtained from NCBI against the poplar genome (JGI, version 1.1, August 2005). There are several paralogs for each gene in the poplar genome; thus, we chose one or two gene models from each gene type for the experiment (Table 1). Our selections were based on the criterion of strong gene expression in xylem and phloem observed in poplar microarrays (Table S1), as we were interested in modifying stem growth and form in poplar. The microarray data used corresponded to data sets from our laboratory generated using a poplar Nimblegen microarray (http://www.nimblegen.com/; Brunner *et al.*, 2007b) and uploaded to GEO (accession numbers GSE21481 and GSE21485). These arrays contained 55 794 predicted transcripts from the *Populus trichocarpa* genome sequencing project, 126 mitochondrial and chloroplast gene models and 9995 unigenes derived from aspen EST sequences (Sterky *et al.*, 2004). Each gene model was represented by three replicated 60mer isothermal probes on the array. Array hybridization, quality control and data extraction were carried out by Nimblegen using their established microarray-processing pipeline (Wang *et al.*, 2006). Microarray data were normalized across all the arrays in the experiment using the Bioconductor—Robust Multiarray Averaging (RMA) protocol. The genes chosen were *GA20ox7*, *GA2ox2*, *GAI1*, *RGL1_1* and *RGL1_2*, and for brevity will hereafter be referred to as *GA20*, *GA2*, *GAI*, *RG1* and *RG2*, respectively.

Vector construction and plant material

Each of the five genes was amplified from *Populus trichocarpa* Nisqually-1 genomic DNA using PCR with the Taq HiFi (Invitrogen, Carlsbad, CA, USA) or Herculase (Stratagene) enzymes. PCR primers were designed to include approximately 2 kb upstream (5′) and 1 kb downstream (3′) of the protein coding sequences (see Figures 11 and S1 for genomic fragment and T-DNA schematics; Table S9 shows cloning primer sequences).

For *GA20*, *GAI* and *RG2*, we cloned the complete PCR products into either the Zero Blunt TOPO4 vector or the TOPO4 TA cloning vector, both from Invitrogen. *RG1* transcript and terminator were amplified and cloned into vector TOPO4 (Invitrogen). The *RG1* promoter and part of the transcript sequence were separately cloned into TOPO4. Both TOPO4 vectors with *RG1* native DNA were transformed into DAM-DCM *E. coli* cells (New England Biolabs, Ipswich, MA, USA), and the DNA was re-extracted. Then, the SbfI-BclI 1914-bp fragment from the promoter was cloned into the same restriction sites of the transcript and terminator vector, from which an SbfI-NotI 4297-bp fragment containing the native *RG1* gene with its promoter and terminator was transferred to the corresponding sites of pBART. *GA2* transcript and terminator were amplified, and the 3030-bp fragment was cloned into TOPO4. The *GA2* promoter had been previously amplified and subcloned into TOPO4, from where the 1190-bp SpeI fragment was removed and cloned into the SpeI site of the transcript and terminator vector. The product contained the *GA2* promoter, transcript and terminator in a 4169-bp fragment. This fragment was cut with EcoRV and PmlI and cloned into the PmeI site of pBART. The resulting clones were screened by PCR with Platinum Taq DNA Polymerase High Fidelity (Invitrogen), then by restriction digestion with NotI and PvuI. We sequenced the entire PCR products in TOPO4 vector to ensure that no mutations were introduced by PCR.

We cut the TOPO constructs containing sequence-verified inserts of the *GAI*, *RG1*, *RG2*, *GA20* and *GA2* genes with NotI and then ligated the GA gene fragments into pBART that was also cut with NotI and treated with Antarctic phosphatase (New England Biolabs). pBART is a binary vector based on pART27 (Gleave, 1992), but which includes a selectable marker for basta rather than kanamycin resistance. The basta gene employed was a *phosphinothricin acetyl transferase* gene, with both a NOS promoter and a terminator. We screened the pBART ligation constructs for the presence of the GA gene by restriction digestion and sequenced the gene-plasmid borders with M13F and M13R sequencing primers.

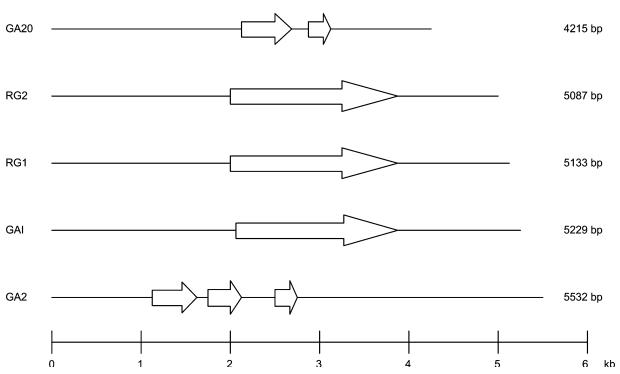

Figure 11 Structure of cisgenic DNA inserts. Sizes and locations of exons (arrows) are shown for each gene used in this study.

Transformation and regeneration

We transformed the pBART constructs into the *Agrobacterium* hypervirulent strain AGL1 (Chabaud *et al.*, 2003) and then co-cultivated these Agrobacteria with leaf discs and stem explants from the clone 717-1B4 from INRA, France (*Populus tremula* × *P. alba*). We performed transformation and regeneration using the methods described previously (Filichkin *et al.*, 2006). Differences from Filichkin's methods include using spectinomycin rather than kanamycin for selection during bacterial growth, and basta (generic name glufosinate-ammonium 5 mg/L, Chem Service West Chester, PA, USA) rather than kanamycin for selection during plant regeneration. Also, leaf discs were regenerated on SIM (shoot-induction media) containing 0.4 μM TDZ (thidiazuron), which is twice what Filichkin used. Explants that exhibited leaf growth in the presence of basta were transferred to SIM (Filichkin *et al.*, 2006) containing 0.1 μM BAP (benzlaminopurine) without basta to allow more rapid growth. Shoot elongation and rooting media are as described by Filichkin, except that basta was used as the selection agent. Explants that produced roots on this media were screened using PCR for the presence of their expected gene insert (primer sequences in Table S10). PCR-positive cisgenics from this screen were propagated to provide 10–15 genetically identical ramets, five of which we transferred to soil at week 0 in propagation medium. We acclimated them over 4 weeks in plastic bags for high humidity with the bags open for the last week, transferred them to 4-L pots containing SunGro Horticulture 'Sunshine' soil with Osmocote 14-14-14 slow-release fertilizer at week 4 for a 2-week acclimation period, and then used them in the greenhouse study.

Greenhouse study

This experiment consisted of six types of GM plants (empty vector control plus GA constructs), each including 19–38 cisgenic events with four ramets each, totalling 562 individual plants. Plants were placed in a greenhouse in a completely randomized design at 30-centimetre spacing between the centres of all plant pots (designated as week 6 after in vitro culture). In the greenhouse, we provided high-pressure sodium artificial light as needed for 16 h a day to supplement the Willamette Valley daylight throughout July 2007. Plants were fertilized once every 2 weeks with 500 ppm 20-10-20 fertilizer (Jack's Professional Water-Soluble Fertilizer, J.R. Peters, Inc, Allentown, PA, USA). Because of the rapid growth of the plants, they were moved to 45-centimetre spacing centre to centre 2 weeks into the experiment (week 8).

DNA and RNA extraction

DNA for PCR screening of cisgenics before the greenhouse experiment was extracted from between 200 and 400 mg of tissue culture–grown leaf and shoot tip material using an isopropanol precipitation method. Tissue was immersed in a 4 : 1 mixture of extraction buffer (2 M NaCl, 200 mM Tris–HCl, pH8, 70 mM EDTA, 20 mM Na-metabisulphite) and 5% N-lauroylsarcosine, ground with a rounded plastic tip in a drill press and incubated at 60 °C for 1 h. The mixture was centrifuged at 14 000 rpm for 15 min, the supernatant was transferred to a clean tube, a 1 : 2 mixture of 5 M NH_4OAc and isopropanol was added, it was cooled at −20 °C for 15 min, centrifuged 14 000 rpm for 10 min, the supernatant was discarded and the pellet was rinsed with cold 70% EtOH. The DNA pellet was air-dried for 15 min and then resuspended in ddH_2O. We tested the DNA using PCR with primer to both the 5′ and 3′ sides of the native genes, including flanking regions from the pBART construct (primer sequences in Table S10). We screened the empty vector control plants, transformed with pBART, for the entire expected insert from pBART in one PCR product.

We collected shoot tip tissue from each of ten events per *GA2*, *GA20*, *RG1* and *GAI* construct for RT-PCR expression analysis. We chose a subset of events based on a preliminary analysis of growth rate so that we included three to four events from the low, high, and average rates of growth in regression analyses. Each sample of 200 mg of shoot tip stem tissue was collected at the end of the greenhouse experiment and was placed immediately in liquid nitrogen and then stored in the −80 °C freezer. We extracted RNA once for each event from tissue that had been pooled from equal amounts of tissue from all ramets in the event. RNA quality was determined by visualization on a 1% agarose gel and by the ratio of absorbencies at 260 and 280 nm. The RNA of two events from *GA2* and one event from *GA20* appeared degraded and was therefore not used in the analysis. RNA was extracted using the RNeasy Plant Mini Kit from Qiagen (Hilden, Germany)

Quantitative real-time PCR

We analysed gene expression for 10 events per construct for *RG1* and *GAI*, eight events for *GA2* and nine events for *GA20*. Expression data were produced by pooling equal amounts of young shoot tissue from all the ramets for an event, extracting RNA from the pooled tissue and producing cDNA, and running RT-PCR on the cDNA. We also included four pooled control plants transformed with pBART, four pooled nontransformed 717 plants and three nontransformed Nisqually plants (from internodes two and three for this sample). RNA was treated with the TURBO DNA-free Kit from Ambion (Foster City, CA, USA) and reverse polymerized using Superscript III from Invitrogen with random hexamers and $(dT)_{16}$ primers to make cDNA. Primers were designed using Clone Manager and Primer Express (http://www.scied.com/pr_cmbas.htm, http://www3.appliedbiosystems.com/AB_Home/index.htm, respectively, accessed April 2009), and sequence identity of the amplified fragments was determined by DNA sequencing.

We created standard curves using cisgenic templates for each sample in a dilution series of 1, 0.2, 0.04, 0.008 and 0.0016. However, for *GA20*, the relatively lower gene expression necessitated the dilution series to be 1, 0.5, 0.25, 0.125 and 0.0625. Samples were run in triplicate with SYBR Green kit (Invitrogen) on an Mx3000p real-time PCR system (Stratagene, http://www.stratagene.com, April 2009). RT-PCR primer sequences can be found in Table S12. Conditions for RT-PCR reactions were as follows: 50 °C for 2 min, 95 °C for 2 min, followed by forty cycles of 95 °C for 30 s, annealing temperature (Table S11) for 30 s, 72 °C for 30 s, then followed by one cycle of 95 °C for 1 min, 55 °C for 30 s and 95 °C for 30 s. Transcript abundance was determined based on changes in C_t values relative to 717 *Ubiquitin* (gene model estExt_fgenesh4_pm.C_LG_XI0348) and were compared with the relative transcript abundance in nontransgenic 717 plants.

Despite extensive selection and testing of dozens of primers, because of the very high levels of DNA sequence similarity between the cisgenes and endogenes (Table S12), only for a single construct (*GAI*) were we able to reliably distinguish among the two (primer sequences in Table S11). Thus, for most constructs, the data presented are total gene expression, and the primers chosen amplified cDNA from 717 and Nisqually equally based on visual inspection of PCR results. Cross-amplification between the two *RGL* genes precluded the generation of useful data from RT-PCR of *RG2*.

Co-suppression qRT-PCR was carried out under the same conditions as the other qRT-PCR reactions with the following exceptions. Two ramets were randomly chosen per event, and the material was pooled before RNA extraction. Two runs were performed with two replicates per sample for a total of four replicates. PCR efficiencies were 98% for both the *RG1* and the *Ubiquitin* primer sets. Primers used for the co-suppression experiment were designed based on the *P. trichocarpa* sequence: forward 5′-cagacactccttggacaacct-3′ and reverse 5′-tgttcgctctgacaaagtacctgaa-3′.

Phenotypic assessment of cisgenic plants

All of the measured and calculated phenotypic traits are described in the online material in Tables S2 and S3 and briefly summarized below. We measured plant height and stem diameter 5 cm above the soil at weeks 6, 8 and 10 of the 4-week experiment (Table S3). At week 8, we observed some leaf chlorosis and thus scored it; plants that exhibited a light green colour in most of their leaves were given a '1', those that had a deeper green

colour were assigned a '0' (Figure 2d). After the 10th week, we harvested and then oven-dried all of the plant tissues and determined their weights. Final measurements also included the length of four internodes located at the 20th through 23rd nodes from the top of the plant; dry mass of stems, leaves and roots; and the mean single leaf area; fibre length and cell area of xylem internode cells between the 20th and 23rd leaves. For mass measurements, the plant tissue was dried in a paper bag and then weighed; the average mass of a paper bag was subtracted from the measured weight.

For fibre length measurements, we selected one tree from each of eight events per construct that were largely the same ones used for gene expression analysis. We collected 1- to 2-cm stem sections from the 20th internode from the top in an FAA fixative solution (10 volume % of 40% formaldehyde, 50 volume % of 95% ethanol and 5 volume % of glacial acetic acid) and then rinsed them with deionized water and soaked them in an aqueous solution that was 20% hydrogen peroxide and 30% glacial acetic acid by volume for about 4 h at 60 °C before storing them at room temperature overnight. We then rinsed the samples three times with deionized water and placed them overnight for staining in an ethanol solution containing approximately 10 mg of safranin. The staining and visualization of the samples was carried out according to the procedure described by Peterson *et al.* (2007). They were then rinsed gently with ethanol the next day. We took an aliquot of fibres from each sample to make three slides (three cover slips on one slide) using 70% glycerine and took 15 pictures at 4× magnification for each sample (five from each cover slip) and used ImageJ (http://rsbweb.nih.gov/ij/, accessed April 2009) to measure the lengths of 150 fibres/sample. We measured cell area for these same plants by slicing cross-sections with a microtome, taking 20× pictures using an Olympus Vanox 2 research grade microscope and using ImageJ software to measure the cell area. We took at least 150 measurements for each plant for fibre length and cell area.

Statistical analysis

We identified three outlier events that were excluded from statistical analysis. Two of the events were from the control group, and one event was from *RG2*. All of them exhibited highly abnormal morphology and growth retardation, with the same symptoms in all ramets that were beyond standard deviations from the mean. For the remaining plants, we averaged the growth responses for the four ramets in each event so that each event has only one value for each growth trait for each analysis, resulting in 145 data points for each growth trait. The exception to this is the generalized linear model fitted in SAS that we used to determine the variance at the construct, event and ramet level, for which we used the data for each of the 562 ramets. All tests performed used an alpha level of 0.05 to determine statistical significance, though *P*-values between 0.05 and 0.1 are sometimes noted.

We visually inspected the data for normality using a box and whisker plot for each construct and for each growth trait. The chlorosis score was the only trait that was clearly not normal in its distribution (it is not a continuous variable), thus we used a series of two-sided, pairwise chi-square tests to compare constructs. Mean values for events were used in linear regression analysis of growth traits vs. gene expression.

We combined some of the raw data to create more comprehensive growth response variables; this included volume index, total mass, root/total mass ratio, stem aspect ratio and the principal components (Table S4). We calculated volume index for each of weeks 6, 8 and 10 by multiplying the height times diameter squared for each individual at each week. We calculated the total mass by adding the stem, root and leaf mass. We used the natural log of the root/total biomass ratio for statistical analysis to better approximate normality. The stem taper ratio is the height divided by the square of the diameter. To reduce dimensionality and help to summarize the data set, we used principal components analysis on data that were scaled to have a mean of zero and a standard deviation of one. The growth traits we used to calculate the principal components were the height, diameter and volume index at weeks 6, 8 and 10; mean single leaf area; internode length; leaf, stem, root and total masses; and the ratio of root/total mass and the natural log of the ratio. Data were scaled in the analysis. The first, second and third principal components were used for further analysis, which together account for 99% of the variance in the data.

We performed most analyses in SAS version 9.2 (http://www.sas.com/technologies/analytics/statistics/stat/, Cary, NC, USA). We used one-way ANOVA analysis to compare the means between the different constructs for each response variable; Tukey's multiple comparisons test to determine differences between event and construct means; and a Brown–Forsythe modified Levene's test on the constructs to compare variances. We found that five traits had significant differences between the variances of at least two of the constructs (modified Levene's test, $P < 0.05$). We therefore compared the construct means by first performing a log transformation on the data for those traits before comparing each combination of two constructs using a two-way *t*-test. This analysis tests for the null hypothesis of homogeneity of variance between the groups by performing an ANOVA on the absolute deviations from group medians. An ANOVA analysis was also performed on the natural log of the data for each trait to correct for unequal variances and found no differences in the comparisons of construct means, but with fewer statistically significantly different means for events within constructs.

To summarize sources of variance, we determined the per cent variance in the data at each level of construct, event and ramet using a generalized linear model on the first principal component where the constructs and events were both treated as random effects. A mixed model was used to determine differences in construct means for stem aspect ratio after accounting for height. All characteristics except fibre length and cell area had 5 and 139 degrees of freedom, while fibre length had 5 and 42 and cell area had 5 and 37.

Tissue culture regeneration time between co-cultivation with *Agrobacterium* and the date that shoots were at least 0.5 cm long was compared between constructs. Because of non-normal distribution, the Wilcoxon rank sum test in SPlus (http://www.splus.com/products/default.asp) was used to compare the construct means for *GA20* and *RG2* individually with the controls. For this test, we used 46 *GA20* events, 21 control events and 18 *RG2* events. Four *GAI* events available at that time were not compared statistically because of the small number of observations. These

events were all co-cultivated within the same 6 weeks and were grown on the same regeneration media. The other constructs were not comparable to the controls because of a media change part-way through transformation.

For study of phenotypic traits in relation to gene expression, we selected three to four events from the middle and extremes of the growth distribution for analysis and performed linear regression analyses on RT-PCR expression vs. growth responses. Log transformations were employed where the data suggested it gave a more linear response in transformed data. The one-way linear regression tested the hypothesis of positive relationships between *GA20* gene expression and growth traits (height, diameter, single leaf area, volume index, internode length, mass, as well as the composite response PC1), and negative correlation between root/total mass ratio and PC2. It tested the hypothesis of opposite response in the other three constructs, which are expected to inhibit rather than increase shoot growth traits. Two-tailed tests were also run to detect statistically significant but unexpected correlations. *GA20* had 7 degrees of freedom, *GA2* had 6, and both *RG1* and *GAI* had 8 degrees of freedom.

We calculated relative growth rate for each event from weeks 6 to 8 and from weeks 8 to 10 by dividing the later volume index data point by the earlier one for each time period for each ramet, then taking the means. We compared the relative growth of the constructs with the larger plants (*GA20*, control, *GAI* and *RG2*) to that of the constructs with the smaller plants (*GA2* and *RG1*) using a one-sided *t*-test.

Acknowledgements

This research was supported by the US Department of Energy grant DE-FG02-05ER64113 and by industrial members of the Tree Biosafety and Genomics Research Cooperative based at Oregon State University. We would also like to thank Cathleen Ma for help with every aspect of plant growth, Meghan Allen for help with cloning, Lisa Ganio and Vikas Vikram for statistical help, Barbara Lachenbruch for help with fibre length analysis, Barbara Taylor for her help in taking microscope pictures, Carol Rivin and Dave Harry for their constructive feedback on the first draft of the paper, and Hai Yue Han, Luke Sugie and James Summerton for help with the greenhouse experiment.

References

Arias, R.S., Filichkin, S.A. and Strauss, S.H. (2006) Divide and conquer: development and cell cycle genes in plant transformation. *Trends Biotechnol.*, **24**, 7.

Biemelt, S., Tschiersch, H. and Sonnewald, U. (2004) Impact of altered gibberellin metabolism on biomass accumulation, lignin biosynthesis, and photosynthesis in transgenic tobacco plants. *Plant Physiol.*, **135**, 254–265.

Björklund, S., Antti, H., Uddestrand, I., Moritz, T. and Sundberg, B. (2007) Cross-talk between gibberellin and auxin in development of *Populus* wood: gibberellin stimulates polar auxin transport and has a common transcriptome with auxin. *Plant J.*, **52**, 499–511.

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

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

Brunner, A., Li, J., DiFazio, S., Shevchenko, O., Montgomery, B., Mohamed, R., Wei, H., Ma, C., Elias, A., VanWormer, K. and Strauss, S. (2007a) Genetic containment of forest plantations. *Tree Genet. Genomes*, **3**, 75–100.

Brunner, A.M., DiFazio, S.P., Crasta, O., Fei, Z., Mane, S.P., Sobral, B. and Dharmawardhana, P. (2007b) Microarray expression analysis of poplar development, In *Plant and animal genomes XV conference*. San Diego, CA, USA: Scherago Assoc., NY.

Busov, V.B., Meilan, R., Pearce, D.W., Ma, C., Rood, S.B. and Strauss, S.H. (2003) Activation tagging of a dominant gibberellin catabolism gene (GA 2-oxidase) from Poplar that regulates tree stature. *Plant Physiol.*, **132**, 1283–1291.

Busov, V., Meilan, R., Pearce, D.W., Rood, S.B., Ma, C.P., Tschaplinski, T.J. and Strauss, S.H. (2006) Transgenic modification of gai or rgl1 causes dwarfing and alters gibberellins, root growth, and metabolite profiles in *Populus*. *Planta*, **224**, 288–299.

Busov, V.B., Brunner, A.M. and Strauss, S.H. (2008) Genes for control of plant stature and form. *New Phytol.*, **177**, 589–607.

Chabaud, M., de Carvalho-Niebel, F. and Barker, D.G. (2003) Efficient transformation of *Medicago truncatula* cv. Jemalong using the hypervirulent *Agrobacterium tumefaciens* strain AGL1. *Plant Cell Rep.*, **22**, 46–51.

Chapple, C., Ladisch, M. and Meilan, R. (2007) Loosening lignin's grip on biofuel production. *Nat. Biotech.*, **25**, 746–748.

Eriksson, M.E. and Moritz, T. (2002) Daylength and spatial expression of a gibberellin 20-oxidase isolated from hybrid aspen (*Populus tremula* L. × *P. tremuloides* Michx.). *Planta*, **214**, 920–930.

Eriksson, M.E., Israelsson, M., Olsson, O. and Moritz, T. (2000) Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nat. Biotech.*, **18**, 784–788.

Fagoaga, C., Tadeo, F.R., Iglesias, D.J., Huerta, L., Lliso, I., Vidal, A.M., Talon, M., Navarro, L., Garcia- Martinez, J.L. and Pena, L. (2007) Engineering of gibberellin levels in citrus by sense and antisense overexpression of a GA 20-oxidase gene modifies plant architecture. *J. Exp. Bot.*, **58**, 1407–1420.

Feschotte, C., Jiang, N. and Wessler, S.R. (2002) Plant transposable elements: where genetics meets genomics. *Nat. Rev. Genet.*, **3**, 329–341.

Filichkin, S.A., Meilan, R., Busov, V., Ma, C.P., Brunner, A. and Strauss, S.H. (2006) Alcohol-inducible gene expression in transgenic *Populus*. *Plant Cell Rep.*, **25**, 8.

Gazzarrini, S. and McCourt, P. (2003) Cross-talk in plant hormone signalling: What arabidopsis mutants are telling us. *Ann. Bot.*, **91**, 605–612.

Gleave, A.P. (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol. Biol.*, **20**, 1203–1207.

Hay, A., Kaur, H., Phillips, A., Hedden, P., Hake, S. and Tsiantis, M. (2002) The gibberellin pathway mediates KNOTTED1-type homeobox function in plants with different body plans. *Curr. Biol.*, **12**, 1557–1565.

Hedden, P. (2003) The genes of the green revolution. *Trends Genet.*, **19**, 5.

Hedden, P., Phillips, A.L., Rojas, M.C., Carrera, E. and Tudzynski, B. (2002) Gibberellin biosynthesis in plants and fungi: a case of convergent evolution? *J. Plant Growth Regul.*, **20**, 319–331.

Hussain, A. and Peng, J. (2003) DELLA proteins and GA signalling in Arabidopsis. *J. Plant Growth Regul.*, **22**, 134–140.

ISAAA (2010) Crop Biotech Update.

Israelsson, M., Sundberg, B. and Moritz, T. (2005) Tissue-specific localization of gibberellins and expression of gibberellin-biosynthetic and signaling genes in wood-forming tissues in aspen. *Plant J.*, **44**, 494–504.

Jacobsen, E. and Nataraja, K.N. (2008) Cisgenics – facilitating the second green revolution in India by improved traditional plant breeding. *Curr. Sci.*, **94**, 1365–1366.

Jacobsen, E. and Schouten, H. (2008) Cisgenesis, a new tool for traditional plant breeding, should be exempted from the regulation on genetically modified organisms in a step by step approach. *Potato Res*, **51**, 75–88.

Jing, Z.P., Gallardo, F., Pascual, M.B., Sampalo, R., Romero, J., Navarra, A.Td. and Canovas, F.M. (2004) Improved growth in a field trial of transgenic hybrid poplar overexpressing glutamine synthetase. *New Phytol.*, **164**, 137–145.

Kaneko, M., Itoh, H., Inukai, Y., Sakamoto, T., Ueguchi-Tanaka, M., Ashikari, M. and Matsuoka, M. (2003) Where do gibberellin biosynthesis and gibberellin signaling occur in rice plants? *Plant J.*, **35**, 104–115.

Lee, S., Cheng, H., King, K.E., Wang, W., He, Y., Hussain, A., Lo, J., Harberd, N.P. and Peng, J. (2002a) Gibberellin regulates Arabidopsis seed germination via RGL2, a GAI/RGA-like gene whose expression is up-regulated following imbibition. *Genes Dev.*, **16**, 646–658.

Li, L., Zhou, Y., Cheng, X., Sun, J., Marita, J.M., Ralph, J. and Chiang, V.L. (2003) Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proc. Natl Acad. Sci. USA*, **100**, 4939–4944.

Li, J., Brunner, A., Shevchenko, O., Meilan, R., Ma, C., Skinner, J. and Strauss, S. (2008) Efficient and stable transgene suppression via RNAi in field-grown poplars. *Transgenic Res.*, **17**, 679–694.

Li, J., Brunner, A.M., Meilan, R. and Strauss, S.H. (2009) Stability of transgenes in trees: expression of two reporter genes in poplar over three field seasons. *Tree Physiol.*, **29**, 299–312.

Matzke, A.J.M. and Matzke, M.A. (1998) Position effects and epigenetic silencing of plant transgenes. *Curr. Opin. Biotechnol.*, **1**, 7.

Morgante, M., Brunner, A., Pea, G., Fengler, K., Zuccolo, A. and Rafalski, A. (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat. Genet.*, **37**, 6.

Nam, Y.K., Noh, J.K., Cho, Y.S., Cho, H.J., Cho, K.-N., Kim, C.G. and Kim, D.S. (2001) Dramatically accelerated growth and extraordinary gigantism of transgenic mud loach *Misgurnus mizolepis*. *Transgenic Res.*, **10**, 353–362.

Oikawa, T., Koshioka, M., Kojima, K., Yoshida, H. and Kawata, M. (2004) A role of OsGA20ox1, encoding an isoform of gibberellin 20-oxidase, for regulation of plant stature in rice. *Plant Mol. Biol.*, **55**, 687–700.

Olszewski, N., Sun, T.-p. and Gubler, F. (2002) Gibberellin signaling: biosynthesis, catabolism, and response pathways. *Plant Cell*, **14**, S61–S80.

Peterson, M., Dietterich, H. and Lachenbruch, B. (2007) Do douglas-fir branches and roots have juvenile wood? *Wood Fiber Sci.*, **39**, 651–660.

Poupin, M. and Arce-Johnson, P. (2005) Transgenic trees for a new era. *In Vitro* Cell. Dev. Biol., Plant, **41**, 91–101.

Pysh, L.D., Wysocka- Diller, J.W., Camilleri, C., Bouchez, D. and Benfey, P.N. (1999) The GRAS gene family in Arabidopsis: sequence characterization and basic expression analysis of the *SCARECROW-LIKE* genes. *Plant J.*, **18**, 111–119.

Ragauskas, A.J., Williams, C.K., Davison, B.H., Britovsek, G., Cairney, J., Eckert, C.A., Frederick, W.J., Hallett, J.P., Leak, D.J., Liotta, C.L., Mielenz, J.R., Murphy, R., Templer, R. and Tschaplinski, T. (2006) The path forward for biofuels and biomaterials. *Science*, **311**, 484–489.

Remington, D.L. and Purugganan, M.D. (2002) GAI homologues in the hawaiian silversword alliance (Asteraceae-Madiinae): molecular evolution of growth regulators in a rapidly diversifying plant lineage. *Mol. Biol. Evol.*, **19**, 1563–1574.

Sakamoto, T., Kamiya, N., Ueguchi-Tanaka, M., Iwahori, S. and Matsuoka, M. (2001a) KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. *Genes Dev.*, **15**, 581–590.

Sakamoto, T., Kobayashi, M., Itoh, H., Tagiri, A., Kayano, T., Tanaka, H., Iwahori, S. and Matsuoka, M. (2001b) Expression of a gibberellin 2-oxidase gene around the shoot apex is related to phase transition in rice. *Plant Physiol.*, **125**, 1508–1516.

Schouten, H.J. and Jacobsen, E. (2008) Cisgenesis and intragenesis, sisters in innovative plant breeding. *Trends Plant Sci.*, **13**, 260–261.

Schouten, H., Krens, F. and Jacobsen, E. (2006) Cisgenic plants are similar to traditionally bred plants. *EMBO Rep.*, **7**, 4.

Shahmuradov, I.A., Gammerman, A.J., Hancock, J.M., Bramley, P.M. and Solovyev, V.V. (2003) PlantProm: a database of plant promoter sequences. *Nucleic Acids Res.*, **31**, 114–117.

Sponsel, V.M. and Hedden, P. (2007) Gibberellin biosynthesis and inactivation. In *Plant Hormones* (Davies, P.J. ed) pp 63–94: Netherlands: Springer.

Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P., Brunner, A.M., Charbonnel-Campaa, L., Lindvall, J.J., Tandre, K., Strauss, S.H., Sundberg, B., Gustafsson, P., UhlÃ©n, M., Bhalerao, R.P., Nilsson, O., Sandberg, G., Karlsson, J., Lundeberg, J. and Jansson, S. (2004) A *Populus* EST resource for plant functional genomics. *Proc. Natl Acad. Sci. USA*, **101**, 13951–13956.

Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C., Tyler-Smith, C., Carter, N., Scherer, S.W., Tavare, S., Deloukas, P., Hurles, M.E. and Dermitzakis, E.T. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*, **315**, 6.

Strauss, S.H. and Brunner, A.M. (2004) Tree biotechnology in the 21st century: transforming trees in the light of comparitive genomics. In *The BioEngineered Forest: Challenges to Science and Society* (Bradshaw, S.H.Sa.H.D. ed), pp 76–97, Washington, D.C., USA: Resources for the Future.

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

Talukder, K. (2006) Low-lignin wood-a case study. *Nat. Biotechnol.*, **24**, 2.

Thomas, S.G. and Sun, T.P. (2004) Update on gibberellin signaling. A tale of the tall and the short. *Plant Physiol.*, **135**, 668–676.

Tuskan, G.A., 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, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., 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, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., 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, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y. and Rokhsar, D. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr& Gray). *Science*, **313**, 1596–1604.

Tzfira, T., Zuker, A. and Altman, A. (1998) Forest-tree biotechnology: genetic transformation and its application to future forests. *Trends Biotechnol.*, **16**, 8.

Wang, X., He, H., Li, L., Chen, R., Deng, X.W. and Li, S. (2006) NMPP: a user-customized NimbleGen microarray data processing pipeline. *Bioinformatics*, **22**, 2955–2957.

Wen, C.K. and Chang, C. (2002) Arabidopsis *RGL1* encodes a negative regulator of gibberellin responses. *Plant Cell*, **14**, 4.

Yamaguchi, S. (2008) Gibberellin metabolism and its regulation. *Ann. Rev. Plant Biol.*, **59**, 225–251.

Zhang, J.Z. (2003) Evolution by gene duplication: an update. *Trends Ecol. Evol.*, **18**, 292–298.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Schematic of T-DNA region in the transgenic constructs.

Figure S2 Amino acid sequence alignment for RGL1_2.

Figure S3 Amino acid sequence alignment for GAI1.

Figure S4 Event means for chlorosis score.

Figure S5 Examples of principal component differentiation of constructs.

Table S1 Relative expression values from Nisqually-1 microarray studies.

Table S2 Means and statistical test results of measured characteristics for each construct.

Table S3 Measured Traits.

Table S4 Calculated Traits.

Table S5 Comparison of variances among events within constructs for growth traits.

Table S6 Levene's test results for analyses of homogeneity of variances among groups of constructs.

Table S7 Correlation coefficients between untransformed gene expression and growth traits for event means.

Table S8 Correlation coefficients between the natural log of gene expression and untransformed values of growth traits for event means.

Table S9 Primer sequences for cloning entire genomic fragments from poplar.

Table S10 Primer sequences for PCR screening of transformants.

Table S11 Primer sets for RT-PCR used to characterize gene expression in selected transgenic events.

Table S12 Number of paralogs and per cent similarity of Nisqually transgenes to their 717 homologues and Nisqually paralogs.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

