Identification of Quantitative Trait Loci Influencing Wood Property Traits in Loblolly Pine (*Pinus taeda* L.). III. QTL Verification and Candidate Gene Mapping

Garth R. Brown,* Daniel L. Bassoni,[†,1] Geoffrey P. Gill,* Joseph R. Fontana,[†] Nicholas C. Wheeler,[‡] Robert A. Megraw,[‡] Mark F. Davis,[§] Mitchell M. Sewell,[†,] Gerald A. Tuskan** and David B. Neale*,[†,2]**

**Department of Environmental Horticulture, University of California, Davis, California 95616, [†]Institute of Forest Genetics, Pacific Southwest Research Station, USDA Forest Service, Davis, California 95616, [‡]Weyerhaeuser Company, Weyerhaeuser Technical Center, Tacoma, Washington 98477, [§]National Renewable Energy Laboratory, Golden, Colorado 80401 and **Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831*

Manuscript received November 1, 2002
Accepted for publication April 4, 2003

ABSTRACT

A long-term series of experiments to map QTL influencing wood property traits in loblolly pine has been completed. These experiments were designed to identify and subsequently verify QTL in multiple genetic backgrounds, environments, and growing seasons. Verification of QTL is necessary to substantiate a biological basis for observed marker-trait associations, to provide precise estimates of the magnitude of QTL effects, and to predict QTL expression at a given age or in a particular environment. Verification was based on the repeated detection of QTL among populations, as well as among multiple growing seasons for each population. Temporal stability of QTL was moderate, with approximately half being detected in multiple seasons. Fewer QTL were common to different populations, but the results are nonetheless encouraging for restricted applications of marker-assisted selection. QTL from larger populations accounted for less phenotypic variation than QTL detected in smaller populations, emphasizing the need for experiments employing much larger families. Additionally, 18 candidate genes related to lignin biosynthesis and cell wall structure were mapped genetically. Several candidate genes colocated with wood property QTL; however, these relationships must be verified in future experiments.

A continuous, as opposed to discrete, distribution of phenotypic values is a feature of many traits important to animal and plant breeding, as well as many traits impacting human health. Variation in these quantitative, or complex, traits is influenced by multiple genetic loci with relatively small effects coupled with environmental and epistatic interactions. Quantitative trait locus (QTL) mapping is a well-developed discipline that dissects the inheritance of complex traits into discrete Mendelian genetic factors. The number and location of chromosomal regions affecting trait variation and the magnitude of their effects can be determined by associating genotypes with phenotypes in a segregating population. In a limited number of cases, QTL mapping has identifed genetic markers suitable for the improvement of breeding populations by marker-assisted selection (Bernacchi *et al.* 1998; Hardin 2000) as well as genes with direct influence on phenotype (Cormier *et al.* 1997; Frary *et al.* 2000; Steinmetz *et al.* 2002).

Numerous factors influence the ability to detect a QTL. Using computer simulations, Beavis (1994) first showed the impact of sample size. Small segregating populations (100–200 progeny) typical of many QTL mapping experiments resulted in the detection of few loci with disproportionately large effects on phenotype and with poor congruence of QTL position between simulations. These findings were supported empirically in maize by Melchinger *et al.* (1998) and by Utz *et al.* (2000). In addition, genetic background, environment, and interactions among QTL affect QTL detection. For example, the expression of a QTL in long-lived organisms, such as perennial plants, is likely to be modified by changing biotic and abiotic factors on a seasonal or yearly basis. With particular reference to sample size, many experimental studies have ignored these limitations (Utz *et al.* 2000), which has left an incomplete portrayal of the genetic architecture of many complex traits. Larger, replicated follow-up experiments that sample time, space, and genotypes are required. Furthermore, these experiments provide the opportunity to verify QTL detected in earlier experiments to support an underlying biological basis for observed marker-trait associations.

Among forest trees, QTL mapping has focused on wood properties and traits related to adaptation and

[1]*Present address:* Virtual Arrays, Sunnyvale, CA 94809.

[2]*Corresponding author:* Institute of Forest Genetics, Pacific Southwest Research Station, USDA Forest Service, Department of Environmental Horticulture, University of California, 1 Shields Ave., Davis, CA 95616. E-mail: dneale@dendrome.ucdavis.edu

growth (reviewed in Sewell and Neale 2000). In loblolly pine (*Pinus taeda* L.), the leading timber species of North America, the physical and chemical properties of wood have been studied extensively in a single population of modest size (Sewell *et al.* 2000, 2002). The physical properties of wood, which have a major influence on the quality and end use of sawed lumber, include wood-specific gravity (*wsg*) and microfibril angle (*mfa*). The chemical properties of wood, which impact the pulping process, are largely determined by the relative amounts of cellulose, hemicellulose, and lignin. Biologically, wood is essentially a matrix of cell walls and the lumen of secondary xylem (Megraw 1985). It can be considered as the end product of the collective action of many genes modulating the morphology and composition of secondary xylem cell walls in response to environmental and developmental signals. Consistent with this hypothesis, Sewell *et al.* (2000) identified 39 QTL for *wsg* and 7 for *mfa*, each accounting for 5.4–15.7% of the phenotypic variance.

Verification of the findings of Sewell *et al.* (2000, 2002) is required to assess how well the genetic architecture of wood properties in loblolly pine has been described. True QTL verification implies complete replication of all experimental parameters. For many conifer species, low to moderate rates of clonal propagation, coupled with environmental heterogeneity inherent in large test sites, complicates the design of such experiments. In this report, QTL verification is defined as the repeated detection, at a similar position on the genetic map, of a QTL controlling a trait under more than one set of experimental conditions. The goals of this study were (1) to perform a QTL analysis on a larger (N = 457), independent population of the original QTL detection population to verify the existence of previously detected QTL; (2) to obtain accurate estimates of the percentage of phenotypic variation explained by the QTL; (3) to compare QTL detected in an unrelated pedigree with QTL in the original detection pedigree to determine if a similar suite of QTL are expressed and if additional QTL are revealed; and (4) to identify positional candidate genes underlying wood property phenotypes by genetic mapping and colocation with QTL influencing wood properties.

MATERIALS AND METHODS

Mapping populations: Three populations from two three-generation outbred pedigrees of loblolly pine were considered (Table 1). QTL influencing wood properties were initially identified in the *detection* population of the *QTL* pedigree, which consists of 172 progeny located at six sites in Oklahoma and Arkansas (Groover *et al.* 1994; Sewell *et al.* 2000, 2002). The *verification* population consists of 457 progeny derived from remating the parents of the *QTL* pedigree. The *unrelated* population consists of 445 progeny derived from the *base* pedigree (Devey *et al.* 1994). Both populations were established by Weyerhaeuser in 1994 on adjacent sites near New Bern, North Carolina. A smaller population of the *base* pedigree and the *detection* population are reference populations used for routine genetic mapping in loblolly pine (http://dendrome.ucdavis.edu/Synteny/refmap.html).

Genotypic data and map construction: Methods pertaining to restriction fragment length polymorphism (RFLP) and expressed sequence tag polymorphism analyses in loblolly pine followed Devey *et al.* (1991) and Temesgen *et al.* (2001), respectively. Evenly spaced markers were chosen for each mapping population, with preference given to those that segregated in both parents (*i.e.*, fully informative markers). Sex-averaged linkage maps of the *verification* population and the *unrelated* population were constructed from genotypic segregation data using Joinmap 1.4 (Stam 1993) according to Sewell *et al.* (1999). A consensus map, which allowed the relative placement of QTL onto a single integrated map, was constructed according to Sewell *et al.* (1999).

Phenotypic measurements: A 5-mm radial wood core was taken for each progeny of the *verification* and *unrelated* populations at ~1.4 m above ground and cropped at the pith and outer ring. Earlywood and latewood measurements for a variety of physical wood properties (Table 1) were determined for either one or three growth rings (rings 4–6 from the pith, respectively) and their averages were calculated (composite traits) as described in Sewell *et al.* (2000). Rings 4–6 represent predominantly juvenile wood in loblolly pine, which requires 7–10 years of growth before the onset of mature wood production (Megraw 1985). A number of cores were excluded due to an excess of compression wood; as a result, the actual number of phenotypic data points per trait ranged from 381 to 409 in the *verification* population and 408 to 434 in the *unrelated* population. A 12-mm core was also taken from 280 progeny of only the *verification* population for analysis of chemical properties of earlywood and latewood in the fifth growth ring according to Sewell *et al.* (2002). Models for projection to latent structures (PLS) were used to predict the chemical composition of cell walls (*i.e.*, α-cellulose, lignin, galactan, xylan, and mannan content) from pyrolysis molecular beam mass spectrometry data using multivariate statistics (Davis *et al.* 1999). PLS estimates from two independently analyzed subsamples were averaged and used as traits in QTL analysis. Chemical properties were measured by chemical content per unit weight rather than per unit volume. Because wood is ~97% lignin and holocellulose (*i.e.*, α-cellulose and hemicellulose), an inverse relationship exists between lignin and cellulose content on a per-unit-weight basis. As a result, an increase in lignin content could actually be due to a reduction in α-cellulose and vice versa. Therefore, the QTL detected are described as cell wall chemistry (*cwc*) traits (Table 1) rather than as QTL associated with any specific wood chemistry component (Sewell *et al.* 2002).

QTL analysis: Associations between segregating genetic markers and phenotypic variability for wood property traits in the *verification* and *unrelated* mapping populations were detected using the interval mapping approaches of Knott *et al.* (1997) and QTL Express (Seaton *et al.* 2002), a World Wide Web-based interface for the method of Haley *et al.* (1994). A minor modification to the genotype file was required to enable running the F_2 QTL Analysis Servlet of QTL Express with data from a three-generation outbred pedigree (C. S. Haley, personal communication). Estimated QTL positions and associated F-statistics were identical between software packages.

Each linkage group was scanned at 5-cM intervals for locations explaining a high proportion of the phenotypic variance using a one-QTL model interval analysis. Only regions of the genome that exceeded chromosome-wide $P < 0.05$ (suggestive level) or $P < 0.01$ (significant level) significance in support

TABLE 1

Loblolly pine mapping populations and phenotypic traits for QTL analyses of physical and chemical wood properties

	Pedigree		
	Detection	*Verification*	*Unrelated*
Grandparents	$G_1 \times G_2$ $G_3 \times G_4$	$G_1 \times G_2$ $G_3 \times G_4$	$G_5 \times G_6$ $G_7 \times G_8$
Parents	$P_1 \times P_2$	$P_1 \times P_2$	$P_5 \times P_6$
Progeny	172	457	445
Trait and rings analyzed			
Wood-specific gravity (*ewsg, lwsg*)[a]	Rings 2–11	Rings 4–6	Rings 4–6
Percentage of late wood (*%lw*)	Rings 2–11	Rings 4–6	Rings 4–6
Microfibril angle (*emfa, lmfa*)	Rings 3, 5, 7	Ring 6	Ring 6
Cell wall chemistry (*ecwc, lcwc*)[b]	Ring 5	Ring 6	Not assayed

[a] *wsg* is a measure of the total amount of cell wall substance and within an annual ring has three main determinants: *wsg* of earlywood (xylem cells having thin walls and large lumens: *ewsg*), *wsg* of latewood (xylem cells with thicker walls and smaller lumens: *lwsg*), and the percentage of latewood (*%lw*).

[b] Mass peaks collected from the pyrolysis molecular beam mass spectrophotometer were associated with the amounts of α-cellulose, galactan, mannan, xylan, and lignin, collectively termed *cwc* traits.

of the existence of a QTL are reported. These thresholds were determined by performing 1000 permutations of the data as implemented in QTL Express. Note that these thresholds correspond approximately to genome-wide significance levels of 0.6 and 0.12, respectively, following Bonferroni correction (Lynch and Walsh 1998). Therefore, it is probable that some QTL are false positives, but are reported to the mapping community as recommended by Lander and Kruglyak (1995). To compare these results to QTL observed previously in the *detection* population, it was also necessary to assess marker-trait associations using the suggestive ($0.01 > P > 0.005$) and significant ($P < 0.005$) thresholds employed by Sewell *et al.* (2000, 2002). Only a few differences in accepting or rejecting the null hypothesis by either method were found; therefore, comparisons between experiments were considered valid and a reanalysis of the *detection* population was not performed.

A two-dimensional analysis at 5-cM intervals was also performed to fit a two-QTL model for each linkage group. Permutation tests have not been implemented for this model in QTL Express and the suggestive and significant levels of Sewell *et al.* (2000, 2002) were used in determining significance.

The model used to test the effect of QTL alleles as reported in Sewell *et al.* (2000, Table 3) assumes that the grandparents of each parent have divergent *wsg* phenotypes. This assumption is not valid for other traits measured for progeny of the *detection* and *verification* populations since grandparent phenotypes were not assessed. For the same reason, the *unrelated* population may violate this assumption. Thus, comparisons of the magnitude and direction of the parental effects and interaction effects among populations are difficult to interpret and these effects are not reported.

Candidate gene mapping: Candidate gene selection emphasized structural genes of phenylpropanoid metabolism involved in monolignol synthesis, including phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), coumarate 3-hydroxylase (C3H), caffeate *O*-methyltransferase, 4-coumarate:CoA ligase (4CL), caffeoly CoA *O*-methyltransferase (CCoAOMT), and cinnamyl alcohol dehydrogenase (CAD). Also included were loblolly pine genes homologous to (1) laccase, a gene potentially involved in polymerization of lignin monomers; (2) three genes involved in supplying methyl groups for lignin biosynthesis via *S*-adenosyl methionine (SAM), including SAM synthetase (SAMS), *S*-adenosyl homocysteine hydrolase (SAHH), and glycine hydroxymethyltransferase; and (3) five genes encoding arabinogalactan proteins (AGPs). These AGPs are abundantly and differentially expressed in differentiating xylem (Loopstra and Sederoff 1995; Loopstra *et al.* 2000; Zhang *et al.* 2003) and may play critical roles in wood development. PCR amplification primers can be viewed at the Genetics website (http://www.genetics.org/supplemental).

Primer pairs were designed with CPrimer using individual loblolly pine expressed sequence tags (ESTs) or contig assemblies of EST sequences accessed through the National Center for Biotechnology Information web server. Contig assembly was done using Sequencher (Gene Codes, Ann Arbor, MI). For several genes, it was possible to distinguish different members of a gene family within a contig, which allowed the selective amplification and mapping of individual gene family members. PCR amplication was performed as described previously (Harry *et al.* 1998) except HotStarTaq DNA Polymerase (QIAGEN, Valencia, CA) and a 60° annealing temperature were used. Genotypic data for markers segregating in any of the loblolly pine reference populations were obtained primarily by denaturing gradient gel electrophoresis according to Temesgen *et al.* (2001). Genotypic data for SAHH were generated by pyrosequencing (http://www.pyrosequencing.com) with the oligonucleotide 5′-CGGCGAGTATCAAGTT-3′ following PCR amplification. For PAL-1, a segregating banding pattern was obtained after restriction digestion with *Nla*III (New England Biolabs, Beverly, MA). cDNA clones encoding an arabinogalactan-like protein (Pta3HZ) and CAD were mapped previously by RFLP analysis (Sewell *et al.* 1999).

RESULTS

Linkage map construction: The sex-averaged linkage map of the *verification* population consists of 103 markers distributed across 12 linkage groups (LG) of loblolly pine ($2n = 24$). The map spans 1305 cM, slightly larger than that of Sewell *et al.* (2000, 2002). Only one fully

informative marker on LG 8 was available. In this case, a sex-averaged map was not constructed and QTL analysis was performed on the individual parental maps. The linkage map of the *unrelated* population, which covers 890 cM, consists of 73 markers distributed across 10 of the 12 LGs. Marker order on the four parental maps and the consensus map are essentially identical to those published previously (Sewell *et al.* 1999, 2002; Brown *et al.* 2001).

QTL mapping: All QTL observed in the *detection* population using both the one-QTL and two-QTL models were reported in Sewell *et al.* (2000, Tables 4–7; 2002, Tables 3–4). QTL detected in the *verification* and *unrelated* populations can be viewed at the Genetics website at http://www.genetics.org/supplemental. Many of these QTL are independent estimations of the same QTL since many of these traits are highly associated (*i.e.*, the same trait measured from annual rings and the composite trait derived from these rings). Accordingly, unique QTL are defined as the subset of QTL influencing the same traits that map within ~15 cM of one another (Sewell *et al.* 2000, 2002). Further interpretations are based solely on these unique QTL.

Verification population: A total of 44 unique QTL were detected in the *verification* population using the one- and two-QTL models, including 10 QTL for earlywood *wsg*, 8 QTL for latewood *wsg*, 12 QTL for the percentage volume of latewood (%*lw*), 4 QTL for latewood *mfa*, 5 QTL for earlywood *cwc*, and 5 QTL for latewood *cwc*. No QTL were detected for earlywood *mfa*. With the exception of a QTL on LG 5, which accounted for as much as 15.9% of the phenotypic variation in latewood *wsg*, the percentage of variance explained by each QTL was generally small, ranging from 1.7 to 5.7%. These effects are two- to threefold smaller than those reported by Sewell *et al.* (2000, 2002) for the same traits and likely represent more accurate estimates of the true QTL effects owing to the larger segregating population analyzed.

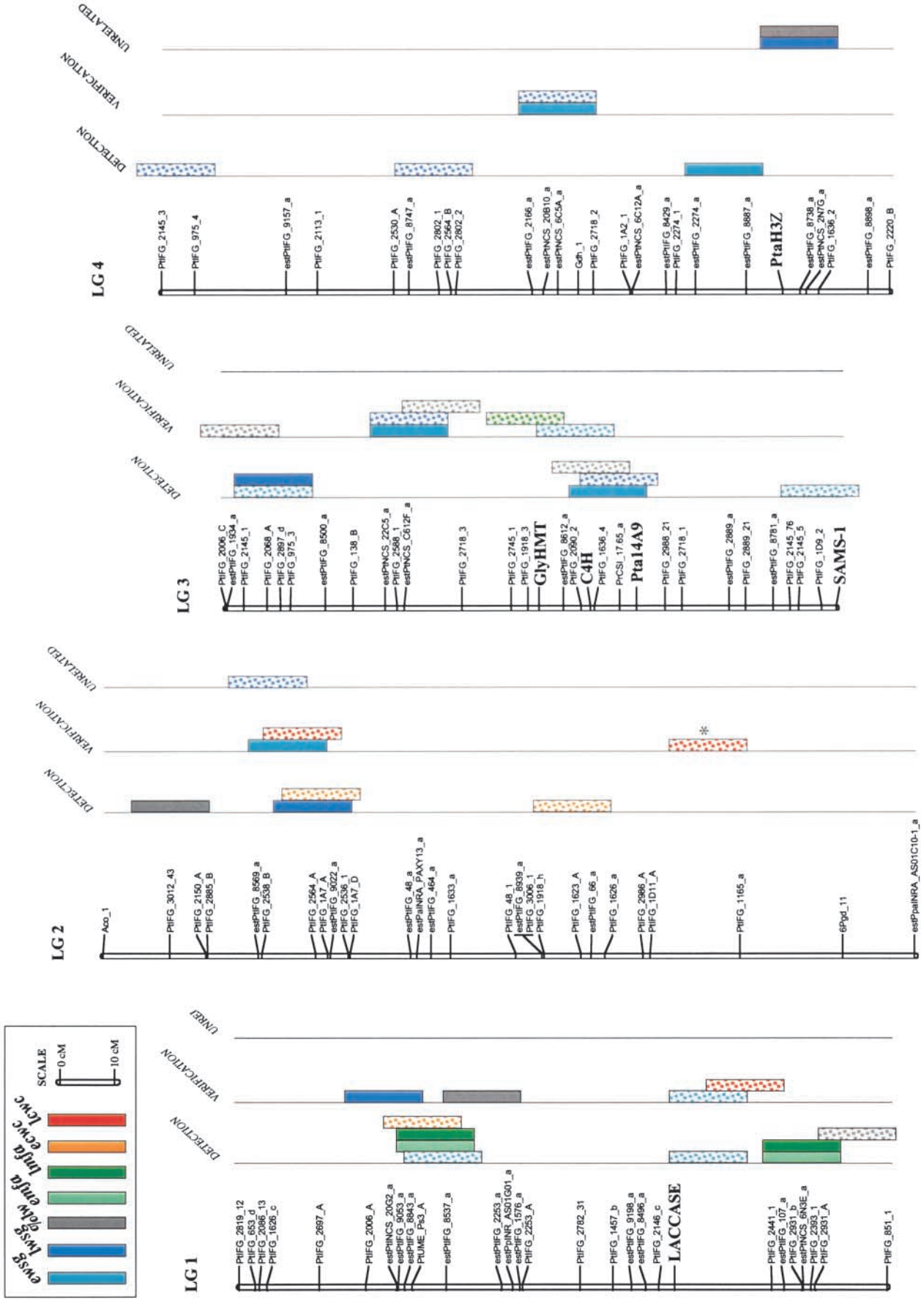
Unrelated population: A total of 12 unique QTL for physical wood properties were detected in the *unrelated* population using the one- and two-QTL models, including 5 QTL for latewood *wsg*, 5 QTL for %*lw*, and 2 QTL for latewood *mfa*. No QTL were detected for *wsg* or *mfa* of earlywood. The percentage of the phenotypic variance explained by each QTL was also small, ranging from 1.8 to 4.4%. Fewer QTL were detected in this pedigree in part due to less complete genome coverage. In addition, the grandparents of the *unrelated* population were not chosen on the basis of divergent phenotypic values and, as a result, fewer QTL may be segregating in the mapping population.

QTL verification: For comparative purposes, unique QTL in the three populations were placed into 15-cM regions of the consensus genetic map of loblolly pine on the basis of the position of homologous flanking markers (Figure 1). In some cases (*e.g.*, LG 7 of the *unrelated* population), this assignment is only approximate due to insufficient numbers of such markers. The 95% confidence interval of each QTL likely varies between experiments and in many cases will be considerably >15 cM; however, this bin size was chosen for illustrative purposes and in keeping with the definition of a unique QTL used here and previously. QTL verification (*i.e.*, the repeated detection of a QTL) was possible at three different levels: (1) across growing seasons, (2) between mapping populations of the same pedigree, and (3) between unrelated pedigrees.

Across growing seasons: Within each population, the colocation of QTL detected over multiple annual rings for a given trait represents a form of verification across a developmental gradient. For traits where QTL were estimated in more than one growing season (*e.g.*, *wsg* and %*lw* in all populations and *mfa* in the *detection* populations), 56 of 91 (62%) QTL were detected in more than one ring. For example, the majority of *wsg* QTL in the *verification* population were supported by both composite trait and individual ring analyses.

Between mapping populations of the same pedigree: Sewell *et al.* (2000, 2002) identified 61 unique QTL influencing physical and chemical wood properties in the *detection* population. Of the 44 QTL detected in the *verification* population, 12 (27%) are potentially repeated detections of the same QTL (Table 2 and Figure 1). For example, 6 QTL for earlywood *wsg* were found in similar locations in both populations on LGs 1, 3, 5 (2 QTL), 6, and 12. Despite the increased power of the larger *verification* population to detect QTL of small effect, fewer QTL than reported in the *detection* population were found. Several confounding factors may have contributed to these findings, including differences in both the number of annual rings sampled and the test sites for each of the two populations. Sewell *et al.* (2000) suggested that the onset of mature wood production might induce the expression of a new suite of QTL not detected in juvenile wood. The overall success of QTL

Figure 1.—Verification of QTL influencing wood properties in three populations of loblolly pine. Unique QTL for earlywood and latewood traits in each pedigree are presented as ~15-cM bars on the consensus genetic map at left. An evenly spaced subset of the markers shown on the consensus map was used for QTL analysis in each population. Shaded bars denote QTL verified by repeated detection across multiple growing seasons; stippled bars represent QTL observed or measured in only one growing season. Asterisks (*) show QTL whose position is estimated due to insufficient markers in common to allow more accurate placement on the consensus map. Candidate genes are capitalized and in boldface type. Genetic maps of the *verification* and *unrelated* populations will be submitted to TreeGenes, the forest tree genome database (http://dendrome.ucdavis.edu/TreeGenes).


LG 4
LG 3
LG 2
LG 1
UNRELATED
VERIFICATION
DETECTION
PtaH3Z
GlyHMT
C4H
Pta14A9
SAMS-1
LACCASE
SCALE
0 cM
10 cM
ewse
lwse
ofdw
emfa
lmfa
ecwc
lcwc

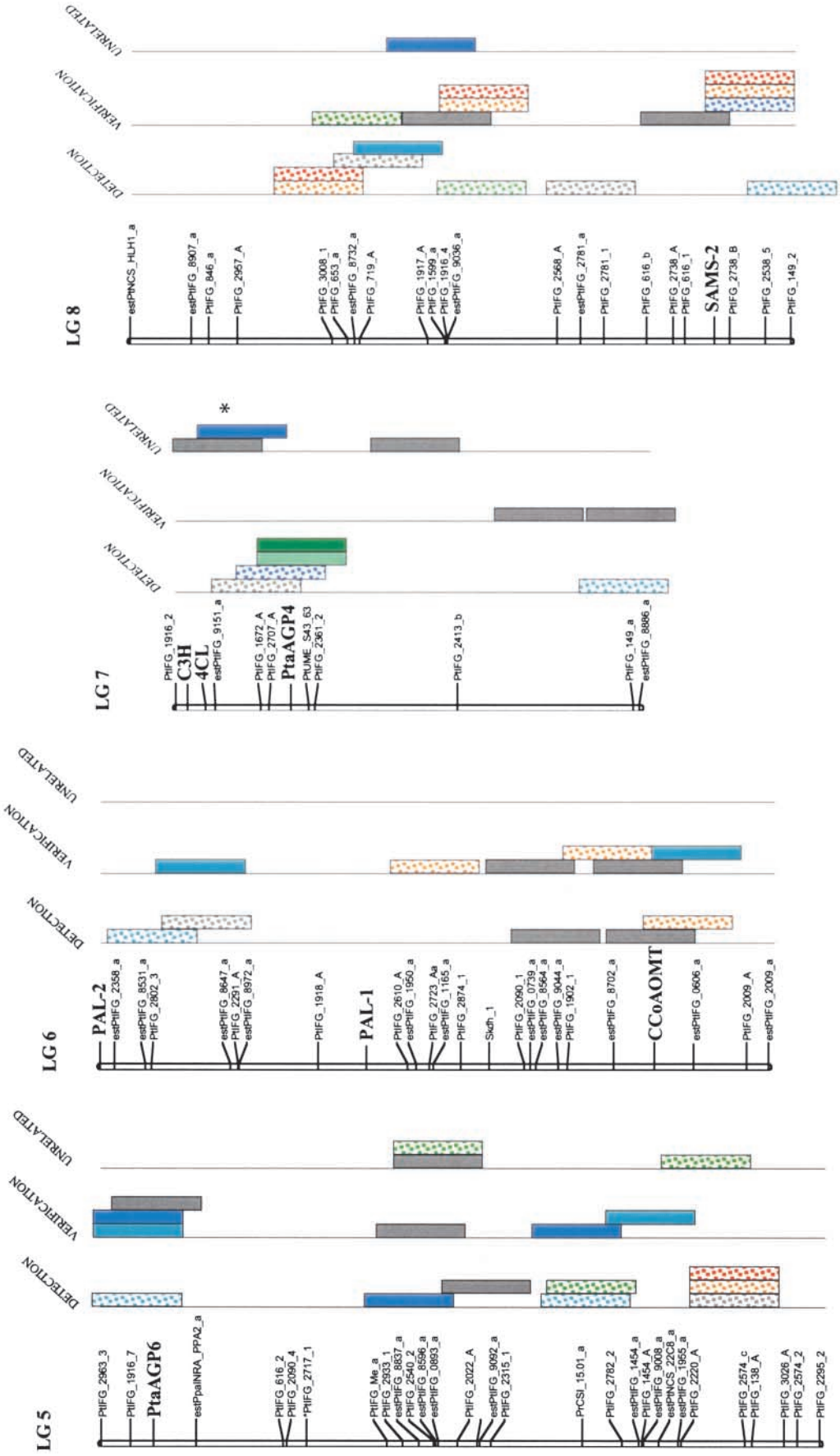


FIGURE 1.—*Continued.*

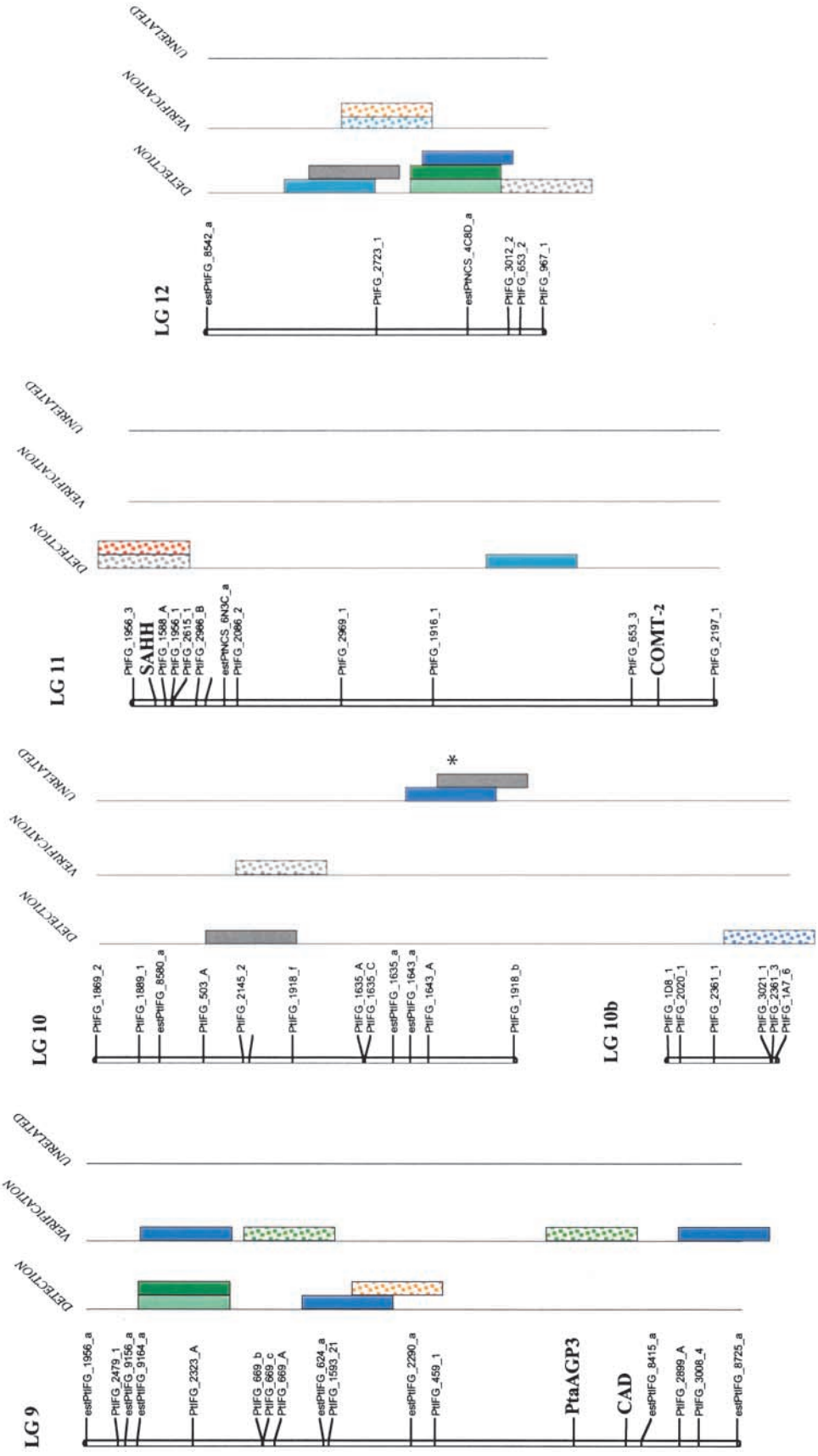


FIGURE 1.—*Continued.*

TABLE 2

Summary of QTL verified by repeated detection in loblolly pine

LG	Trait	Interval[a]	*P* value[b]		PVE[c]			Candidate genes
			Verification	*Unrelated*	*Detection*	*Verification*	*Unrelated*	
1	*ewsg*	2146_c:2441_1	0.004*		7.2	3.0		laccase
2	*lwsg*	2150_A:1A7_A		0.005*	5.4		2.3	
3	*ewsg*	2090_2:17.65_a	0.006*		6.6	3.2		C4H, GlyHMT, Pta14A9
5	*ewsg*	15.01_a:2220_A	0.003*		6.0	2.9		
5	*ewsg*	2963_3:2090_4	0.004**[d]		5.6	3.5[d]		AGP6
5	*%lw*	2933_1:15.01_a	0.002**[d]	0.006*	7.2	3.1[d]	2.0	
6	*ewsg*	2802_3:8972_a	0.0008**		6.6	3.5		PAL-2
6	*%lw*	2874_1:8702_a	0.007*[e]		11.0[e]	3.1[e]		
		8702_a:2009_a						CCoAOMT
6	*ecwc*	2874_1:8702_a	0.006*[d]		6.4	4.4[d]		
7	*lwsg*	1916_2:2361_2		0.0105*	5.9		2.4	C3H, 4CL, PtaAGP4
7	*%lw*	1916_2:2361_2		0.0006**	6.1		2.4	C3H, 4CL, PtaAGP4
8	*%lw*	719_A:1916_4	0.004*		5.8	1.8		
10	*%lw*	2145_2:1635_A	0.005*		8.7	2.5		
12	*ewsg*	8542_a:3012_2	0.015*		5.4	2.3		

[a] Marker interval on the consensus genetic map of loblolly pine. Markers that are underlined are common to the genetic maps of the populations compared. Markers not underlined denote interval boundaries inferred from homologous flanking markers.

[b] * and ** represent chromosome-wide significance at $P < 0.05$ and 0.01, respectively, except for marker-trait associations detected by the two-QTL model (see footnote [d] below).

[c] Percentage of the phenotypic variance explained by a QTL.

[d] A QTL detected only by the two-QTL model in the *verification* population with significance levels as in Sewell *et al.* (2000): *, $0.01 > P > 0.005$; **, $P < 0.005$. PVE refers in this case to that explained by the pair of QTL detected.

[e] QTL detected by the two-QTL model in both the *detection* and *verification* populations.

verification, therefore, is probably biased downward since not all QTL found in the *detection* pedigree, in particular those detected in rings 8–11, may be directly comparable to those detected in the juvenile wood cores of the *verification* population.

Between unrelated pedigrees: Of 12 (33%) unique QTL-influencing physical wood properties in the *unrelated* population, 4 mapped to similar locations in the *detection* population (Table 2 and Figure 1). These included latewood *wsg* QTL on LGs 2 and 7 and *%lw* QTL on LGs 5 and 7. The *unrelated* population was not phenotyped for *cwc* traits. It was rare to observe a QTL common to all populations with only a single *%lw* QTL on LG 5 being found. Its detection suggests that this QTL is affected less by genetic background (*e.g.*, lack of epistasis) and potentially by genotype × environment interactions than all other detected QTL.

Candidate gene mapping: The map positions of 18 candidate genes with known or putative roles in the biosynthesis of lignin or components of the cell wall are shown in Figure 1. Of particular interest were those candidate genes colocating with QTL verified by repeated detection (Table 2). A laccase mapped near a verified QTL on LG 1 influencing earlywood *wsg*. C4H, GlyHMT, and Pta14A9 on LG 3 and PtaAGP6 on LG 5 also mapped near verified QTL for earlywood *wsg*. CCoAOMT mapped to a region on LG 6 containing verified QTL influencing *%lw*. C3H, 4CL, and PtaAGP4 mapped to a region on LG 7 possibly influencing latewood *wsg* and *%lw*. Finally, a member of the SAM synthetase gene family (SAMS-2) mapped to LG 8 near a cluster of QTL affecting latewood *wsg* and *cwc*.

DISCUSSION

The size and expense of experiments to verify QTL has proven to be an obstacle to their widespread implementation. In a review of QTL mapping in forest trees (Sewell and Neale 2000), QTL verification was a component of only 2 of 20 experiments described (Wilcox *et al.* 1997; Frewen *et al.* 2000). QTL verification is essential to substantiate a biological basis for observed marker-trait associations, to provide precise estimates of the magnitude of QTL effects, and to predict whether a QTL will be expressed at a given developmental age or in a particular environment. Despite a number of confounding factors, these experiments lead to several conclusions regarding the genetic control of wood properties in loblolly pine.

Unlike agronomic crop species, which develop to maturity within a single season, forest trees are long lived, experiencing both seasonal cycles and maturation processes over decades. Half of the QTL influencing *wsg* and *%lw* were consistently detected over multiple growing seasons. QTL controlling *mfa* were equally stable when measured across multiple years. The structural

and regulatory genes underlying these QTL may be the primary determinants of the physical properties of juvenile wood whereas QTL detected in only a single year may represent physiological processes activated in response to biotic or abiotic variation. However, it is not known which, if any, of these QTL contribute to mature wood properties. Once these populations have grown sufficiently to ensure the production of mature wood, a second QTL analysis is necessary to determine the consistency of QTL expression at maturity.

The components of *wsg*, in general, were detected more consistently than those of *mfa* or *cwc* in both the *verification* and *unrelated* pedigrees. This may be a reflection of high heritabilities for *wsg* ($0.2 > h^2 > 1$; Zobel and Jett 1995), although estimates of h^2 for *mfa* and *cwc* are limited. Between the *detection* and *verification* populations, factors in addition to the problem of sampling different annual rings may also have contributed to inconsistent estimates of QTL. First, overlapping but not identical marker sets were used for construction of the genetic maps. Coupled with differences in family sizes, different recombination distances between markers common to both populations were observed in some cases, which may have obscured the orthologous relationship betweeen QTL. Second, statistical significance thresholds are not static but vary among linkage groups and experiments according to sample size, marker density, and the proportion and pattern of missing data, among other factors (Churchill and Doerge 1994). Although our preliminary analyses showed that the different statistical criteria gave rise to similar results, it is possible that at least some of the QTL of suggestive significance in the *detection* population may arise from type I error. Finally, the impact of the environment on QTL expression and detection remains to be addressed. Forest trees grow under conditions of great environmental heterogeneity, which impacts tree physiology and growth and the properties of wood (Zobel and Jett 1995). Inconsistent QTL detection between environments may reflect differing environmental influences on specific metabolic pathways in the formation of wood. Appropriately designed experiments deploying large amounts of clonal material over multiple test sites need to be performed.

The repeated detection of QTL in unrelated families was difficult. Given the outcrossing mating system of pine and other conifers, that is not surprising since both a QTL and its genetic marker will not be polymorphic in every family. (As a corollary, QTL must be identified in multiple families to account for all genomic regions affecting trait variation.) The populations used for QTL mapping in agronomic crops, such as F_2 intercrosses or recombinant inbred lines, are considerably more efficient for QTL detection since nearly all genetic markers and QTL segregate. These differences are apparent when comparing the extent of QTL verification between unrelated genotypes in agronomic crops and loblolly pine. For example, Moncada *et al.* (2001) reported that 11 of 25 (44%) QTL detected in an interspecific cross of rice were identified previously in similar positions. Lan and Paterson (2001) reported 9 of 17 (53%) QTL controlling plant size in *Brassica oleracea* were common between F_2 populations derived from two different intervarietal crosses, and 27% of QTL were common among three crosses. The existence of QTL that exert major effects on phenotypes, such as those controlling fruit size and shape in tomato (reviewed by Grandillo *et al.* 1999) and height and flowering time in maize (Lin *et al.* 1995), have further facilitated QTL verification in agronomic crops.

Successful QTL verification raises the prospect of marker-assisted selection (MAS). Strauss *et al.* (1992) discussed several serious obstacles to its implementation in forest tree breeding programs. While technical concerns, such as the cost of marker development, have largely vanished, practical considerations remain. The domestication of forest trees has begun only recently; thus, breeding programs contain a wide diversity of germplasm. Unlike crop plants, in which strong linkage disquilibrium (LD) is created by inbreeding or hybridization, breeding populations of forest trees show little LD, giving rise to inconsistent marker-trait associations among genotypes. Plantation site heterogeneity and the extended time and variable climatic conditions that trees experience before harvesting may result in unpredictable QTL effects. Nonetheless, MAS within families for juvenile wood property traits seems feasible, given the modest level of QTL verification between the *detection* and *verification* populations. One circumstance in which QTL mapping and MAS may be warranted is when clonal deployment over large areas is anticipated for a small number of highly valued families. Family selection will be possible under two scenarios: if the gene underlying a QTL is identified or if a genetic marker in linkage disequilibrium with the molecular polymorphism causing trait variation at the population level is discovered.

Isolating the gene underlying a QTL is an enormous undertaking even in species with small genomes (*e.g.*, Frary *et al.* 2000), and it is unlikely that map-based cloning will be used in pines [the C value of loblolly pine is 21–23 pg (Wakamiya *et al.* 1993)]. A population-based association approach using positional candidate genes is an alternate strategy. Our screen of genes involved in lignin biosynthesis or encoding arabinogalactan-like proteins provided plausible candidates controlling wood property traits. For example, C4H, C3H, 4CL, and CCoAOMT of monolignol synthesis may have influences on *wsg*. However, the populations used for this analysis are in strong linkage disequilibrium, and, as such, a very large region of DNA including many additional genes is implicated by these findings. Validation of these results is being undertaken in a natural population of loblolly pine, in which linkage disequilibrium

between any two sites is minimized due to historical recombination events. A successful association test of single nucleotide polymorphisms in these candidate genes with wood property phenotypes promises to enable family selection at the allele level, regardless of pedigree or family relationships.

The authors thank C. Dana Nelson and three anonymous reviewers for critical reading of the manuscript and Robert Saich for technical assistance. This research was funded in part by the United States Department of Agriculture National Research Initiative Plant Genome grant 96-35300-3719, National Science Foundation grant 9975806, and Department of Energy Agenda 2020 grant DE-AC05-96OR22464.

LITERATURE CITED

Beavis, W. D., 1994 The power and deceit of QTL experiments: lessons from comparative QTL studies, pp. 250–266 in *Proceedings of the 49th Annual Corn and Sorghum Industry Research Conference*. American Seed Trade Association, Washington, DC.

Bernacchi, D., T. Beck-Bunn, Y. D. Emmatty, Y. Eshed, S. Inai *et al.*, 1998 Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. Theor. Appl. Genet. **97:** 170–180.

Brown, G. R., E. E. Kadel, D. L. Bassoni, K. L. Kiehne, B. Temesgen *et al.*, 2001 Anchored reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomics. Genetics **159:** 799–809.

Churchill, G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. Genetics **138:** 963–971.

Cormier, R. T., K. H. Hong, R. B. Halberg, T. L. Hawkins, P. Richardson *et al.*, 1997 Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis. Nat. Genet. **17:** 88–91.

Davis, M., C. Elam, R. Evans, A. Wiselogel, N. Wheeler *et al.*, 1999 Application of pyrolysis molecular beam mass spectrometery for the determination of loblolly pine and hybrid poplar cell wall composition, pp. 1077–1082 in *TAPPI Pulping Conference*. TAPPI Press, Atlanta.

Devey, M. E., K. D. Jermstad, C. G. Tauer and D. B. Neale, 1991 Inheritance of RFLP loci in a loblolly pine three-generation pedigree. Theor. Appl. Genet. **83:** 238–242.

Devey, M. E., T. A. Fiddler, B.-H. Liu, S. J. Knapp and D. B. Neale, 1994 An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. Theor. Appl. Genet. **88:** 273–278.

Frary, A., T. C. Nesbitt, A. Frary, S. Grandillo, E. van der Knapp *et al.*, 2000 *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science **289:** 85–88.

Frewen, B. E., T. H. Chen, G. T. Howe, J. Davis, A. Rohde *et al.*, 2000 Quantitative trait loci and candidate gene mapping of bud set and bud flush in Populus. Genetics **154:** 837–845.

Grandillo, S., H. M. Ku and S. D. Tanksley, 1999 Identifying the loci responsible for natural variation in fruit size and shape in tomato. Theor. Appl. Genet. **99:** 978–987.

Groover, A., M. Devey, J. Lee, R. Megraw, T. Mitchell-Olds *et al.*, 1994 Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics **138:** 1293–1300.

Haley, C. S., S. A. Knott and J. M. Elsen, 1994 Mapping quantitative trait loci in crosses between outbred lines using least squares. Genetics **136:** 1195–1207.

Hardin, B., 2000 Rice breeding gets marker assists. Agric. Res. **48:** 11.

Harry, D. E., B. Temesgen and D. B. Neale, 1998 Codominant PCR-based markers for *Pinus taeda* developed from mapped cDNA clones. Theor. Appl. Genet. **97:** 327–336.

Knott, S. A., D. B. Neale, M. M. Sewell and C. S. Haley, 1997 Multiple marker mapping of quantitative trait loci in an outbred pedigree of loblolly pine. Theor. Appl. Genet. **94:** 810–820.

Lan, T.-H., and A. H. Paterson, 2001 Comparative mapping of QTLs determining the plant size of *Brassica oleracea*. Theor. Appl. Genet. **103:** 383–397.

Lander, E., and L. Kruglyak, 1995 Genetic dissection of complex traits: guidelines for intrepreting and reporting linkage results. Nat. Genet. **11:** 241–247.

Lin, Y.-R., K. F. Schertz and A. H. Paterson, 1995 Comparative analysis of QTLs affecting plant height and maturity across the Poaceae, in reference to an interspecific sorghum population. Genetics **141:** 391–411.

Loopstra, C. A., and R. R. Sederoff, 1995 Xylem-specific gene expression in loblolly pine. Plant Mol. Biol. **27:** 277–291.

Loopstra, C. A., J. D. Puryear and E.-G. No, 2000 Purification and cloning of an arabinogalactan-protein from xylem of loblolly pine. Planta **210:** 686–689.

Lynch, M., and B. Walsh, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.

Megraw, R. A., 1985 *Wood Quality Factors in Loblolly Pine*. TAPPI Press, Atlanta.

Melchinger, A. E., H. F. Utz and C. C. Schon, 1998 Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics **149:** 383–403.

Moncada, P., C. P. Martinez, J. Borrero, M. Chatel, H. Gauch, Jr. *et al.*, 2001 Quantitative trait loci for yield and yield components in an *Oryza sativa* × *Oryza rufipogon* BC_2F_2 population evaluated in an upland environment. Theor. Appl. Genet. **102:** 41–52.

Seaton, G., C. S. Haley, S. A. Knott, M. Kearsay and P. M. Visscher, 2002 QTL express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics **18:** 339–340.

Sewell, M. M., and D. B. Neale, 2000 Mapping quantitative traits in forest trees, pp. 407–424 in *Molecular Biology of Wood Plants*, edited by S. M. Jain and S. C. Minocha. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Sewell, M. M., B. K. Sherman and D. B. Neale, 1999 A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigree. Genetics **151:** 321–330.

Sewell, M. M., D. L. Bassoni, R. A. Megraw, N. C. Wheeler and D. B. Neale, 2000 Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). I. Physical wood properties. Theor. Appl. Genet. **101:** 1273–1281.

Sewell, M. M., M. F. Davis, G. A. Tuskan, N. C. Wheeler, C. C. Elam *et al.*, 2002 Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). II. Chemical wood properties. Theor. Appl. Genet. **104:** 214–222.

Stam, P., 1993 Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J. **3:** 739–744.

Steinmetz, L. M., H. Sinha, D. R. Richards, J. I. Spiegelman, P. J. Oefner *et al.*, 2002 Dissecting the architecture of a quantitative trait locus in yeast. Nature **416:** 326–330.

Strauss, S. H., R. Lande and G. Namkoong, 1992 Limitations of molecular-marker-aided selection in forest tree breeding. Can. J. For. Sci. **22:** 1050–1061.

Temesgen, B., G. R. Brown, D. E. Harry, C. S. Kinlaw, M. M. Sewell *et al.*, 2001 Genetic mapping of expressed sequence tag polymorphism (ESTP) markers in loblolly pine (*Pinus taeda* L.). Theor. Appl. Genet. **102:** 664–675.

Utz, H. F., A. E. Melchinger and C. C. Schon, 2000 Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics **154:** 1839–1849.

Wakamiya, I. R., R. J. Newton, J. S. Johnston and H. J. Price, 1993 Genome size and environmental factors in the genus *Pinus*. Am. J. Bot. **80:** 1235–1241.

Wilcox, P. L., T. E. Richardson and S. D. Carson, 1997 Nature of quantitative trait variation in *Pinus radiata*: insights from QTL detection experiments, pp. 304–312 in *Proceedings of IUFRO '97: Genetics of Radiata Pine*, edited by R. D. Burdon and J. M. Moore. FRI Bulletin No. 203, Rotorua, New Zealand, December 1–4, 1997.

Zhang, Y., G. Brown, R. Whetten, C. A. Loopstra, D. Neale *et al.*, 2003 An arabinogalactan protein associated with secondary cell wall formation in differentiating xylem of loblolly pine. Plant Mol. Biol. **52:** 91–102.

Zobel, B. J., and J. B. Jett, 1995 *Genetics of Wood Production*. Springer-Verlag, Berlin/Heidelberg/New York.

Communicating editor: A. H. D. Brown