Theor Appl Genet (2004) 109: 451–463
DOI 10.1007/s00122-004-1653-5

ORIGINAL PAPER

T. M. Yin · S. P. DiFazio · L. E. Gunter ·
D. Riemenschneider · G. A. Tuskan

Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map

Received: 26 August 2003 / Accepted: 25 February 2004 / Published online: 27 May 2004

Abstract We report the most complete genetic map to have been constructed for the genus *Populus*. This map includes 544 markers mapped onto 19 linkage groups, equivalent to the *Populus* chromosome number, with all markers displaying internally consistent linkage patterns. We estimate the genome length to be between 2,300 and 2,500 cM, based both on the observed number of crossovers in the maternal haplotypes, as well as the total observed map length. Genome coverage was estimated to be greater than 99.9% at 20 cM per marker. We did not detect obvious recombination repression in the maternal tree (a hybrid of *Populus trichocarpa* Hooker × *P. deltoides* Marsh.) compared to the paternal tree (pure *P. deltoides*). Finally, most markers exhibiting segregation distortion were derived from the donor parent in this backcross, and generally occurred in large contiguous blocks on two linkage groups. We hypothesize that divergent selection has occurred on chromosomal scales among the parental species used to create this pedigree, and explore the evolutionary implications of this observation. This genetic linkage map provides the most comprehensive view of the *Populus* genome reported to date and will prove invaluable for future inquiries into the structural and functional genomics, evolutionary biology, and genetic improvement of this ecologically important model species.

Communicated by D.B. Neale

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00122-004-1653-5

T. M. Yin · S. P. DiFazio (✉) · L. E. Gunter · G. A. Tuskan
Environmental Sciences Division, Oak Ridge National Laboratory,
P.O. Box 2008 Oak Ridge, TN 37831-6422, USA
e-mail: difazios@ornl.gov
Tel.: +1-865-5747844
Fax: +1-865-5769939

T. M. Yin
The Key Laboratory of Forest Tree Genetic Engineering of the State Forestry Administration, Nanjing Forestry University, Nanjing, JS 210037, China

D. Riemenschneider
USDA Forest Service, North Central Forest Experiment Station, Forestry Sciences Laboratory,
5985 Highway K,
Rhinelander, WI 54501, USA

Introduction

The genus *Populus* possesses many characteristics that are conducive to functional genomic studies, and is therefore widely accepted as a model system in tree genome research (Bradshaw et al. 2000; Taylor 2002; Wullschleger et al. 2002a). Consequently the collection of *Populus* genomic resources has grown rapidly in recent years, culminating in the sequencing of the entire *Populus* genome by the Joint Genome Institute of the U.S. Department of Energy (Wullschleger et al. 2002b; Bhalerao et al. 2003). In addition, as of 26 December 2003, 154,746 *Populus* EST sequences had been deposited in GenBank and a 10× BAC library has been fingerprinted, end-sequenced and assembled into approximately 4,000 scaffolds, thus providing a physical map on which the genomic information can be assembled (M. Marra, BC Genome Sciences Center, personal communication, February 2003). A dense genetic map will prove to be an invaluable resource for coalescing this genomic information into chromosomal units, thus providing a platform for functional and evolutionary comparisons with other species.

The first investigation of the *Populus* genome was made in 1921, in which the haploid chromosome number was erroneously reported as four (Graf 1921). By 1924, it became clear that the base chromosome number in *Populus* was 19, based on observations in seven species (Harrison 1924). Since then, examination by various scientists has revealed that all *Populus* species exist in the diploid form with $2n$=38 (Smith 1943), with occasional cases of triploid or tetraploid genets (Einspahr et al. 1963; Bradshaw and Stettler 1993). Based on cytology studies,

Van Dillewijin (1940) hypothesized that the ancestral chromosome number of *Populus* was eight. However, because *Populus* chromosomes are mostly small and lacking in distinctive morphological features (Smith 1943), there is scant information on chiasma frequencies and chromosomal dynamics to substantiate this claim.

Genetic linkage mapping provides a powerful means for understanding genomic structure, as well as many other promising applications in tree improvement (Dinus and Tuskan 1997; Neale et al. 2002). The first linkage map in *Populus* was constructed by Bradshaw et al. (1994) with 50% coverage of the genome. Using AFLP and RAPD markers, three other studies have achieved 19 main linkage groups in *Populus* (Wu et al. 2000; Yin et al. 2001, 2002). Recently, Cervera et al. (2001) reported dense genetic maps for *P. deltoides*, *P. trichocarpa* and *P. nigra*, utilizing AFLP and more than 100 simple sequence repeat (SSR) markers. These maps provide an excellent starting point for comparative mapping.

Populus occurs throughout the northern hemisphere, and extensive hybridization occurs in large, persistent zones of overlap between *Populus* species (Eckenwalder 1984; Rood et al. 1986). The genetic and ecological factors maintaining such hybrid zones have garnered much attention from evolutionary biologists, and the relative roles of gene flow and endogenous and exogenous selection against hybrids have been extensively debated (Anderson 1949; Barton and Hewitt 1985; Orr 1996; Arnold 1997; Rieseberg and Carney 1998). We present here a mapping pedigree derived from a hybrid *P. trichocarpa* × *P. deltoides* female crossed to an alternate *P. deltoides* male genotype. This family structure provides an opportunity to examine the chromosomal dynamics associated with the early stages of introgression between these two divergent species (Rieseberg et al. 2000). Therefore, one of the objectives of this study was to search for evidence of recombination repression due to the heterogeneous genetic background of the hybrid. Additionally, Martinsen et al. (2001) found that introgression from a natural hybrid zone of *P. angustifolia* James and *P. fremontii* S. was directional, and only certain genomic regions had penetrated from the *P. fremontii* zone deep into the *P. angustifolia* zone, suggesting that the potential for introgression varies across the genome. Our pedigree involves species from the same sections of the genus as those studied by Martinsen et al. (2001) (*Tacamahaca* and *Aigeiros*), and our linkage map provides some potential insights into the chromosomal signatures of evolutionary divergence between these sections.

Materials and methods

Plant material and DNA preparation

The mapping pedigree (family 13) in this study was an interspecific backcross consisting of 180 offspring. The female, '52-225' was derived from a hybrid cross between clones '93-968' (*P. trichocarpa*) × 'ILL-101' (*P. deltoides*). The backcross male, 'D109', is a pure *P. deltoides* selected within a half-sib progeny test from an open-pollinated seed collection from a tree in southeastern Minnesota. Total DNA was extracted from young leaves sampled from seedlings in the greenhouse using the DNeasy plant miniprep kit (QIAGEN, Valencia, Calif., USA). DNA concentration was estimated on the FLUOROSKAN Ascent (Thermo Labsystems, Vantaa, Finland).

SSR analysis and marker nomenclature

The SSR primers used in this study came from four sources: (1) the *Populus* Molecular Genetics Cooperative ("P_" prefix), (2) Oak Ridge National Laboratory ("O_" prefix) (Tuskan et al. 2004), (3) the Center for Plant Breeding and Reproduction Research (Van der Schoot et al. 2000), "W_" prefix) and (4) the University of Ghent ("R_" prefix). Initially, 108 SSRs were selected based on a framework SSR map from another pedigree (family 331; Tuskan et al. 2004). SSRs were analyzed with fluorescent dye-labeled primers (Hex and Fam). PCR reactions were carried out in a total volume of 15 μl containing 25 ng of template DNA, 20 ng forward and reverse oligonucleotide primers (Operon Technologies, Alameda, Calif., USA); 200 μM of each dNTP, 0.5 U *Taq* DNA polymerase (New England Biolabs, Beverly, Mass., USA), 1.5 μl 10× buffer containing 100 μM Tris-HCl, pH 8.3, 500 mM KCl, 20 mM $MgCl_2$ and 10.0 g l^{-1} BSA. PCR was conducted in a Perkin-Elmer Cetus thermocycler 9700 (Applied Biosystems, Foster City, Calif., USA) with 30 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min combined with a final extension at 72°C for 7 min. PCR products were detected on an ABI 3700 sequencer using the standard genotyping module. Before each analysis, the appropriate dilution factor was determined for each primer set using a subsample of the amplifications. Two to four primer pairs with different product sizes and/or dye colors were mixed in appropriate ratios and diluted 1:10 with loading buffer [91% deionized formamide, 9% internal standard GeneScan 450ROX (Applied Biosystems)], then denatured at 95°C for 5 min followed by rapid cooling on ice.

AFLP analysis and marker nomenclature

The AFLP procedure was performed following the approach of Vos et al. (1995), with the following modifications. Pre-amplification reactions (15 μl) were performed for 3 μl of the diluted DNA template using 20 pmol each of a pair of AFLP primers (Operon Technologies) with no selective 3′ nucleotides on the 'E' primer and one "C" selective 3′ nucleotide on the 'M' primer. Reaction conditions were otherwise identical to those described for SSRs. PCR was conducted with 20 cycles of 94°C for 30 s, 56°C for 60 s and 72°C for 60 s. The pre-amplified products were diluted 1:30 as DNA template for selective amplification. Selective amplification was carried out in a volume of 15 μl containing 3 μl

diluted pre-amplification product, 0.5 pmol 'E' primer with two selective nucleotides (Hex labeled) and 5 pmol 'M' primer with three selective nucleotides (Operon Technologies), 200 μM each dNTP, 0.5 U *Taq* DNA polymerase (Promega, Madison, Wis., USA), 1.5 μl 10× buffer (100 μM Tris-HCl, pH 8.3, 500 mM KCl, 20 mM $MgCl_2$), 10.0 g l^{-1} BSA and 1% (v/v) deionized formamide. Thermocycling conditions for selective amplification were 12 cycles of 94°C for 30 s, 65°C for 30 s decreasing by 0.7°C per cycle, and 72°C for 60 s, followed by 23 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 60 s.

Twenty-four AFLP primer combinations were selected based on previous performance in a different pedigree for which 256 AFLP primer combinations were initially screened (Yin et al. 2002). PCR products were separated on an ABI 3700 sequencer following the same procedure as for SSR analysis, except for 14 primer combinations that yielded fragments greater than 500 bp in size, which were analyzed using the X-Rhodamine MapMakers 1000 Standard (BioVentures, Murfreesboro, Tenn., USA) with a run voltage of 5,000. All AFLP markers were named using a code for each primer combination (Electronic Supplementary Material, Table 1), followed by sequential numbers for scored bands, beginning with the lowest molecular weight.

GeneScan and Genotyper software (Applied Biosystems) were used to extract data and score the traces, and the resulting data tables were further processed by PERL scripts (http://www.esd.ornl.gov/PGG/scripts.htm) to detect null alleles, anomalous alleles, aneuploidy, discrepancies in repeated samples, segregation distortion, and to infer parental origins of alleles. Discrepancies in the data were corrected by returning to the original traces and, in some cases, repeated genotyping.

Map construction

The linkage map was constructed using MapMaker version 3.0 (Lander et al. 1987) following the two-way pseudo-testcross mapping strategy (Grattapaglia and Sederoff 1994), generating a separate map for each of the parents. The data sets of the dominant markers were duplicated and both alleles of the codominant makers were included to allow the detection of linkage of markers in repulsion phase. Linkage groups were assigned with thresholds for a minimum LOD score of 10.0 and a maximal recombination fraction (r) of 0.30. For each assigned linkage group, the "LOD table" command was used to identify markers in bins (map distance equal to 0.0 cM). The marker with the highest LOD score from the bin was designated a 'framework' marker, and other markers in the bin were designated 'alternative' markers. The markers were ordered using the "Order" command iteratively with a default LOD of 3.0. The first-sequence order was confirmed using the "Ripple" command permuting five markers at a time. All markers that were not ordered in the first pass were placed again using the "Try" command. New markers were placed at appropriate positions if LOD scores descended toward both ends of the linkage group from the insert position. If the change in LOD was less than 1.5 for adjacent markers, the unplaced markers were placed in a bin with the closest framework marker, and designated 'alternative' markers. Maps were drawn with the program MapChart 2.1 (Voorrips 2002).

Tests for recombination repression

We tested for recombination repression due to the hybrid genetic background in the female parent by comparing the observed and estimated genome length with that of the *P. deltoides* male parent. The relationship between the number of observed chiasmata (μ) in a diploid cell, recombination frequency (r) and genetic distance in Haldane units (M) were given by Haldane (1919) as follows:

$$r = (1 - e^{-2M})/2 \quad (1)$$

$$\mu = -\ln(1 - 2r) \quad (2)$$

$$M = \ln(1 - 2r)/2 \quad (3)$$

Thus, the genome length is $\mu/2$, which can be determined directly by counting the chiasmata observed during meiosis. We estimated this value by examining the linkage phase of alleles in the progeny. A change in linkage phase along a linkage group was counted as a chiasma, and the total genome length was calculated from the average number of chiasmata among the progeny. In addition, we evaluated the effects of using alternative markers rather than framework markers by resampling loci without replacement and allowing only one locus per framework position. These calculations were repeated 1,000 times each for a range of locus numbers, sampling either from framework markers only, or from all markers. Programs for performing these analyses are available at http://www.esd.ornl.gov/PGG/scripts.htm.

We calculated observed genome length simply as $G=\sum G_I$, where G_I is the total genetic distance of linkage group I. We also calculated the estimated genome length using the method of Hulbert et al. (1988), based on partial linkage data, assuming a random distribution of markers, m,

$$G = m(m - 1)X/K \quad (4)$$

where X is the map distance corresponding to the LOD threshold Z for declaring linkage, and K is the number of marker pairs having LOD values at or above Z. We estimated the genome length for a range of LOD criteria. In addition, we analyzed the effects of the number of mapped markers on length estimates by resampling loci without replacement and performing the calculations for

1,000 iterations. This facilitated comparison between the female and male maps, which had substantially different numbers of markers.

Genome coverage was estimated by the function given by Lange and Boehnke (1982), assuming a random marker distribution.

$$c = 1 - e^{-2md/\hat{L}} \quad (5)$$

where c is the proportion of the genome within d cM of a marker, $\hat{L}$ is the estimated genome length and m is the number of markers.

Chromosomal-scale segregation distortion

Departure from Mendelian expectations for the segregation of individual markers was assessed using χ^2-tests. We graphically represented species-specific segregation distortion for chromosomal segments by plotting the difference between the observed and expected number of progeny that inherited chromosomal regions from *P. trichocarpa*, as inferred by P_1 genotypes and linkage phases. We assumed that regions between markers of different origin were split equally between each species. The total size of *P. trichocarpa* (B_{Ti}) and *P. deltoides* (B_{Di}) segments in the genome were estimated as

$$B_{\mathrm{Ti}} = \sum IT_{ij} + \frac{1}{2}\sum ITD_{ik} \quad (6)$$

and

$$B_{\mathrm{Di}} = \sum ID_{il} + \frac{1}{2}\sum ITD_{ik} \quad (7)$$

where IT_{ij} is the jth interval formed by markers from *P. trichocarpa* on linkage group I, ID_{il} is the lth interval formed by markers from *P. deltoides* on linkage group I, ITD_{ik} is the kth interval between markers of different origin on linkage group I. The size of the distorted regions was calculated in an analogous manner. A paired t-test was used to test whether the proportion of alleles or intervals from *P. trichocarpa* was equal to that from *P. deltoides* with $df=n-1$, where n is the number of individuals.

Marker distribution analysis

Marker distribution among linkage groups was evaluated by comparing the marker density with expectations under the Poisson distribution using the method described by Remington et al. (1999). Under the assumption of equal marker density for all linkage groups, the expected marker number λ_I in linkage group I would be a sample from a Poisson distribution, $\lambda_I = mL_I/\sum_I L_I$, where m is the total number of markers and L_I is the observed map length plus two times the size of the average interval in linkage group I. The probabilities $P(m_I \leq \lambda_I)$ and $P(m_I \geq \lambda_I)$ were evaluated under the cumulative Poisson distribution using a one-tailed test at $\alpha \leq 0.05$ and $\alpha \leq 0.01$. Since the SSR markers were selected based on preliminary linkage information, only AFLP markers were used for marker density comparisons.

We also examined clustering of both AFLP and SSR markers on a finer scale to identify gaps in the coverage of the current map. We evaluated marker distribution along each linkage group by examining markers in windows of variable sizes. Window boundaries were defined by a change in spacing from clustering (interval size less than the average interval size for the entire map) to dispersion (interval size greater than the average interval). The number of markers in each window was compared to the null expectation for evenly dispersed markers under a cumulative Poisson distribution using a one-tailed test with $\alpha \leq 0.05$ and $\alpha \leq 0.01$.

Results

Marker analysis

Of the 110 SSR primer pairs selected based on preliminary linkage analyses, four (3.7%) did not amplify, 14 (12.7%) did not segregate, and 92 (83.6%) generated 119 segregating loci. Among these loci, 56 (47.5%) were maternally informative [1:1], eight (7.5%) were paternally informative [1:1], 49 (41%) were fully informative [1:1:1:1] and six (4.2%) were intercross informative [3:1] loci. The number of loci amplified by each primer set was as follows: 77 (83.7%) produced one locus, nine (9.8%) two loci, three (3.3%) three loci, two (2.2%) four loci and one (1.1%) seven loci. SSR markers were distributed on all linkage groups. Based on multiple loci with more than two alleles per locus, nine of the original 180 progeny were identified as possible aneuploids or triploids and excluded from subsequent analyses.

The twenty-four AFLP primer combinations yielded a total of 653 segregating loci, an average of 27 loci per primer combination. There was considerable variation in the number of polymorphic AFLP markers revealed by different primer combinations, ranging from 11 to 51 markers. Among the AFLP markers, 451 (69%) were maternally informative, 118 (18%) were paternally informative, and 84 (13%) were intercross markers.

Linkage map construction

All intercross informative AFLP and SSR markers (segregating 3:1) were excluded from map construction, because these markers provide little information for linkage analysis (Maliepaard et al. 1997). Thus, linkage analysis in the hybrid female parent was based on 556 markers, including 451 AFLPs and 105 SSRs. We did not exclude markers with significant departures from Mende-

lian segregation because we were interested in patterns of segregation distortion in the genome and did not want to bias the data set. Using a LOD threshold of 10.0, 544 markers were initially assigned to 20 groups, leaving 12 ungrouped AFLP markers. Linkage group *X* separated into two groups under these criteria, but these groups could be joined at a LOD threshold of 5.98. Therefore, 544 markers were mapped with high confidence to 19 linkage groups. Linkage groups ranged from 73 to 262.5 cM in size (Fig. 1).

The 12 ungrouped AFLP markers were all seriously distorted due to an excess of visible alleles, with χ^2 values ranging from 34.5 to 72. In contrast, the maximum χ^2 value of mapped markers was 19.6. Furthermore, linkage analysis indicated that each of these 12 markers was loosely linked with markers in multiple linkage groups, suggesting that these markers might be genetically heterogeneous due to comigration of bands from multiple loci. These markers were therefore excluded from further consideration.

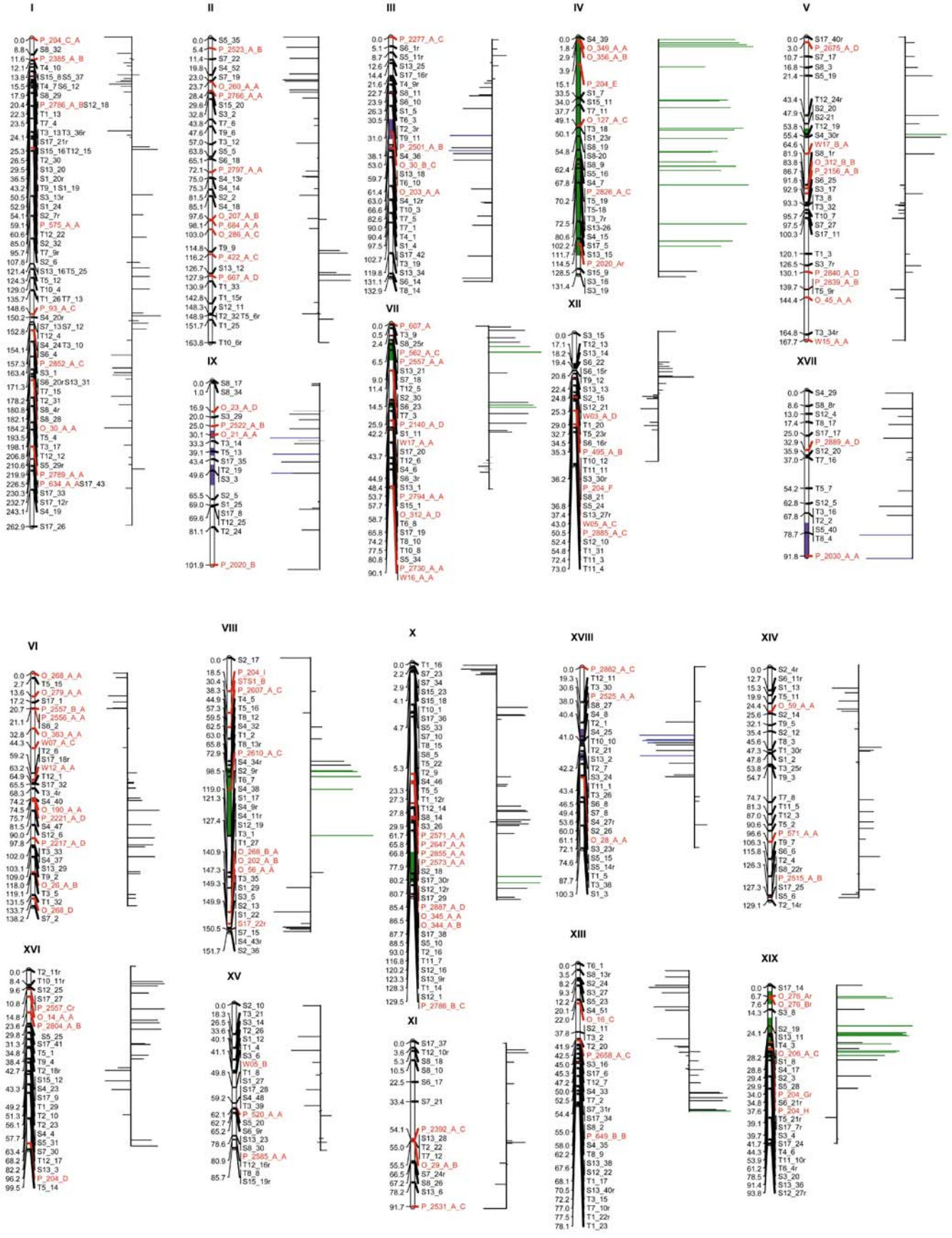
The framework map consisted of 445 markers. In addition, 99 alternative markers mapped in bins with a framework marker. The average distance between the alternative markers and the closest framework marker was 0.56cM (range 0-2.42cM). Chiasmata observed among randomly sampled loci were equivalent to those estimated from framework markers only, demonstrating that the framework map was an accurate representation of the full data set (Fig. 2).

A total of 175 markers were available for mapping in the paternal tree, and 158 of these could be mapped. The paternal map consisted of 33 linkage groups representing 1,046.2 cM. These groups were aligned with the 19 linkage groups from the female map using shared codominant SSR markers. There was excellent synteny between the male and female linkage groups, with full conservation of order and high correspondence in distance between markers (r=0.97, P≤0.001; paired t=−1.39, P=0.19, for 15 pairs of markers) (Electronic Supplementary MaterialFig. S1).

Table 1 Marker distribution and derivation for the maternal map. Clustered and dispersed regions refer to AFLP markers only, but species inheritance data are for all species. The map length is the observed map length (*G*) plus two times the size of the average interval in the linkage group. The probabilites for the expected AFLPs were evaluated under the cumulative Poisson distribution. The clustered regions are the number of regions showing significant clustering of AFLP within linkage groups. See Electronic Supplementary Material (Table 2) and Methods for details. Cumulative percentages of the linkage group encompassed by clustered markers are given. The dispersed regions are given as the number of regions showing significant dispersion (i.e., greater-than-average spacing) of AFLP within linkage groups. See Electronic Supplementary Material (Table 2) and Methods for details. The markers inherited from each species gives the mean number of markers observed in maternal haplotypes that were derived from *Populus trichocarpa* (*T*) or *P. deltoides* (*D*), for all 171 progeny. The blocks inherited from each species are represented by the mean size (*G*) of linkage blocks observed among the progeny. Linkage blocks are defined as regions in which adjacent markers are derived from the same species

Linkage group	Map length	Expected AFLP	Observed AFLP	Clustered regions		Dispersed regions		Markers inherited from each species		Blocks inherited from each species	
				Number	Percent	Number	Percent	D	T	D	T
I	270.58	47.34	56	1	7.9	1	23.2	24.18	23.77	123.66	124.73
II	173.73	30.39	24	4	5.7	0	-	15.90	15.02	80.34	76.19
III	142.07	24.85	25	1	7.1	0	-	13.13	11.63	67.80	58.49
IV	140.46	24.57	21	2	3.7	0	-	7.76	13.00***	45.26	78.38***
V	179.27	31.36	21*	2	11.9	0	-	12.37	13.01	75.67	81.71
VI	147.41	25.79	18	3	2.7	0	-	12.15	12.91	63.57	67.92
VII	96.31	16.85	20	2	3.8	1	30.7	8.59	10.15	39.48	46.55
VIII	160.62	28.10	25	3	10	1	37.8	9.11	9.74	66.60	73.45
IX	113.89	19.92	13	2	5.4	0	-	8.02	6.77	52.17	43.38
X	137.40	24.04	31	4	9.2	1	24.3	11.58	12.84	54.24	64.59
XI	103.93	18.18	12	2	2.3	0	-	6.33	6.52	42.59	43.14
XII	78.41	13.72	22*	2	7.3	2	39.1	10.25	10.39	34.56	34.58
XIII	83.31	14.57	27**	1	1.4	1	20.2	12.77	12.80	38.95	36.59
XIV	139.03	24.32	23	2	2.9	0	-	11.41	11.37	60.82	61.86
XV	93.49	16.36	19	1	1.2	0	-	7.05	6.74	42.36	38.98
XVI	107.46	18.80	21	2	3.0	0	-	9.76	10.87	45.25	50.14
XVII	104.04	18.20	13	1	4.5	0	-	6.88	5.76	48.47	38.57*
XVIII	108.02	18.90	23	3	12.5	1	37.9	9.08	7.67	49.34	45.31
XIX	101.30	17.72	20	2	2.8	1	50.2	8.30	11.44**	39.35	49.68*
Total	2480.72	434.00	434	40	5.5	9	32.9	204.67	212.47	1070.55	1114.32

*P≤0.05; **P≤0.01; ***P≤0.001


I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
XVI
XVII
XVIII
XIX
Scale for distortion:
Scale: 10 observations

◀ **Fig. 1** Genetic linkage map for female '52-225', a *Populus trichocarpa* × *P. deltoides* F_1 hybrid, as determined from 171 progeny of family 13, a backcross to *P. deltoides* male 'D109'. Markers in *black* are AFLP, and markers in *red* are framework microsatellites. Colored chromosome segments are between markers that showed significant segregation distortion: *green* excess of *P. trichocarpa* alleles, *blue* excess of *P. deltoides* alleles. Adjacent bar graphs depict segregation distortion for each marker. The *Y axis* is distance in centimorgans and the *X axis* is the difference between the observed and expected number of progeny with *P. trichocarpa* alleles for each marker. Markers with significant segregation distortion at $P \leq 0.05$ are indicated by *colored bars*

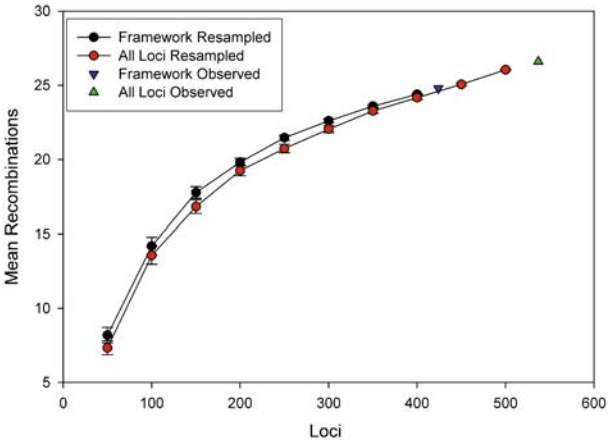


Fig. 2 Recombinations averaged over all progeny for different numbers of loci sampled 1,000 times without replacement. One set of simulations used framework markers only, and the other sampled from all markers, with a maximum of one marker sampled per position. *Bars* represent one standard deviation

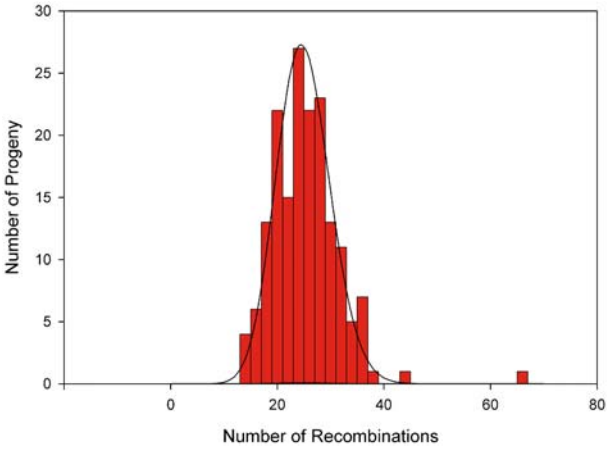


Fig. 3 Distribution of the number of inferred recombinations for 171 progeny of family 13. Recombinations were inferred when adjacent markers on linkage groups originated from different grandparents. The theoretical distribution is a Poisson with a mean of 25

Map length and coverage

The combined length of the 19 linkage groups of the maternal parent was 2,313.9 cM with the error detection function of Mapmaker enabled, and 2564.3 cM with error detection off. The average number of recombinations observed in the progeny was 24.789, which corresponds to a genome size estimate of 2478.9 cM (Eqs. 1, 2 and 3). These crossovers theoretically follow a Poisson distribution with $\lambda = \bar{u}$, where $\bar{u}$ is the average number of chiasmata. Since $\bar{u} = 24.866$ in this study, the distribution approaches a normal distribution of $\mu(\bar{u}, \sqrt{\bar{u}})$ (Fig. 3).

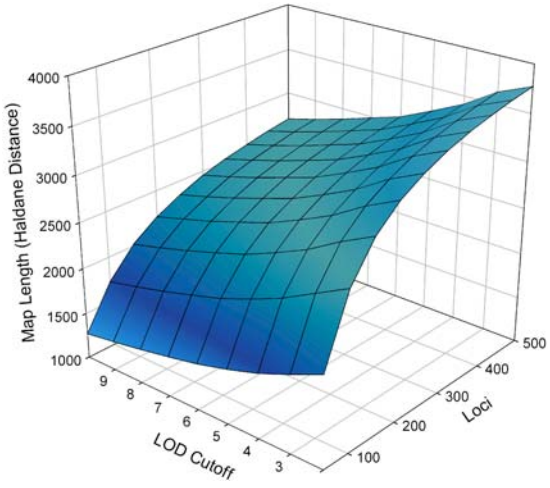


Fig. 4 Estimates of total map length for female '52-225' with different LOD cutoffs for determining pairwise linkage, and different numbers of loci sampled randomly without replacement from the entire data set. Based on 1,000 iterations

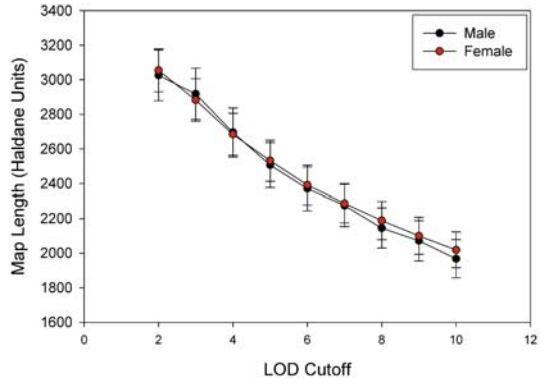


Fig. 5 Comparison of total map length estimates using the method of Hulbert et al. (1988) for female '52-225' and male 'D109' using 150 loci randomly sampled without replacement, and a range of LOD score cutoffs for determining pairwise linkage. *Bars* represent one standard deviation based on 1,000 iterations

Based on resampling loci without replacement, the genome length estimates varied greatly depending on the number of loci and the LOD cutoff used in the calculation (Fig. 4). The proper LOD cutoff for a given number of loci may be empirically determined based on the observed map length. For example, if the genome length is estimated to be 2,500 cM, the LOD cutoff would be three if 100 loci are involved and five if 150 loci are used for the estimate. Furthermore, our simulation demonstrates that at least 100 loci should be used to derive a proper genome length estimate for *Populus*. Thus, relying on the relationship between LOD cutoff, the number of loci, and the estimated map length, alternate genetic maps can be resolved among different studies. In our study, based on resampling 150

loci at different LOD cutoffs, the estimates of the hybrid female and male *P. deltoides* map lengths were quite similar (Fig. 5).

Using the function given by Lange and Boehnke (1982; Eq. 4), we estimate that the maternal framework map covers greater than 99.9% of the genome at 20 cM intervals. Furthermore, the linkage groups from our female map were largely colinear with those of a previously published map for a *P. deltoides* female, 'S9-2', from family 87001 (Cervera et al. 2001). In addition, the sizes of the linkage groups were quite similar between maps. Thus, we have adopted Cervera et al.'s (2001) linkage group nomenclature and propose that this be adopted as a standard for *Populus*.

Chromosomal-scale clustering tests revealed that the AFLP marker density was significantly higher on linkage groups XII and XIII, and significantly less than expected on linkage group V (Table 1) under the expectation that marker density did not vary among linkage groups. Analysis of clustering of AFLP and SSR markers within linkage groups detected localized clustering on all linkage groups, with 13 groups showing multiple clustered regions. An average of 45% (95% CI: 38–51%) of all markers mapped in clustered regions, but these represented only 5.5% (95% CI: 3.9–7.1%) of the map distance on each linkage group on average (Table 1; Electronic Supplementary Material, Table 2). Significant dispersion of markers was also detected on eight of the linkage groups. These regions involved only 12.8% (95% CI: 9.5–16.0%) of the markers, yet covered an average of 32.9% (95% CI: 25.9–40.0%) of the map distance on each linkage group (Table 1; Electronic Supplementary Material, Table–2).

Chromosomal-scale segregation distortion

Forty-four markers were significantly distorted from the expected 1:1 segregation ratio ($P<0.05$). Forty-one of the distorted markers (94%) were clustered on six linkage groups. On a whole-genome scale, in keeping with Mendelian expectations, an average of approximately 50% of the chromosomal segments in the offspring were derived from *P. trichocarpa*, but with substantial variation among individual offspring (Electronic Supplementary Material, Fig. S3). Likewise, significant differences in transmission of alleles from the parent species were detected on linkage groups IV (92.1% of chromosome distorted), XVII (20.4% distorted) and XIX (31.1% distorted) in terms of block size, and on linkage groups IV and XIX in terms of the number of distorted markers (Table 1). Markers from the donor parent (*P. trichocarpa*) exceeded those of the recurrent parent (*P. deltoides*) in 79% of the cases of segregation distortion, which is significantly greater than expected ($\chi^2=15.4$, $P\leq0.001$). The predominant alleles on linkage group IV and XIX were from *P. trichocarpa*, whereas those on linkage group XVII were from *P. deltoides*. Small regions of distortion were also detected on other linkage groups: five with *P. trichocarpa* alleles overabundant, and three with *P. deltoides* alleles overabundant (Fig. 1).

Discussion

Map characteristics and genome length

We have reported a dense genetic map consisting of 19 linkage groups, corresponding to the number of *Populus* chromosomes. To our knowledge, this is the first forest tree mapping project that has produced the correct number of linkage groups for a single genotype with no unlinked or accessory markers [though there are examples of maps with the correct number of major linkage groups obtained by aligning maps from multiple individuals (e.g., Remington et al. 1999; Wu et al. 2000; Cervera et al. 2001; Yin et al. 2002, 2003)]. We attribute the high quality of this map to several factors: (1) the types of markers used, (2) the type of pedigree and (3) the genotyping methodology. The use of a shotgun library may enhance the evenness of coverage compared to maps composed solely of SSRs derived from enriched libraries, and the use of different marker types enhances the chances of sampling different parts of the genome (Tuskan et al. 2004). Furthermore, the use of an interspecific backcross pedigree, which is the most efficient mapping design for dominant markers, allowed us to obtain a large number of segregating loci with a modest number of AFLP primer combinations. Finally, automated data collection and post-processing with PERL scripts helped identify inconsistencies in the data and facilitated the discovery and correction of genotyping errors.

Genotyping errors, including misscoring, mislabeling or inconsistent PCR amplification, hamper all aspects of map development. Genotyping error is enhanced in outbred pedigrees, since up to four alleles can segregate per locus, increasing the chances for PCR primer competition and the probability of encountering chromosomal anomalies such as aneuploidy (Bradshaw and Stettler 1993). It is almost impossible to compare map length derived from different studies with differing rates of genotyping error. Errors inflate the number of apparent recombinations and expand map distances (Harald et al. 2000). This is especially severe when markers are tightly linked, since a misordered marker with genotyping error is most likely to be interpreted as a double crossover in regions with high marker density. For example, it has been shown that a 3% error rate in genotyping can double the genetic map length (Brzustowicz et al. 1993), and a 5% error rate precludes detection of QTLs of moderate effect (Abecasis et al. 2001).

The reported mapping distance in *Populus* varies considerably from different studies (Bradshaw et al. 1994; Wu et al. 2000; Cervera et al. 2001; Yin et al. 2002). Some discrepancies may be due in part to differences in genome coverage, the choice of mapping function, and differences in recombination rates in the parents of the crosses (Plomion and O'Malley 1996; Echt

and Nelson 1997; Remington et al. 1999). Additionally, some overestimates of linkage map length may be attributed to genotyping errors. Double crossovers and possibly misscored individuals or loci can be identified by specific commands in various mapping software packages (e.g., Lincoln and Lander 1992). Therefore, a comparison of map length with and without error detection enabled gives some indication of the level of error in the data set. For the current map, our estimates differed by only 250.4 cM, while in a previous map, the difference was 794 cM (Yin et al. 2002). Another indication of the level of genotyping error is the number of markers that cannot be properly mapped. All accessory markers in the current map were completely substitutable for framework markers mapped to the same bin. All of this suggests that genotyping error is quite low in our map and our estimates of the *Populus* map length are highly accurate.

Hulbert et al.'s (1988) method-of-moment estimator is the most widely used function for deducing genome length. This estimator is derived under the assumption of an even distribution of markers (Chakravarti et al. 1991; Gerber and Rodolphe 1994; Yin et al. 2003). Comparison of our resampling results to the observed map length derived from the recombination rate demonstrates that genome size can be overestimated if the LOD cutoff is too low for the number of loci analyzed. Given the potentially large range of variation with LOD cutoffs, genome size estimates based on Hulbert et al.'s (1988) estimator should always be provided in the context of independent estimates of genome size, or at least reported for a range of LOD cutoffs. It appears that comparisons among pedigrees can be accurately accomplished using Hulbert et al.'s (1988) estimator derived by resampling from an equivalent number of loci for all maps.

Genome dynamics

Hybrids might be expected to have suppressed recombination compared to pure species because of differentiation of the homologous chromosomes of the parental species (Jackson 1985; Tenhoopen et al. 1996; Chetelat et al. 2000). We found no evidence of recombination repression due to the heterogeneous genetic background of the female (*P. trichocarpa* × *P. deltoides*) compared to the male (pure *P. deltoides*) parent in our pedigree. However, recombination repression may have occurred on individual chromosomes, which would be beyond the achievable precision of the map distance estimates in the male parent. In fact, three of the female linkage groups showed significantly greater-than-expected marker density, which could be an indication of recombination repression (although these linkage groups may merely have an abnormally high concentration of AFLP recognition sites). Furthermore, the apparent tendency for angiosperm females to display higher recombination rates than males (DeVicente and Tanksley 1991; Ganal et al. 1995) could have partially offset repressed recombination due to the hybrid nature of our female (see Plomion and O'Malley 1996). Finally, the divergence of the parental species used in this cross may be inadequate to cause recombination repression. Although they belong to different sections of the genus, these species freely interbreed, and introgression has likely occurred throughout much of their evolutionary history (Eckenwalder 1984; Stettler et al. 1996). Furthermore, phylogenetic analyses based on morphological, chemical, and molecular characteristics have shown a close relationship between these two sections of the genus (Smith and Sytsma 1990; Eckenwalder 1996; Shi et al. 2001). Further supporting evidence is provided by the lack of rearrangements of marker order observed between the male and female maps, and the high correlation of their map distances. However, estimates of the physical length of each chromosome would be needed to explore recombination repression in *Populus* hybrids in more detail, so this remains an open question.

One of the more striking features of our map was the extensive occurrence of segregation distortion in favor of the heterospecific alleles (i.e., alleles originating from the donor parent in the backcross). Segregation distortion is commonly observed in interspecific crosses, though usually this occurs in favor of the recurrent species in a backcross (reviewed by Rieseberg and Carney 1998; Burke and Arnold 2001). In cases where heterospecific alleles are in the majority, the distortion usually occurs over small chromosomal regions (Rieseberg et al. 1996; Jiang et al. 2000; but see Kim and Rieseberg 1999). In contrast, we observed extensive regions of distortion favoring heterospecific alleles, covering nearly the entire length of a linkage group in one case (linkage group IV). Segregation distortion in hybrid pedigrees is commonly attributed to factors such as pollen-pistil incompatibilities (Lord and Russell 2002), gametic competition (Snow 1984; Carney et al. 1996; Lu et al. 2002), negative epistatic interactions among alleles (Li et al. 1997; Fishman et al. 2001), the presence of 'segregation distorter' loci that result in the destruction of alternate gametes (meiotic drive: Hartl 1974; Sano 1990), or positive selection for the introgressing alleles (Jiang et al. 2000; Burke and Arnold 2001). It appears that positive selection is the most likely explanation for our pedigree for several reasons: (1) the direction of introgression argues against prezygotic barriers to introgression and negative epistatic interactions, because these would typically favor the recurrent alleles (Kim and Rieseberg 1999; Fishman et al. 2001) and (2) the pattern of segregation distortion, which consists of a series of peaks of moderate height distributed across the linkage groups, is inconsistent with expectations for meiotic drive (Chetelat et al. 2000; Fishman et al. 2001).

This raises the question of why alleles from the *P. trichocarpa* parent would have been favored in our cross. We have begun investigating the genetic composition of the regions exhibiting segregation distortion. Interestingly, the distorted region of linkage group IV contains a locus conferring resistance against the leaf rust pathogen *Melampsora* × *columbiana* (Yin et al., unpublished;

Stirling et al. 2001) and the distorted region of linkage group XIX contains another locus conferring resistance against *Melampsora larici-populina* (Yin et al., unpublished; Zhang et al. 2001). Our pedigree experienced rust attack during the propagation phase, so it is possible that some selection was exerted in favor of resistance genes carried by the *P. trichocarpa* parent, which is not native to the region of propagation. However, the size of the distorted regions suggests that multiple loci were involved in causing the segregation distortion, especially since substantial efforts were exerted to minimize the effects of the rust on propagation success and thus minimize the selection pressure.

There are a variety of other potential selective factors, including photoperiod, frost tolerance and rooting ability, that would likely have segregated in our pedigree and may have resulted in differential survival of progeny carrying *P. trichocarpa* alleles. In the case of photoperiod and frost tolerance, selection against the maternal *P. deltoides* alleles might be expected because these were derived from an accession originating in southern Illinois (ca. 39°30′N latitude), whereas the *P. trichocarpa* alleles were derived from an accession from approximately the same latitude as the site where the pedigree was propagated (ca. 48°00′N latitude). Photoperiod is under oligogenic control in *Populus*, with individual *P. trichocarpa* loci causing earlier bud set and later bud flush in interspecific crosses (Frewen et al. 2000), which could have conferred an adaptive advantage in the severe test site in northern Minnesota. However, eight of the nine bud phenology QTL identified by Frewen et al. (2000), and all five spring bud break QTL identified by Bradshaw and Stettler (1995) could be associated with our linkage groups based on shared markers (data not shown), and none of these occurred on groups IV or XIX.

Another possible selective difference between the parent species is rooting ability: *P. trichocarpa* is much more proficient than *P. deltoides* at rooting from dormant cuttings (Riemenschneider et al. 2001) and this difference is also apparently under oligogenic control (Han et al. 1994). Since rooting from hardwood cuttings was the main method of propagation for the field trials for this pedigree, there was substantial opportunity for selection in favor of *P. trichocarpa* rooting alleles as well. However, the major rooting QTL identified by Han et al. (1994) was on linkage group I, so there is currently no evidence that rooting QTL are located in the regions showing segregation distortion. However, the Han et al. (1994) study was performed in tissue culture with a different pedigree, and this should be further investigated directly using hardwood cuttings for family 13.

It is particularly interesting that the segregation distortion in favor of heterospecific alleles occurred primarily on two linkage groups only. If this segregation distortion was caused by alleles that have been subjected to differential selection over evolutionary time, this finding suggests that genetic linkage has played a role in the development and/or maintenance of genetic isolation between *P. trichocarpa* and *P. deltoides*. A classic model of speciation would invoke differentiation of the species due to drift and/or divergent selection on reproductively isolated populations (Mayr 1942), with the accumulation of incompatible alleles randomly spread throughout the genome (Orr 1996). However, there is no evidence for prolonged allopatry for poplar species from sections *Aigeiros* and *Tacamahaca*, and hybrids between these sections are present in the fossil record for much of the evolutionary history of the genus (Eckenwalder 1984). Furthermore, this genus is apparently capable of gene flow by pollen over tremendous distances (DiFazio et al. 2004), and past and present introgression are potentially extensive (Stettler et al. 1996; DiFazio et al. 1999; Martinsen et al. 2001). Therefore, it is more likely that a parapatric or sympatric speciation model should be invoked for this species complex. Recombination is increasingly recognized as a counteracting force in most plausible scenarios of sympatric and parapatric speciation, because it breaks up associations between loci causing reproductive isolation and loci under divergent selection (Felsenstein 1981; Barton 1983; Hawthorne and Via 2001). Therefore, loci that are subject to divergent selection and which cause reproductive isolation might be expected to accumulate on individual linkage groups during speciation. This expectation is consistent with our observation of chromosomal-scale segregation distortion due to multiple loci of moderate effect.

There are, of course, many possible alternative explanations for the patterns of segregation distortion that we observed. For example, these patterns could be caused by opposing epistatic interactions among a very small number of major genes peculiar to our pedigree. However, we have seen similar patterns of segregation distortion for linkage group IV for another mapping pedigree, an F_2 family involving the same *P. trichocarpa* P_1 tree but a different *P. deltoides* genotype (family 331: Bradshaw et al. 1994). Furthermore, Cervera et al. (2001) observed segregation distortion for linkage group IV in a completely unrelated pedigree. It is therefore likely that these patterns represent evolutionarily significant differentiation between the species.

Utility of our approach for poplar genomics

Mapping of AFLP in a backcross pedigree is the most efficient means currently available for deriving a detailed picture of the genomes of highly heterozygous, long-lived organisms such as forest trees. Pure *P. deltoides* and *P. trichocarpa* have AFLP heterozygosity levels around 20–30% (Cervera et al. 2001). Therefore heterozygosity of hybrids of the two species should be approximately three times higher than in the pure species, assuming a minimum of loci fixed for the same allele in both species. In our study, 83.5% of segregating AFLPs derived from the maternal parent and 33.4% derived from the paternal parent were heterozygous, a 2.5-fold difference. Furthermore, most of the visible alleles were linked in coupling phase and originated from *P. trichocarpa*. Therefore,

mapping of AFLP markers in a backcross pedigree would be an efficient way to anchor BAC contigs to chromosomal scaffolds by screening BAC pools with AFLP primers to identify mapped fragments. Furthermore, AFLP markers could potentially be used to anchor genomic sequence to chromosomal scaffolds in silico using the recognition sites and product sizes as sequence tags. This approach requires the use of three selective nucleotides per primer, and an analysis system that yields highly precise and accurate band size estimates. However, one caveat is that AFLP markers are not evenly distributed throughout the genome, as indicated by our analysis. This may be due to an uneven distribution of the restriction enzyme recognition sites, an assertion that is supported by uneven coverage of the *Populus* BAC physical map, which was based on *Hind*III fragments (M. Marra, personal communication). If this is the case, genetic maps should be constructed from AFLP derived from multiple restriction enzymes for the purpose of genome assembly.

In the past 10 years, extensive genetic maps have been established in dozens of tree species, and QTLs have been identified for a variety of traits (e.g., Bradshaw and Stettler 1995; Grattapaglia et al. 1996; Krutovskii et al. 1998; Frewen et al. 2000; Sewell et al. 2002). However, most of these maps were built using anonymous and dominant markers like RAPD and AFLP. Although these marker types are very useful for quickly developing a genetic map for a particular cross, they are not as efficient for constructing maps for comparison. The most interesting QTLs for tree breeders will be those that could be defined as 'general' QTLs (i.e., those expressed in different genetic backgrounds; Lerceteau et al. 2000). It is difficult or perhaps impossible to establish a significant association between a marker and a QTL at a population level using maps constructed with anonymous markers (Strauss et al. 1992). Moreover, marker-QTL associations will be disrupted by recombination, so the association may not be robust in different genetic backgrounds. Consequently, the practical application of genome mapping in forest trees has frequently been questioned. However, the problem of linkage equilibrium at the population level can be greatly reduced by exploiting the genome sequence and functional genomics information to directly target candidate genes putatively involved in control of the trait of interest, thus increasing the power of marker-assisted selection (Strauss et al. 1992). SSR markers provide an ideal bridge for map comparison and direct links to the genomic sequence. Therefore, genetic maps will be a powerful tool for exploring the function of candidate genes in the post-genome era for *Populus* and other genera.

Acknowledgements We thank M. Sewell, T. Bradshaw, and W. Boerjan for providing data and advice, and two anonymous reviewers for helpful comments on the manuscript. Funding for this research was provided by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory (ORNL), and the U.S. Department of Energy, Office of Science, Biological and Environmental Research Carbon Sequestration Program. ORNL is managed by UT-Battelle, LLC for the U. S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

Abecasis GR, Cherny SS, Cardon LR (2001) The impact of genotyping error on family-based analysis of quantitative traits. Eur J Hum Genet 9:130−134

Anderson E (1949) Introgressive hybridization. Wiley, New York

Arnold ML (1997) Natural hybridization and evolution. Oxford University Press, Oxford

Barton NH (1983) Multilocus clines. Evolution 37:454–471

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. Annu Rev Ecol Syst 16:113–148

Bhalerao R, Nilsson O, Sandberg G (2003) Out of the woods: forest biotechnology enters the genomic era. Curr Opin Biotech 14:206–213

Bradshaw HD III, Stettler RF (1993) Molecular genetics of growth and development in *Populus*. I. Triploidy in hybrid poplars. Theor Appl Genet 86:301–307

Bradshaw HD III, Stettler RF (1995) Molecular genetics of growth and development in *Populus*. 4. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. Genetics 139:963–973

Bradshaw HD III, Villar M, Watson BD, Otto KG, Stewart S, Stettler RF (1994) Molecular genetics of growth and development in *Populus*. 3. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. Theor Appl Genet 89:167–178

Bradshaw HD III, Ceulemans R, Davis J, Stettler R (2000) Emerging model systems in plant biology: poplar (*Populus*) as a model forest tree. J Plant Growth Regul 3:306–313

Brzustowicz LM, Merette C, Xie X, Townsend L, Gilliam TC, Ott J (1993) Molecular and statistical approaches to the detection and correction of errors in genotype databases. Am J Hum Genet 53:1137–1145

Burke JM, Arnold ML (2001) Genetics and the fitness of hybrids. Annu Rev Genet 35:31–52

Carney SE, Hodges SA, Arnold ML (1996) Effects of differential pollen-tube growth on hybridization in the Louisiana irises. Evolution 50:1871–1878

Cervera MT, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V, Slycken JV, Montagu MV, Boerjan W (2001) Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. Genetics 158:787–809

Chakravarti A, Lasher LK, Reefer JE (1991) A maximum-likelihood method for estimating genome length using genetic linkage data. Genetics 128:175–182

Chetelat RT, Meglic V, Cisneros P (2000) A genetic map of tomato based on BC_1*Lycopersicon esculentum* × *Solanum lycopersicoides* reveals overall synteny but suppressed recombination between these homologous genomes. Genetics 154:857–867

DeVicente MC, Tanksley SD (1991) Genome-wide reduction in recombination of backcross progeny derived from male versus female gametes in an interspecific cross of tomato. Theor Appl Genet 83:173–178

DiFazio SP, Leonardi S, Cheng S, Strauss SH (1999) Assessing potential risks of transgene escape from fiber plantations. In: Lutman PJ (ed) Gene flow and agriculture: relevance for transgenic crops. British Crop Protection Council, Farnham, pp 171–176

DiFazio SP, Slavov GT, Burczyk J, Leonardi S, Strauss SH (2004) Gene flow from tree plantations and implications for transgenic risk assessment. In: Walter C, Carson M (eds) Forest biotechnology for the 21st century. Research Signpost, Kerala

Dinus RJ, Tuskan GA (1997) Integration of molecular and classical genetics: a synergistic approach to tree improvement. In: Klopfenstein NB, Chun YW, Kim M-S, Ahuja MR (eds) Micropropagation, genetic engineering, and molecular biology of *Populus*. General Technical Report RM-GTR-297, USDA Forest Service, Fort Collins, pp 220–235

Echt CS, Nelson CD (1997) Linkage mapping and genome length in eastern white pine (*Pinus strobus* L). Theor Appl Genet 94:1031–1037

Eckenwalder JE (1984) Natural intersectional hybridization between North American species of *Populus* (Salicaceae) in sections *Aigeiros* and *Tacamahaca*. III. Paleobotany and evolution. Can J Bot 62:336–342

Eckenwalder JE (1996) Systematics and evolution of *Populus*. In: Stettler RF, Bradshaw HD Jr, Heilman PE, Hinckley TM (eds) Biology of *Populus* and its implications for management and conservation. NRC Research Press, Ottawa, pp 7–32

Einspahr D, Benson MK, Peckham JR (1963) Natural variation and heritability in triploid aspen. Silvae Genet 12:51–58

Felsenstein J (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? Evolution 35:124–138

Fishman L, Kelly AJ, Morgan E, Willis JH (2001) A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. Genetics 159:1701–1716

Frewen BE, Chen THH, Howe GT, Davis J, Rohde A, Boerjan W, Bradshaw HD III (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. Genetics 154:837–845

Ganal MW, Simon R, Brommonschenkel S, Arndt M, Phillips MS, Tanksley SD, Kumar A (1995) Genetic mapping of a wide spectrum nematode resistance gene (Hero) against *Globodera rostochiensis* in tomato. Mol Plant Microbe Interact 8:886–891

Gerber S, Rodolphe F (1994) An estimation of the genome length of Maritime pine (*Pinus pinaster* Ait.). Theor Appl Genet 88:289–292

Graf J (1921) Beitrage zur Kenntnis der Gattung *Populus*. Beih Bot Central 38:405–434

Grattapaglia D, Sederoff RR (1994) Genetic-linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross-mapping strategy and RAPD markers. Genetics 137:1121–1137

Grattapaglia D, Bertolucci FLG, Penchel R, Sederoff RR (1996) Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. Genetics 144:1205–1214

Haldane JBS (1919) The combination of linkage values, and the calculation of distances between the loci of linked factors. J Genet 8:299–309

Han K, Bradshaw HD Jr, Gordon MP, Han KH (1994) Adventitious root and shoot regeneration in vitro is under major gene control in an F_2 family of hybrid poplar (*Populus trichocarpa* × *P. deltoides*). For Genet 1:139–146

Harald H, Goring H, Terwilliger JD (2000) Linkage analysis in the presence of errors II: marker-locus genotyping errors modeled with hypercomplex recombination fractions. Am J Hum Genet 66:1107–1118

Harrison JWH (1924) A preliminary account of the chromosomes and chromosome behavior in the Salicaceae. Ann Bot 38:361–378

Hartl DL (1974) Genetic dissection of segregation distortion. I. Suicide combinations of *Sd* genes. Genetics 76:477–486

Hawthorne DJ, Via S (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature 412:904–907

Hulbert S, Ilott T, Legg EJ, Lincoln S, Lander E, Milchelmore R (1988) Genetic analysis of the fungus, *Bremia lactucae*, using restriction length polymorphisms. Genetics 120:947–958

Jackson RC (1985) Genomic differentiation and its effect on gene flow. Syst Bot 10:391–404

Jiang CX, Chee PW, Draye X, Morrell PL, Smith CW, Paterson AH (2000) Multilocus interactions restrict gene introgression in interspecific populations of polyploid *Gossypium* (cotton). Evolution 54:798–814

Kim SC, Rieseberg LH (1999) Genetic architecture of species differences in annual sunflowers: implications for adaptive trait introgression. Genetics 153:965–977

Krutovskii KV, Vollmer SS, Sorensen FC, Adams WT, Knapp SJ, Strauss SH (1998) RAPD genome maps of Douglas-fir. J Hered 89:197–205

Lander E, Green P, Abrahamson J, Barlow A, Daly M, Lincoln S, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lange K, Boehnke M (1982) How many polymorphic genes will it take to span the human genome? Am J Hum Genet 24:842–845

Lerceteau E, Plomion C, Andersson B (2000) AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in *Pinus sylvestris*: a preliminary study. Mol Breed 6:451–458

Li Z, Pinson SRM, Paterson AH, Park WD, Stansel JW (1997) Genetics of hybrid sterility and hybrid breakdown in an intersubspecific rice (*Oryza sativa* l.) population. Genetics 145:1139–1148

Lincoln S, Lander E (1992) Systematic detection of errors in genetic linkage data. Genomics 14:604–610

Lord EM, Russell SD (2002) The mechanisms of pollination and fertilization in plants. Annu Rev Cell Dev Biol 18:81–105

Lu H, Romero-Severson J, Bernardo R (2002) Chromosomal regions associated with segregation distortion in maize. Theor Appl Genet 105:622–628

Maliepaard C, Jansen J, Van Ooijen JW (1997) Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. Genet Res 70:237–250

Martinsen GD, Whitham TG, Turek RJ, Keim P (2001) Hybrid populations selectively filter gene introgression between species. Evolution 55:1325–1335

Mayr E (1942) Systematics and the origin of species. Columbia University Press, New York

Neale DB, Sewell MM, Brown GR (2002) Molecular dissection of the quantitative inheritance of wood property traits in loblolly pine. Ann For Sci 59:595–605

Orr HA (1996) Dobzhansky, Bateson, and the genetics of speciation. Genetics 144:1331–1335

Plomion C, O'Malley DM (1996) Recombination rate differences for pollen parents and seed parents in *Pinus pinaster*. Heredity 77:341–350

Remington DL, Whetten RW, Liu B-H, O'Malley DM (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theor Appl Genet 98:1279–1292

Riemenschneider DE, Stanton BJ, Vallee G, Perinet P (2001) Poplar breeding strategies. In: Dickmann D, Isebrands JG, Eckenwalder JE, Richardson J (eds) Poplar culture in North America. NRC Research Press, Ottawa, pp 43–76

Rieseberg LH, Carney SE (1998) Plant hybridization. New Phytol 140:599–624

Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM (1996) Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. Science 272:741–745

Rieseberg LH, Baird SJE, Gardner KA (2000) Hybridization, introgression, and linkage evolution. Plant Mol Biol 42:205–224

Rood SB, Campbell JS, Despins T (1986) Natural poplar hybrids from southern Alberta. I. Continuous variation for foliar characteristics. Can J Bot 64:1382–1388

Sano Y (1990) The genic nature of gamete eliminator in rice. Genetics 125:183–191

Sewell MM, Davis MF, Tuskan GA, Wheeler NC, Elam CC, Bassoni DL, Neale DB (2002) Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L). II. Chemical wood properties. Theor Appl Genet 104:214–222

Shi QL, Zhuge Q, Huang MR, Wang MX (2001) Phylogenetic relationship of *Populus* sections by ITS sequence analysis. Acta Bot Sin 43:323–325

Smith EC (1943) A study of cytology and speciation in the genus *Populus* L. J Arnold Arboretum 24:275–305

Smith RL, Sytsma KJ (1990) Evolution of *Populus nigra* (Sect Aigeiros) - introgressive hybridization and the chloroplast contribution of *Populus alba* (Sect Populus). Am J Bot 77:1176–1187

Snow AA (1984) Mate choice in plants-tactics, mechanisms, and consequences. Ecology 65:1025–1026

Stettler RF, Zsuffa L, Wu R (1996) The role of hybridization in the genetic manipulation of *Populus*. In: Stettler RF, Bradshaw HD III, Heilman PE, Hinckley TM (eds) Biology of *Populus* and its implications for management and conservation. NRC Research Press, Ottawa, pp 87–112

Stirling B, Newcombe G, Vrebalov J, Bosdet I, Bradshaw HD (2001) Suppressed recombination around the *MXC3* locus, a major gene for resistance to poplar leaf rust. Theor Appl Genet 103:1129–1137

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

Taylor G (2002) Populus: *Arabidopsis* for forestry. Do we need a model tree? Ann Bot 90:681–689

Tenhoopen R, Robbins TP, Fransz PF, Montijn BM, Oud O, Gerats AGM, Nanninga N (1996) Localization of T-DNA insertions in petunia by fluorescence in situ hybridization: physical evidence for suppression of recombination. Plant Cell 8:823–830

Tuskan GA, Gunter LE, Yang ZK, Yin TM, Sewell MM, DiFazio SP (2004) Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. Can J For Res 34:85–93

Van der Schoot J, Pospíšková M, Vosman B, Smulders MJM (2000) Development and characterization of microsatellite markers in black poplar (*Populus nigra* L). Theor Appl Genet 101:317–322

Van Dillewijin C (1940) Zytologische studien in der gattung *Populus*. Genetica 22:131–182

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Vos P, Hogers R, Bleeker M, Reijans M, Vandelee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wu RL, Han YF, Hu JJ, Fang JJ, Li L, Li ML, Zeng ZB (2000) An integrated genetic map of *Populus deltoides* based on amplified fragment length polymorphisms. Theor Appl Genet 100:1249–1256

Wullschleger SD, Janssson S, Taylor G (2002a) Genomics and forest biology: *Populus* emerges as the perennial favorite. Plant Cell 14:2651–2655

Wullschleger SD, Tuskan GA, DiFazio SP (2002b) Genomics and the tree physiologist. Tree Physiol 22:1273–1276

Yin TM, Huang MR, Wang MX, Zhu LH, Zeng ZB, Wu RL (2001) Preliminary interspecific genetic maps of the *Populus* genome constructed from RAPD markers. Genome 44:602–609

Yin TM, Zhang XY, Huang MR, Wang MX, Zhuge Q, Tu SM, Zhu LH, Zeng ZB, Wu RL (2002) Molecular linkage maps of the *Populus* genome. Genome 45:541–555

Yin TM, Wang XR, Andersson B, Lerceteau-Kohler E (2003) Nearly complete genetic maps of *Pinus sylvestris* L (Scots pine) constructed by AFLP marker analysis in a full-sib family. Theor Appl Genet 106:1075–1083

Zhang J, Steenackers M, Storme V, Neyrinck S, Van Montagu M, Gerats T, Boerjan W (2001) Fine mapping and identification of nucleotide binding site/leucine-rich repeat sequences at the MER locus in *Populus deltoides* 'S9-2'. Phytopathology 91:1069–1073