Whole-Genome Characterization of Embryonic Stage Inbreeding Depression in a Selfed Loblolly Pine Family

David L. Remington*,† and David M. O'Malley*

**Department of Forestry and †Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695*

Manuscript received June 7, 1999
Accepted for publication January 10, 2000

ABSTRACT

Inbreeding depression is important in the evolution of plant populations and mating systems. Previous studies have suggested that early-acting inbreeding depression in plants is primarily due to lethal alleles and possibly epistatic interactions. Recent advances in molecular markers now make genetic mapping a powerful tool to study the genetic architecture of inbreeding depression. We describe a genome-wide evaluation of embryonic viability loci in a selfed family of loblolly pine (*Pinus taeda* L.), using data from AFLP markers from an essentially complete genome map. Locus positions and effects were estimated from segregation ratios using a maximum-likelihood interval mapping procedure. We identified 19 loci showing moderately deleterious to lethal embryonic effects. These loci account for >13 lethal equivalents, greater than the average of 8.5 lethal equivalents reported for loblolly pine. Viability alleles show predominantly recessive action, although potential overdominance occurs at 3 loci. We found no evidence for epistasis in the distribution of pairwise marker correlations or in the regression of fitness on the number of markers linked to deleterious alleles. The predominant role of semilethal alleles in embryonic inbreeding depression has implications for the evolution of isolated populations and for genetic conservation and breeding programs in conifers.

THE key role of inbreeding depression in the evolution of plant populations and mating systems has long been recognized (Charlesworth and Charlesworth 1987). Levels of inbreeding depression above critical thresholds will effectively counteract the fitness advantages of self-fertilization in plants and thus prevent the evolution of selfing. Conversely, increased inbreeding in populations that undergo bottlenecks is likely to reduce the extent of inbreeding depression through purging of deleterious recessive alleles and favors the evolution of self-fertilization (Lande and Schemske 1985; Schemske and Lande 1985). Average levels of inbreeding depression have been reported lower in predominantly selfing species than in outcrossing taxa (Husband and Schemske 1996), although a recent analysis by Byers and Waller (1999) calls these findings into question.

Observed levels of inbreeding depression provide insights about other important genetic and evolutionary parameters. The equilibrium level of inbreeding depression is a function of the genomic mutation rate U and the average dominance coefficient h of deleterious alleles. If information on h is available, the degree of inbreeding depression provides a direct estimate of U in obligately selfing populations (Charlesworth *et al.* 1990). In outcrossing plants, high values of U appear necessary to explain the levels of inbreeding depression found in many species (Kärkkäinen *et al.* 1996; Kondrashov 1998). If U is sufficiently large, the resulting high level of inbreeding depression may be maintained even with relatively high selfing rates, because selection against individual alleles will be ineffective (Lande *et al.* 1994).

The evolutionary dynamics of inbreeding depression will be profoundly influenced by the contribution of overdominant loci, the roles of lethal alleles *vs.* those with slight deleterious effects, and the extent of epistatic interactions (Charlesworth *et al.* 1991; Husband and Schemske 1996). Overdominance would favor much higher values of equilibrium heterozygosity, compared with dominance or partial dominance, and would allow high levels of inbreeding depression to occur with very low mutation rates (Charlesworth and Charlesworth 1987; Namkoong and Bishir 1987). The strength of selection against homozygous deleterious alleles largely determines the effectiveness with which they are purged in inbred populations. Models developed by Charlesworth *et al.* (1991) predict that negative synergistic epistasis will tend to increase the mean fitness of individuals but lead to greater inbreeding depression, relative to strictly multiplicative models. The relative contribution of each of these factors remains largely unresolved. Overdominant loci with respect to fitness appear to be rare (Charlesworth and Charlesworth 1987; Ritland 1996), but some studies have found substantial evidence for overdominance (Stuber *et al.* 1992; Mitchell-Olds 1995). Husband and

Corresponding author: David L. Remington, Department of Genetics, Campus Box 7614, North Carolina State University, Raleigh, NC 27695-7614. E-mail: dlreming@unity.ncsu.edu

Schemske (1996) found indirect evidence that early-acting inbreeding depression in plants is due primarily to lethal mutations, while effects later in life are probably due to a large number of mildly deleterious mutations. Data from Drosophila, however, suggest that loci with small effects do make a large contribution to overall egg-to-adult viability of homozygous chromosomes (Charlesworth and Charlesworth 1987). Several studies in Mimulus and various gymnosperms have indicated a major role for negative synergistic epistasis in early viability and in fertility (Griffin and Lindgren 1985; Willis 1993; Fu and Ritland 1996).

Gymnosperms are especially useful systems for studying inbreeding depression in early life stages. They are predominantly outcrossing, but lack genetic self-incompatibility or a separately pollinated endosperm, which would confound estimates of genetic load (Charlesworth and Charlesworth 1987; Husband and Schemske 1996). Most gymnosperms exhibit high levels of inbreeding depression, especially at the embryonic stage. Embryonic lethality generally results in empty seed, and embryonic inbreeding depression is typically measured by the ratio of filled seed frequencies under inbreeding compared with outcrossing (Williams and Savolainen 1996). Conifers typically carry an average of 8–10 embryonic lethals, although a few species show little inbreeding depression (Fowler 1965; Sorensen 1969; Franklin 1972; Park and Fowler 1982; Fowler and Park 1983; Griffin and Lindgren 1985). Actual primary selfing rates in conifers may actually be quite high, but the high lethality of selfed embryos results in low percentages of selfs observed at the seedling stage (Lande *et al.* 1994).

Molecular markers can provide a powerful supplement to the traditional techniques of quantitative genetics to investigate the genetic architecture of inbreeding depression in plants (Ritland 1996). Fu and Ritland (1996) used a modest set of unlinked isozyme markers to test for epistatic interactions among viability loci in Mimulus. Stuber *et al.* (1992) and Mitchell-Olds (1995) used comprehensive genetic maps to characterize genetic mechanisms for heterosis in crosses between inbred lines of maize and of Arabidopsis, respectively. However, the necessity of a complete genetic map and large sample sizes has made this approach infeasible for most nonmodel systems to date (Ritland 1996).

Recent advances in marker technology, however, now allow much more efficient construction of complete genetic maps and marker scoring in large progeny sets. In this article, we describe a genome-wide evaluation of embryonic viability loci in a selfed family of loblolly pine (*Pinus taeda* L.), using data from mapped markers. We previously constructed a linkage map from amplified fragment length polymorphism (AFLP) markers in megagametophytes from outcrossed seed of the same parent (Remington *et al.* 1999). In the present study, we scored this set of markers in germinating and nongerminating individuals in the selfed family. The essentially complete genome coverage of this map allows us to make inferences about the entire set of loci with major effects on viability in this individual. We estimate the total inbreeding depression explained by these loci using a maximum-likelihood interval mapping procedure and compare these estimates with those obtained from filled-seed frequencies. We also evaluate epistatic interactions between the viability loci, both by evaluating the distribution of pairwise interactions and by regressing fitness on the number of markers linked to deleterious alleles. Finally, we discuss the significance of our findings from the standpoint of evolutionary processes in long-lived woody plants and consider their implications for conservation and breeding programs.

MATERIALS AND METHODS

Selfed progeny: Selfed seed from loblolly pine clone 7-56 (NCSU-Industry Cooperative Tree Improvement Program) were provided by Westvaco Corporation and the USDA Forest Service. A total of 373 filled seeds were soaked for 5 min in 70% ethanol to kill seed coat pathogens, nicked at the micropyle end, and then soaked for 24 hr in 0.03% hydrogen peroxide. Seeds were then surface sterilized in household bleach diluted to 15% of commercial concentration for 5 min and rinsed three times for 5 min each in sterile water. Seeds were germinated under sterile conditions on 1% agar in tissue culture tubes in a growth chamber. Root emergence was checked every 2–3 days, and the date on which emergence was first observed was recorded for each seed.

Germinated seedlings were transferred to Ray Leach super cells, with peat-vermiculite growing medium when cotyledons had fully emerged, and placed in a greenhouse. Embryos were extracted from seeds that had not germinated after 44 days and stored at −80° in microfuge tubes. Surviving germinated seedlings were transferred to 10-liter containers in the summer of the first growing season and transplanted to a field site (provided by the South Carolina Forestry Commission) during the second growing season.

DNA preparations: DNA was prepared both from germinated seedlings and embryos of nongerminating seeds. Succulent stem tissue was collected from seedlings that germinated but subsequently died. Needles were collected from surviving seedlings in the second growing season. Approximately 100 mg of these tissues was placed in 1.5-ml polypropylene tubes with a ceramic ball and cylinder, in 800 μl FastDNA CLS-VF lysis solution (BIO 101, Vista, CA) with 1% soluble polyvinylpyrollidone added, and 200 μl protein precipitation solution (BIO 101). The tissues were then ground in a FastPrep instrument (BIO 101) at speed 4.5 for 15–30 sec. The ground samples were centrifuged at 14,000 × *g* for 5 min, and 700 μl supernatant was mixed with an equal volume of chloroform:isoamyl alcohol (24:1). This mixture was centrifuged at 20,000 × *g* for 5 min. The aqueous phase was transferred to fresh microfuge tubes containing 700 μl cold isopropanol to precipitate nucleic acids and centrifuged at 20,000 × *g* for 12 min. The supernatant was poured off and the pellet was rinsed twice, using 1 ml of 70% ethanol for the first rinse, and 1 ml of 95% ethanol for the second rinse, followed each time by centrifugation at 20,000 × *g* for 5 min prior to discarding the ethanol. Pellets were air dried and then resuspended in 100 μl TE buffer containing 1 μg/ml RNase A.

DNA was prepared in a similar manner from the nongermi-

nating embryos, except that the supernatant from the cell lysis and protein precipitation step was directly precipitated in isopropanol without the additional chloroform:isoamyl alcohol step. Due to the lower yields expected from this tissue, the volume of resuspension buffer was reduced to 50 μl.

AFLP template preparation and reactions: All AFLP template preparation and reactions (Vos *et al.* 1995) were carried out in 96-well format. Template preparation, preamplifcations, and selective amplifications were done as described by Remington *et al.* (1999) except as follows. Reaction mixtures for preamplification were increased to 30 μl, using 15 μl diluted restriction-ligation mixture as template, 1.8 units *Taq* polymerase (Boehringer Mannheim, Indianapolis), 45 ng *Eco*RI (E) primer, 45 ng *Mse*I (M) primer, 10 mm Tris-HCl pH 8.3, 1.5 mm $MgCl_2$, 50 mm KCl, and 0.2 mm each of all four dNTPs. Preamplification products were diluted 1:20 or 1:40 prior to use as templates for selective amplifications. The PCR profile for selective amplification was changed to 36 cycles of 10 sec denaturation at 94°, 30 sec annealing (see below), and extension (see below) at 72°. Annealing temperature was 65° for the first cycle, was reduced by 0.7° for each of the next 12 cycles, and was 56° for the remaining 23 cycles. Extension times were 60 sec for the first 13 cycles and were increased by 1 sec/cycle for subsequent cycles. Primer combinations for preamplification and selective amplification, and infrared dye (IRD) labeling of the E primers, were the same as described by Remington *et al.* (1999).

Detection and scoring of AFLP fragments: AFLP reaction products were resolved on 25-cm denaturing gels with 0.25-mm thickness, containing 8% Long Ranger polyacrylamide (FMC, Rockland, ME), 7.5 m urea, and 0.8× TBE (71.2 mm tris, 71.2 mm boric acid, 1.6 mm EDTA). Loading buffer (10 μl) consisting of 95% deionized formamide, 20 mm EDTA pH 8.0, and 1 mg/ml bromophenol blue (United States Biochemical, Cleveland) was added to each selective amplification product. This mixture was heated at 94° for 3 min, then quickly cooled on ice before loading 1.0 μl of each sample on the gel. IRD-labeled molecular weight markers (Li-Cor, Lincoln, NE) were loaded in two lanes as a standard. Electrophoresis was carried out by the NCSU DNA sequencing facility on Li-Cor 4000 and 4200L-2 automated sequencers using 0.8× TBE running buffer, with run parameters of 1500 V, 35 mA, 42 W, signal channel 3, motor speed 3, 50° plate temperature, and 16-bit pixel depth for collection of TIFF image files. Scoring of framework, alternate, and some accessory markers from the clone 7-56 linkage map from TIFF image files was done using RFLPscan Version 3.0 software (Scanalytics) as described by Remington *et al.* (1999).

Estimation of viability loci positions and effects: Scored markers were evaluated for deviations from the 3:1 ratios expected for dominant markers in F_2 configuration. All germinating and nongerminating individuals for which adequate DNA could be obtained were used for AFLP analysis. Markers for which the recessive band-absent phenotype is linked to a recessive allele affecting embryonic viability (coupling-phase markers) are expected to have an excess of band-present individuals, while repulsion-phase markers are expected to show a corresponding deficit of band-present individuals. Deviations in repulsion-phase markers are much smaller than those in coupling-phase markers. The expected frequency of band-present individuals increases from 0.75 to 1.00 if the marker is perfectly linked in coupling to a recessive lethal, but decreases to only 0.67 if the linkage is in repulsion. Consequently, the power to detect viability alleles in repulsion phase is relatively limited. Chromosomal regions where at least two adjacent markers in the same linkage phase showed a significant excess of band-present individuals ($P \leq 0.05$) were considered candidate viability loci and were evaluated further. Regions near ends of linkage groups with single distorted markers were also treated as candidate viability loci if confounding effects from comigrating fragments could be ruled out.

We used a maximum-likelihood interval mapping procedure to estimate the position and relative fitness (w) of deleterious homozygotes at each candidate viability locus. The model used a pair of dominant coupling-phase markers presumed to flank the viability locus. Our data do not have enough degrees of freedom to estimate the dominance coefficient (h) from the data, so we assumed completely recessive gene action ($h = 0$) in this model and subsequently tested the validity of this assumption with other data (see below). Estimates for w and recombination fractions r_1 and r_2 to the flanking markers that jointly maximize the likelihood of the marker data set for the two markers were generated. The model is conceptually similar to model 3 of Cheng *et al.* (1996) except that dominant markers are used and the likelihood expression is modified to obtain a legitimate multinomial likelihood function with integer counts that depend only on the observations and not on the parameters being estimated.

We designate the band-present and band-absent alleles at marker loci 1 as *A* and *a*, respectively, and the alleles at marker locus 2 as *B* and *b*. We also designate the dominant and deleterious recessive alleles at the viability locus as *Q* and *q*. The expected frequencies of the genotypes *QQ*, *Qq*, and *qq* are $1/(3 + w)$, $2/(3 + w)$, and $w/(3 + w)$. The joint probabilities for the combined marker/viability locus genotypes are given in Table 1. The observed numbers of each marker type n_i are apportioned to the viability locus genotypes using the relationship $n_{ij} = n_i g_{ij}/g_{i.}$, where g_{ij} represents the joint probability of marker type i with viability locus genotype j, and $g_{i.}$ is the sum of the g_{ij} elements over the viability locus genotypes. The likelihood function is given by

$$L(w, r_1, r_2|N) = \frac{n!}{\Pi_{i=1}^{4} n_i!} \prod_{i=1}^{4} g_{i.}^{n_i}, \tag{1}$$

and the log likelihood,

$$\log L = \log c + \sum_{i=1}^{4} n_i \log g_{i.}, \tag{2}$$

where c is the quotient of factorials from (1).

The derivatives with respect to r_1, r_2, and w are

$$\frac{\partial \log L}{\partial r_1} = \sum_{i=1}^{4} \left[n_i \frac{\sum_{j=1}^{3} \partial g_{ij}/\partial r_1}{g_{i.}} \right], \tag{3a}$$

$$\frac{\partial \log L}{\partial r_2} = \sum_{i=1}^{4} \left[n_i \frac{\sum_{j=1}^{3} \partial g_{ij}/\partial r_2}{g_{i.}} \right], \tag{3b}$$

$$\text{and } \frac{\partial \log L}{\partial w} = \sum_{i=1}^{4} n_i \frac{3(g_{i3}/w) - g_{i1} - g_{i2}}{g_{i.}(3 + w)}$$

$$= \frac{1}{3 + w} \left[\frac{3n_{.3}}{w} - n_{.1} - n_{.2} \right]. \tag{3c}$$

The maximum-likelihood estimator (MLE) for w, obtained by setting $\partial \log L/\partial w = 0$, is

$$\hat{w} = \frac{3\sum_{i=1}^{4} n_{i3}}{\sum_{i=1}^{4} (n_{i1} + n_{i2})}. \tag{4}$$

Because the values for n_{i3} themselves depend on w as well as r_1 and r_2, Equation 4 must be solved iteratively until successive values for $\hat{w}$ converge. The MLEs for r_1 and r_2 ($\hat{r}_1$ and $\hat{r}_2$) must be solved numerically.

Values for $\hat{w}$, $\hat{r}_1$, and $\hat{r}_2$ at each candidate viability locus were obtained by fixing r_1 (here defined as the distance to the most distorted marker) at a small value, finding the value for $\hat{r}_2$ to the nearest 0.01 that maximized the log likelihood, and then

TABLE 1

Joint probabilities of marker and viability locus genotypes

Marker genotype (i)	Viability locus genotype (j)			
	QQ	*Qq*	*qq*	Totals
*A_B_*	$(1 - r_1^2)(1 - r_2^2)$	$2(1 - r_1 + r_1^2)(1 - r_2 + r_2^2)$	$w(2r_1 - r_1^2)(2r_2 - r_2^2)$	$(3 + w)g_{1.}$
A_bb	$(1 - r_1^2)r_2^2$	$2(1 - r_1 + r_1^2)(r_2 - r_2^2)$	$w(2r_1 - r_1^2)(1 - 2r_2 + r_2^2)$	$(3 + w)g_{2.}$
*aaB_*	$r_1^2(1 - r_2^2)$	$2(r_1 - r_1^2)(1 - r_2 + r_2^2)$	$w(1 - 2r_1 + r_1^2)(2r_2 - r_2^2)$	$(3 + w)g_{3.}$
aabb	$r_1^2\, r_2^2$	$2(r_1 - r_1^2)(r_2 - r_2^2)$	$w(1 - 2r_1 + r_1^2)(1 - 2r_2 + r_2^2)$	$(3 + w)g_{4.}$
Totals	1	2	w	$3 + w$

All values shown must be divided by $3 + w$ to obtain probabilities.

solving iteratively for $\hat{w}$. The value of $\hat{r}_2$ was then reestimated and the cycle was repeated if $\hat{r}_2$ changed by 0.01 or more. Next, the value for r_1 was incremented by 0.01 and the process repeated. The values of $\hat{w}$, $\hat{r}_1$, and $\hat{r}_2$ at which the log likelihood was maximized were accepted as the MLEs under the hypothesis H_1 that a viability allele exists. These were tested against the null hypothesis H_0 of no viability allele by setting $w = 1$ and finding the value for $\hat{r}_2$ that maximized the log likelihood for a given r_1. A likelihood-ratio (LR) test statistic of LR = $2(\log L|H_1 - \log L|H_0)$ is approximately distributed as a chi-square random variable with 1 d.f. under H_0 (Lander and Botstein 1989). A test statistic threshold of LOD 2.8 (≈LR of 12.89, $P \approx 0.00033$) was obtained from figure 4 of Lander and Botstein (1989), using an average marker spacing of 7.5 cM and a genome length of 1700 cM (Remington *et al.* 1999). As we are applying a one-sided test that only declares viability loci when there is an excess of band-present individuals, the above P value was doubled to 0.00066, reducing the threshold LR to 11.60. This LR threshold corresponds approximately to a probability of 0.05 of declaring a false viability locus somewhere within the genome.

Multiple intervals were tested in the vicinity of the candidate viability locus unless the correct interval was obvious from the marker data. The interval producing the highest LR value was chosen as the most likely location. When w is near zero, coupling-phase markers on opposite sides of the viability locus will be nearly independent, making the correct interval readily apparent. The values of $\hat{r}_1$ and $\hat{r}_2$ were used to interpolate between marker positions on the linkage map to assign the viability locus to an estimated position. In principle, we could have used the linkage map distances to determine $r_1 + r_2$ in lieu of reestimating them from the selfed family (see Mitchell-Olds 1995). This would have provided an extra degree of freedom that could theoretically be used to estimate both h and w from the data, but the precision of the estimate of h from coupling-phase dominant markers would be too low to provide reliable estimates. The sample size of the selfed family was much larger than that of the linkage mapping population, making it preferable to estimate $\hat{r}_1$ and $\hat{r}_2$ from the selfed family data.

Evaluation of dominance hypothesis: If the effects of viability alleles are completely recessive, the expected ratios of the three genotypic classes *QQ*, *Qq*, and *qq* at a single viability locus are 1:2:w. By contrast, an overdominance model would result in expected ratios of w_1:2:w_2, where w_1 and w_2 are the respective relative fitnesses of *QQ* and *qq* relative to *Qq*. If a marker is linked in repulsion to a single recessive viability allele, the expected frequency of band-absent individuals is $(w + (1 - w)r(2 - r))/(3 + w)$, where r in this case is the recombination fraction between the band-present marker allele and the deleterious viability allele. We used this expectation to test segregation at a single repulsion-phase marker closely linked to each identified viability locus. The value for r was derived from the estimated map positions of the viability locus and the selected marker based on the linkage map. A chi-square goodness-of-fit test was used to evaluate deviations from the expected ratios. As we are interested only in deviations in the direction of overdominance (*i.e.*, a deficit of band-absent individuals), this is a one-sided test in which a threshold P value of 2α corresponds to a level α test. This test by itself does not distinguish between true overdominance and recessive alleles at separate loci linked in repulsion (pseudo-overdominance or associative overdominance).

Heterozygous effects of viability alleles were evaluated in data from the original linkage mapping population, which consisted of haploid megagametophytes from outcrossed seed of clone 7-56 (Remington *et al.* 1999). The coupling-phase markers most closely linked to identified and suggestive viability loci were evaluated for total and pooled deviations from 1:1 segregation ratios using chi-square goodness-of-fit tests. An excess of band-present individuals is expected if viability alleles are deleterious in heterozygotes.

Estimated number of lethal equivalents: A *lethal equivalent* is defined as a group of mutant alleles that, if dispersed in different individuals and made homozygous, would on average cause one death (Morton *et al.* 1956). We estimated the number of lethal equivalents represented by detected viability loci by adding the values for the estimated selection coefficients $\hat{s}_i$ (equal to $1 - \hat{w}_i$) for each of the i identified viability alleles. This estimate assumes that effects at each locus are independent and recessive.

Evaluation of biases due to linked viability loci: When two viability loci are linked, the frequency of the homozygous deleterious recessive genotype at a second viability locus will be biased by a factor $(1 - (1 - w_1)(1 - r)^2)/(3 + w_1)$, where w_1 is the fitness coefficient at the first viability locus and r is the recombination fraction between the recessive marker allele and the recessive viability allele (*i.e.*, $r \to 1$ if the alleles are closely linked in repulsion, and $r \to 0$ if the alleles are closely linked in coupling). Consequently, the expected value of the MLE $\hat{w}_2$ for the fitness coefficient at the second locus is equal to $3w_2[1 - (1 - w_1)(1 - r)^2]/(3 - r(2 - r)(1 - w_1)$, where w_2 is the true value. A bias-adjusted estimator, $\hat{w}_2'$, can thus be expressed as

$$\hat{w}_2' = \hat{w}_2 \frac{3 - r(2 - r)(1 - w_1)}{3w_1 + 3r(2 - r)(1 - w_1)}. \qquad (5)$$

Equation 5 was used to adjust fitness estimators for linked loci by initially substituting $\hat{w}_1$ for w_1 to obtain an initial estimator $\hat{w}_{2(1)}'$ and solving reciprocally for $\hat{w}_{1(1)}'$. This process was repeated iteratively by substituting the revised estimator $\hat{w}_{1(t)}'$

in Equation 5 when estimating $\hat{w}'_{2(t+1)}$, and vice versa for estimating $\hat{w}'_{1(t+1)}$, until successive pairs of estimators converge. This adjustment is only approximate, as it does not consider the effect that linkage may have on the estimated map positions of the viability loci. Maximum-likelihood methods to correct for the effects of multiple linked loci (Kao *et al.* 1999; Zeng 1994) remain to be developed for viability loci. Estimates of lethal equivalents using the bias-adjusted fitness values were compared with those obtained from unadjusted values.

The deviation D from multiplicative effects on viability of selfed progeny produced by linkage of viability loci is estimated as $s_1 s_2[1 - 4(1 - r)^2]/16$, where s_1 and s_2 are the selection coefficients ($1 - w_1$ and $1 - w_2$, respectively) at the two loci, and r is as defined above. This effect can be apportioned multiplicatively to one of the two loci by multiplying s_2 by a factor $4[1 - s_1(1 - r)^2]/(4 - s_1)$. Multiplying $\hat{s}_2$ by this factor provides the adjusted number of lethal equivalents that would be estimated from filled seed frequencies under a multiplicative model. We calculated the adjusted number of lethal equivalents by applying this factor to $\hat{s}$ at one locus of each linked pair of viability loci.

Tests for epistasis: Pairwise combinations of marker scores were evaluated for the coupling-phase markers most closely linked to identified viability loci, using JMP software (SAS Institute 1996). A contingency test was done for each pair of markers in JMP using Fisher's exact test. The distribution of the pairwise correlations between all unlinked pairs of markers was evaluated and tested for deviations from normality in JMP using a Shapiro-Wilk *W*-test. Results were compared with those from a randomly generated set of scores for an equivalent number of samples and markers. Presence-absence scores for each simulated marker were generated using a distribution of band-present frequencies identical to that of the actual marker set.

Epistasis was also evaluated using the regression method of Fu and Ritland (1996). Under the univariate model for recessive viability selection, the frequency of individuals homozygous for i markers linked to viability alleles, divided by its binomial expectation, is log-transformed and regressed on i. A linear relationship is expected if there are no interactions between loci. Negative synergistic interaction is expected to produce a negative coefficient for an i^2 term when added to the model. The same set of markers described in the previous paragraph was used for this evaluation. Only individuals that had been scored for all of the selected markers were used in this analysis.

RESULTS

Scoring of AFLP markers: A total of 279 seeds germinated out of a total of 373 that were sown (74.8%). We were able to make DNA preparations suitable for AFLP templates from 270 out of the 279 germinants and from 57 out of 94 nongerminating embryos. These 327 samples were scored for 226 AFLP markers. These consisted primarily of framework and alternate markers from the clone 7-56 genetic map (Remington *et al.* 1999). Additional accessory markers were scored to obtain additional coupling-phase markers in distorted regions to improve detection and estimation of candidate viability loci and to replace framework markers that could not be scored reliably. This provided one marker in each linkage phase for every 15 cM of the genome on average, a sufficient density to detect major viability loci in either linkage phase over nearly the entire genome. All markers were scored as dominant markers.

We arranged the scored markers in order of their linkage map positions (Remington *et al.* 1999). We identified regions on linkage groups in which multiple markers in one linkage phase, or single reliable markers at the ends of linkage groups, showed a significant ($P < 0.05$) excess of band-present individuals as candidate viability loci. Distorted markers at the ends of two linkage groups could not be adequately distinguished from comigrating fragments from other linkage groups and were not included.

Estimation of viability loci: We identified a total of 19 embryonic viability loci with LR values that exceeded the test threshold of 11.60. All of these loci are still significant at the LR threshold of 12.89 that would be appropriate for a two-tailed rather than a one-tailed test. Each of the 12 linkage groups contains at least one viability locus. Five linkage groups (LGs 1, 4, 5, 8, and 12) have two viability loci and one linkage group (LG 10) has three. One additional interval on LG 9 had an LR value of 9.37, which is suggestive of an additional viability locus. Data for all 20 of these loci are summarized in Table 2.

Three loci appear to harbor complete lethals ($\hat{s} \geq 0.99$). The remaining loci have estimated selection coefficients ranging from 0.42 to 0.93. Three sublethal loci (two on LG 10 and one on LG 12) map to regions near the terminal coupling-phase markers of their linkage groups. The actual positions of these loci could be distal to these markers, so it is possible that these alleles are also completely lethal.

Summing the values for $\hat{s}$ results in an estimate of 13.13 lethal equivalents for significant viability loci. This increases to 13.52 lethal equivalents if the suggestive viability locus on LG 9 is also included. We estimated the biases in these estimates due to effects of linked loci on $\hat{s}$. Coupling-phase linkages bias $\hat{s}$ upward, while repulsion-phase linkages result in a downward bias. This bias occurs because the expected frequency of homozygotes at neutral loci linked to a viability locus deviates from the 0.25 value used in the likelihood model. These effects are most pronounced for moderately deleterious alleles linked to lethals, while lethals themselves are not affected at all. The net effect of this bias is an overestimate of $\Sigma\hat{s}$ by only 0.14, as shown in Table 2.

Table 2 also shows the estimated effects of linked loci on the total reduction in viability due to selfing. A pair of loci linked in coupling will have less effect on viability than the same two loci would have if they were unlinked, and repulsion-phase linkages will have the opposite effect. In contrast with the biases in $\hat{s}$, these effects will be greatest when both loci approach complete lethality. The net effect on the estimate of lethal equivalents, however, is negligible.

The segregation ratios of germinated seedlings and nongerminating embryos show divergent trends at only

TABLE 2

Summary of identified and suggestive viability loci

Linkage group	Estimated position (cM)	Linkage phase[a]	$\hat{w}$	$\hat{s}$	LR	Multiple-locus r[b]	Bias-adjusted $\hat{s}$	$\hat{s}$ adjusted for effects on viability[c]
1	48.3	B	0.01	0.99	40.23*	—	0.99	0.99
1	105.3	A	0.36	0.64	150.79*	0.66	0.71	0.75
2	53.1	B	0.46	0.54	25.83*	—	0.54	0.54
3	38.3	B	0.00	1.00	141.02*	—	1.00	1.00
4	87.0	B	0.41	0.59	31.93*	—	0.64	0.59
4	130.6	A	0.53	0.47	18.61*	0.71	0.55	0.52
5	14.9	B	0.31	0.69	43.46*	—	0.67	0.69
5	119.8	B	0.29	0.71	37.63*	0.44	0.69	0.67
6	22.3	B	0.42	0.58	21.37*	—	0.58	0.58
7	132.8	A	0.57	0.43	12.92*	—	0.43	0.43
8	45.0	A	0.58	0.42	14.36*	—	0.34	0.42
8	91.0	A	0.56	0.44	14.54*	0.30	0.37	0.39
9	5.7	A	0.61	0.39	9.37	—	0.46	0.39
9	83.7	B	0.26	0.74	47.67*	0.61	0.76	0.77
10	2.5	A	0.23	0.77	62.17*	—	0.55	0.77
10	37.5	A	0.00	1.00	92.97*	0.25	1.00	0.70
10	98.5	B	0.07	0.93	108.55*	0.65	0.94	1.09
11	5.6	A	0.36	0.64	41.74*	—	0.64	0.64
12	61.3	A	0.18	0.82	75.50*	—	0.84	0.82
12	128.7	B	0.27	0.73	57.64*	0.63	0.77	0.82
Total				13.52			13.47	13.57
Significant loci only[d]				13.13			12.99	13.15

* Significant at the estimated genome-wide 0.05 level (LR = 11.60, $P = 0.00066$).

[a] Marker linkage phases within each linkage group were arbitrarily designated A or B.

[b] Recombination fraction of viability allele with viability allele at locus in previous row, estimated from linkage map positions. Values exceed 0.5 for repulsion-phase linkages.

[c] Adjustment apportioned to second locus of pair. Adjusted $\hat{s}$ may exceed 1.00 because adjustment applies to both loci of pair.

[d] Adjustments to $\hat{s}$ are not made for linkage to nonsignificant locus.

one of the identified loci (LG 8-b), where the band-absent allele frequency at the closest marker is 0.87 for germinated seedlings but only 0.70 for nongerminating embryos. The embryo DNA preparations are underrepresented in our sample due to a higher failure rate, so discrepancies in the segregation rates could confound effects on germination with effects on embryonic viability. Correcting for this bias, however, does not substantially change the overall band-present frequency, so the identification of an embryonic viability locus appears to be valid.

Tests of dominance model: Our dominant marker data lack sufficient degrees of freedom to estimate a dominance coefficient. Consequently, data from repulsion-phase markers close to the estimated position of each viability locus were evaluated for deviations from the ratios expected under the complete dominance model. All 20 identified or suggestive viability loci were tested. Data are shown in Table 3. For a test of overdominance, a P value of $2\alpha = 0.005$ is expected to result in a type I error probability of 0.05 for the entire set of 20 markers. Two markers showed deviations at this level, both of them in the direction of overdominance (*i.e.*, a deficit of band-absent alleles). Both are on LG 1, where two viability loci are linked in repulsion. The trend of segregation ratios across the linkage group suggested that the observed deviations were due to the effect of the second locus rather than true overdominance. When the test threshold was relaxed to $2\alpha = 0.10$, 4 additional loci showed deviations in the direction of overdominance. Only one of these, on LG 4, appeared to be caused by a second viability locus linked in repulsion. All other loci showed the trends expected with the dominance model, in which band-absent frequencies at repulsion-phase markers rise to a maximum in the vicinity of viability loci.

Viability alleles with deleterious effects in heterozygotes would be expected to cause an excess of band-absent alleles in repulsion-phase markers. No markers showed significant deviations in this direction. Heterozygous fitness effects of viability alleles were not observed in markers scored in megagametophytes from outcrossed 7-56 seed either (Table 4). The coupling-phase markers most tightly linked to viability alleles had a pooled frequency of 0.491, which deviates nonsignificantly from the null expectation of 0.5 and in the oppo-

TABLE 3

Test for overdominance

Linkage group-locus	Estimated position (cM)	Repulsion-phase marker	Marker position	Expected no. band-absent	Observed no. band-absent	*P* value
1-a	48.3	ACG/CCAA-172	43.8	107.78	78	0.0005**,a
1-b	105.3	ACA/CCGC-273	102.8	95.50	65	0.0002**,a
2-a	53.1	ACT/CCAG-364	56.8	90.40	71	0.016*
3-a	38.3	ACA/CCGG-233	39.6	107.32	119	0.167
4-a	87.0	ACA/CCAG-196	64.5	92.60	78	0.072*,a
4-b	130.6	ACG/CCAG-152	126.3	90.01	78	0.135
5-a	14.9	ACA/CCAG-152	11.0	97.19	90	0.383
5-b	119.8	ACG/CCAA-440	120.0	78.72	89	0.165
6-a	22.3	ACC/CCAG-204	23.1	93.27	84	0.254
7-a	132.8	ACG/CCAA-255	128.5	90.97	91	0.997
8-a	45.0	ACG/CCAA-287	48.4	90.18	90	0.982
8-b	91.0	ACG/CCTA-401	98.1	88.87	74	0.063*
9-a	5.7	ACA/CCGC-213	7.8	88.63	72	0.038*
9-b	83.7	ACG/CCCA-62	88.2	97.72	109	0.171
10-t	2.5	ACG/CCTC-356	17.2	97.22	104	0.409
10-m	37.5	ACC/CCAG-84	27.6	105.48	108	0.764
10-b	98.5	ACC/CCAG-66	106.3	103.40	108	0.583
11-a	5.6	ACT/CCCG-177	5.6	97.02	91	0.466
12-a	61.3	ACG/CCAG-270	63.2	99.97	94	0.471
12-b	128.7	ACG/CCTC-385	117.6	96.24	86	0.211

*, Significant at $\alpha = 0.05$ level; **, significant at $\alpha = 0.0025$ level.
[a] Locus is linked in repulsion to a second viability locus.

site direction to that predicted if viability alleles are deleterious in heterozygotes. When these loci are evaluated individually, the total and heterogeneity chi-square values also show insignificant deviations from null expectations ($P = 0.28$ and $P = 0.25$, respectively). If the null hypothesis is changed to reflect expectations under a dominance coefficient of 0.2, however, the total and pooled deviations become highly significant.

Tests for epistasis: To test for epistatic interactions, we selected the single coupling-phase marker closest to the estimated position of each of the 20 identified or suggestive viability loci. A marker linked to an additional possible viability locus on LG 2 ($\hat{s} = 0.25$, LR = 4.52) was also used in this analysis, for a total of 21 loci. We included this locus because s will be ≤ 0.25 for an allele with recessive colethal interactions with a second locus. Contingency tests were done on the four two-locus marker classes for each of the 210 possible pairs of markers, using Fisher's exact test.

Only 8 marker pairs deviated from expected numbers of markers per class at the 0.05 level. This is close to the 10.5 pairs expected to deviate by chance alone. Only 1 pair of markers deviated at the 0.01 level, and these were the markers linked to the two loci on LG 8, which are linked genetically. Three other marker pairs showed deviations approaching the 0.01 level. One of these involved markers that were both linked to complete lethals, for which no interaction effects are possible. The other two pairs both involved LG 4-b. One of these two pairs (LG 4-b-LG 2-a) showed the negative correlation expected for a negative synergistic interaction. Only one individual was band-absent for both the LG 2-a and LG 4-b markers, close to the expectation for complete colethals. Thus, there is no power to distinguish even strong negative synergistic interactions from random variation in a data set of this size.

The distribution of the 200 correlation coefficients of unlinked marker pairs was characterized. Results from this comparison are shown in Table 5. The actual correlations deviated significantly from normality, primarily due to a substantial positive skewness caused by an excess of slightly negative correlations. The mean correlation did not deviate significantly from zero. This effect could potentially be due to a tendency toward slightly negative epistatic interactions or may simply reflect skewness in the shape of the binomial distribution of one or more marker class frequencies. The expected number of double band-absent individuals is very small, which would create a positively skewed distribution. To distinguish between these possibilities, the distribution was compared to those from two sets of 21 randomly generated markers scored for 320 individuals. Each simulated marker was generated using a band-present frequency corresponding to one of the 21 actual markers to make the two-marker sets as similar as possible for factors other than epistatic interactions. As shown in Table 5, the simulated data have comparable distributions with a similar degree of positive skewness.

Finally, we tested the combined scores for all 21 markers using the regression approach of Fu and Ritland

TABLE 4

Test for partial dominance

Marker[a]	Individuals scored[b]	Band-present individuals[b]	Chi square test statistic: Hypothesis $h = 0$	Chi square test statistic: Hypothesis $h = 0.2$[c]	d.f.
ACG/CCCA-192	91	55	3.967	1.632	1
ACA/CCGG-51	92	41	1.087	3.131	1
ACT/CCAG-52	91	45	0.011	0.681	1
ACC/CCAG-154	92	40	1.565	3.915	1
ACG/CCTA-151	93	43	0.527	2.120	1
ACG/CCGC-262	91	49	0.538	0.000	1
ACG/CCAA-293	91	47	0.099	0.164	1
ACG/CCTA-208	93	49	0.269	0.043	1
ACG/CCTA-430	93	39	2.419	5.234	1
ACA/CCTG-283	88	42	0.182	1.290	1
ACA/CCGC-316	92	40	1.565	3.915	1
ACG/CCTA-402	93	47	0.011	0.389	1
ACA/CCTG-422	88	49	1.136	0.130	1
ACA/CCAG-351	92	37	3.522	6.791	1
ACA/CCTG-453	88	53	3.682	1.479	1
ACG/CCCA-476	91	40	1.330	3.521	1
ACT/CCGC-78	90	45	0.000	0.513	1
ACA/CCGG-140	92	44	0.174	1.305	1
ACG/CCTA-306	93	49	0.269	0.043	1
ACA/CCAG-527	92	42	0.696	2.435	1
Totals	1826	896	23.049	38.731	20
P values			0.286	0.007	
Pooled chi square			0.317	16.188	1
P values			0.573	<0.001	
Heterogeneity chi square			22.731	22.543	19
P values			0.249	0.258	

[a] Coupling-phase markers nearest identified or suggestive viability loci.
[b] Markers scored in megagametophytes from outcrossed half-sib progeny of 7-56.
[c] Assumed average $s = 0.7$. Expected band-present frequency is $1/(2 - (0.7)(0.2)) = 0.538$.

(1996). We used two versions of the univariate regression model for recessive viability selection. The first version is a simple linear model $\ln f(i) = a + b_1 i$, in which $f(i)$ is the number of individuals homozygous for i of the 21 markers linked to the identified or possible viability alleles, divided by its binomial expectation without selection. The second version, $\ln f(i) = a + b_1 i + b_{11} i^2$, includes a quadratic term whose coefficient is expected to be negative in the presence of negative synergistic epistasis. Results are summarized in Table 6. We found that the linear coefficient alone provided an excellent fit with the observed data ($R^2 = 0.9973$). Adding

TABLE 5

Comparison of distributions of pairwise correlations for scores of actual and simulated marker sets

Moments/tests	Actual marker set (unlinked pairs only)	Simulated marker set 1	Simulated marker set 2
Mean	−0.0006	−0.0068	−0.0028
Variance	0.0037	0.0028	0.0029
Skewness	0.5823	0.399	0.7318
Kurtosis	0.2087	−0.0318	0.4594
Test for mean = 0			
t-statistic	−0.1434	−1.8650	−0.7553
$P(T > t)$	0.8861	0.0636	0.4509
Shapiro-Wilk normality test			
w-statistic	0.962	0.965	0.948
$P(W < w)$	0.0008	0.0018	0.0001

TABLE 6

Results of regression of the log-transformed function of the number of individuals homozygous for i recessive marker alleles linked to viability alleles, on i and i^2

Coefficients/tests	Simple linear model[a] $\ln f(i) = a + b_1 i$	Linear + quadratic model[a] $\ln f(i) = a + b_1 i + b_{11} i^2$
Intercept (a)	3.498	3.456
t-statistic	43.412	31.267
P value	1.22×10^{-7}	6.24×10^{-6}
Linear coefficient (b_1)	−0.957	−0.907
t-statistic	−42.831	−10.516
P value	1.31×10^{-7}	4.62×10^{-4}
Quadratic coefficient (b_{11})	—	−0.0082
t-statistic	—	−0.599
P value	—	0.581
Degrees of freedom		
Regression	1	2
Residual	5	4
F-statistic	1834.47	799.79
P value	1.31×10^{-7}	6.22×10^{-6}
R^2	0.9973	0.9975

[a]

$$f(i) = \frac{n_i}{\binom{21}{i}(0.25)^i\,(0.75)^{21-i}},$$

where n_i is the number of individuals with band-absent type for i markers.

the quadratic term increased the R^2 only slightly and actually reduced the significance of the model due to the loss of an additional degree of freedom. The coefficient (b_{11}) of the quadratic term was small and insignificant ($P = 0.581$).

The existence of a prior genetic map allows us to obtain estimates for the average selection coefficient $\bar{s}$ and for the strength of synergism $\bar{t}/\bar{s}$ from the regression model as described by Fu and Ritland (1996; Charlesworth *et al.* 1991). These estimates are functions of the regression coefficients, the mean estimated recombination fraction $\bar{r}$ between marker and viability locus and the variance in recombination fraction σ_a^2. Using this method with the simple linear model with our values of $\bar{r} = 0.0305$ and $\sigma_a^2 = 0.0010$ produces an estimate of $\bar{s} = 0.649$, which is very close to the mean $\hat{s}$ of 0.656 obtained from the MLEs when all 21 loci are used. The estimate of $\bar{t}/\bar{s}$ from the quadratic model is ~0.002, which is not significantly different from 0.

DISCUSSION

Estimation of selection coefficients and lethal equivalents: Our mapping identified 19 viability loci plus 1 additional highly suggestive locus. Marker coverage in both linkage phases was sufficient to detect major viability throughout the genome, with the possible exception of some linkage group ends. The power to detect viability alleles with a given effect using our interval mapping procedure did not appear to vary greatly between different regions of the genome. The 19 identified loci alone account for 13.13 lethal equivalents. Effects of slightly deleterious alleles are not detected by mapping, in contrast with estimates from filled-seed frequencies. A comparison of estimates from the two methods can provide an indication of the amount of inbreeding depression contributed by slightly deleterious alleles. We are unaware of any filled-seed frequency data specific to clone 7-56, but the average number of embryonic lethal equivalents for *P. taeda* has been estimated at 8.5 (Franklin 1972). Only 9 of 116 trees in Franklin's study (7.8%) had estimates of >14 lethal equivalents. Thus, it seems unlikely that additional loci with slightly deleterious effects are contributing much to inbreeding depression in this family.

Estimation of lethal equivalents from mapping data can be affected by several other factors. There is limited power to identify loci whose fitness effects are close to the statistical threshold for detection. This will lead to overestimates of the true effects in experiments in which the loci are detected (Wilcox *et al.* 1997; Beavis 1998). Some of the loci we have identified probably belong to this category. However, we have probably failed to detect other viability alleles with marginal effects, so it is unlikely that we have substantially overestimated the number of lethal equivalents explained by lethal or semilethal alleles. Linkage between viability loci can bias the estimated selection coefficients, and linked loci will also not affect viability independently. We evaluated the effects of linkage on our estimates of s and on viability.

The opposing effects of coupling- and repulsion-phase linkages resulted in very little net effect on our estimates. Partial dominance of viability alleles would result in underestimates of selection coefficients, but our data show no evidence that dominance coefficients are greatly different from zero. Epistasis will also result in nonindependence of loci. Negative synergistic epistasis (Charlesworth *et al.* 1991; Fu and Ritland 1996) will cause mapping data estimates to be inflated, as effects detected at multiple loci are in effect contributing to the same embryo death. However, epistasis does not appear to be affecting our results (see below). Finally, only segregating variation can be detected by mapping. Deleterious alleles that have been fixed by inbreeding will not be detected by mapping, but they will contribute to reductions in seed viability.

Our data partially support the conclusion of Husband and Schemske (1996) that inbreeding depression in early life stages is predominantly the result of lethal rather than slightly deleterious mutations. However, only 3 of the 19 alleles we identified appear to be completely lethal, with a slight possibility that 3 other alleles at loci near the ends of linkage groups are lethal as well. The average $\hat{s}$ at these 19 loci is 0.69, and this estimate may be too large if some of the "alleles" we have identified are actually alleles at two or more loci linked in coupling. The preponderance of semilethal viability alleles is consistent with recent findings of Kuang *et al.* (1999) in *P. radiata*.

Dominance *vs.* overdominance: Six of the 20 identified or suggestive viability loci show evidence of marker frequency deviations in the direction of overdominance. In three of these cases the explanation appears to be linkage in repulsion to other identified viability loci and not true overdominance. Deviations at the other 3 loci could be due to true overdominance or to linkage in repulsion to a weaker unidentified recessive viability allele. Some of these deviations could also be due to chance, as none were significant at the multiple-test *P* value of 0.005. Mitchell-Olds (1995) found strong evidence for an overdominant locus in Arabidopsis. He favored true overdominance and not pseudo-overdominance to explain the effects because the apparent selection against the two homozygotes was nearly equal, and markers on either side of the region did not show directional dominance. However, the 3 loci in question in our study show unequal selection against the two homozygotes, and the band-present frequency in the repulsion-phase markers reaches a maximum at a point 20 cM from the estimated viability locus position in one of the three cases. Consequently, pseudo-overdominance appears to be at least as reasonable as true overdominance as an explanation for our data.

Epistatic interactions: Detection of individual negative synergistic interactions is not feasible with the sample size we used. The expectation under negative synergism is for a deficit of individuals that are homozygous recessive at both loci. This will result in negative correlations of scores for the associated markers, similar to repulsion-phase linkages. The information content for dominant loci linked in repulsion is extremely low (Liu 1998) and is even lower when homozygous recessive individuals are underrepresented due to reduced viability. Availability of codominant markers would make little difference because the low power of detection is due to the recessive nature of the viability alleles, not to that of the linked markers.

Neither the overall distribution of the pairwise correlations between loci nor the regression model of Fu and Ritland (1996) showed any overall evidence for pairwise epistasis. The regression model is especially appealing, as it is conceptually similar to levels-of-inbreeding studies without requiring the use of complex field study designs and multiple generations of inbreeding. The power to detect epistasis with the regression approach in our study should be relatively high in spite of the modest sample size, because linkage phases are obvious and markers are generally very close to the estimated viability locus positions. The very strong fit of the linear regression model with only the linear coefficient ($R^2 = 0.9973$, $P = 1.31 \times 10^{-7}$) was unexpected. The support for the linear model and the lack of significance of the quadratic coefficient strongly suggest the absence of appreciable epistasis.

Our results contrast with those of Fu and Ritland (1996) in a marker-based study in *Mimulus guttatus* and those of Griffin and Lindgren (1985) using filled-seed frequency data in *P. radiata* and several other conifers. Woods and Heaman (1989) failed to find evidence of epistasis with filled-seed frequencies in Douglas fir, but a reanalysis of their data by Willis (1993) using log-transformed frequencies obtained a strong quadratic regression coefficient. Our study involves a single and possibly atypical parent, but the number of loci identified represents a large number of potential pairwise interactions. If epistatic interactions among viability alleles are common in pine, they should have been evident among this set of alleles. Stabilizing selection on quantitative traits will produce negative synergism between loci due to the concave shape of the fitness curve (Lande and Schemske 1985). However, embryonic viability seems unlikely to be affected by selection on quantitative characters, so negative synergism may be less likely at the embryonic stage. Mapping studies would be useful to verify the reported evidence for epistasis in embryonic viability in other conifers.

Implications for evolutionary processes: Loblolly pine, like most other gymnosperms, shows high levels of embryonic inbreeding depression. Maintaining these levels of inbreeding depression under dominance models requires a low natural selfing rate and a high value of U per generation. The levels of inbreeding depression found in gymnosperms and many other perennial plants suggest a value of U at least an order of magnitude

higher than those in Drosophila and Arabidopsis (Simmons and Crow 1977; Lande *et al.* 1994; Kärkkäinen *et al.* 1996; Hedrick *et al.* 1999).

Overdominance and epistasis could account for high levels of inbreeding depression without requiring high values of U (Charlesworth and Charlesworth 1987; Namkoong and Bishir 1987; Charlesworth *et al.* 1991). However, we find no support in our data for an important role for epistasis and weak evidence at best for overdominance. Conversely, even weak selection against heterozygotes for deleterious alleles will substantially reduce the equilibrium level of inbreeding depression. A value of $h \approx 0.02$ is commonly assumed for lethals based on studies in *Drosophila melanogaster* (Simmons and Crow 1977; Charlesworth and Charlesworth 1987; Lande *et al.* 1994; Hedrick *et al.* 1999). Such a small value of h would be indistinguishable from completely recessive viability alleles in our data. Dominance coefficients of slightly deleterious alleles in Drosophila have average values of 0.20–0.35 (Simmons and Crow 1977), but our data from outcrossed progeny of 7-56 are inconsistent with heterozygous effects of this magnitude. The existence of high levels of U in trees does not seem unreasonable, as trees have long generation times with many more germline cell divisions between meioses than do annual plants or animals with separation of germlines (Lande *et al.* 1994; Gill *et al.* 1995).

Our finding that numerous linkages among viability loci have almost no net effect on viability estimates may have significance for estimation of equilibrium inbreeding depression. Models typically assume that loci are unlinked, which is not true in our data (Lande and Schemske 1985; Charlesworth *et al.* 1991; Lande *et al.* 1994). If the effects of coupling- and repulsion-phase linkages approximately cancel each other, however, linkage is probably not an important source of bias in these models.

The predominance of semilethal as opposed to completely lethal alleles at viability loci can affect the evolutionary dynamics of populations in the early stages of inbreeding. The expected frequency of homozygous recessive zygotes in the second generation of selfing (S_2) is $(3 - 2s)/(8 - 2s)$, compared with $\frac{1}{4}$ in the first (S_1) generation. Consequently, the inbreeding depression contributed by deleterious alleles with $s < \frac{2}{3}$ will increase rather than decrease in the S_2 generation. Nine of our 19 identified viability loci fall into this category. Lower purging rates for moderately deleterious alleles compared to lethals will lead to greater probabilities of fixation in isolated populations. All but 21 of the 292 progeny included in the regression analysis for epistatic effects were homozygous for at least one of the markers linked to viability alleles.

This study is based on a single *P. taeda* parent, and other individuals need to be studied to evaluate the generality of these results. The viability loci we have identified are a small proportion of the thousands of loci that must harbor deleterious mutations, given the observed levels of inbreeding depression in pines (Hedrick *et al.* 1999). Nevertheless, it seems highly unlikely that the overall range of effects and interactions among this sample of 20 or so loci is atypical for highly deleterious (as opposed to slightly deleterious) alleles in pines.

These results have relevance for genetic conservation and tree improvement programs. Fixation of deleterious alleles could have serious implications for populations that experience bottlenecks due either to human-induced environmental changes or the use of small breeding populations. Current loblolly pine breeding strategies in the southeastern United States call for the use of elite sublines of four trees each (McKeand and Bridgwater 1998). Clone 7-56 has been highly favored by tree breeders due to its exceptional breeding value for growth rate. Marker-assisted breeding may be essential if fixation of deleterious mutations from elite breeding selections is to be avoided in small breeding lines.

Genetic mapping as a tool for population genetic studies: Ritland (1996) recently summarized the opportunities to use genetic markers to gain new insights on inbreeding depression. He proposed sampling genomes with a small set of unlinked markers such as isozymes, as constructing complete linkage maps still appeared unrealistic for species other than important crops and model organisms. In the last few years, however, development of new marker techniques such as AFLPs has allowed *de novo* construction of genetic maps in a short period of time. Effective multiplex ratios of 20–25 polymorphisms per primer combination are common in outbred organisms, and shared polymorphisms will commonly be frequent enough to join maps of different individuals even if multiallelic markers are unavailable (Rouppe van der Voort *et al.* 1997; Marques *et al.* 1998; Remington *et al.* 1999). These advances should now make whole-genome map-based studies accessible to almost any organism.

We thank Ross Whetten, Zander Myburg, and Arthur Johnson for technical advice; Trudy Mackay, Bruce Weir, Steve McKeand, Ben-Hui Liu, Outi Savolainen, and Don Waller for valuable discussions and suggestions; Ron Sederoff for his guidance and encouragement; and Ruth Shaw and two anonymous reviewers for helpful comments on the manuscript. This work was supported by funding from the North Carolina State University Forest Biotechnology Industrial Associates Consortium, the National Institutes of Health (grant GM-45344), and by a United States Department of Agriculture National Needs Fellowship to D.L.R.

LITERATURE CITED

Beavis, W. D., 1998 QTL analyses: power, precision, and accuracy, pp. 145–162 in *Molecular Dissection of Complex Traits*, edited by A. H. Paterson. CRC Press, Boca Raton, FL.

Byers, D. L., and D. M. Waller, 1999 Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. Annu. Rev. Ecol. Syst. **30:** 497–513.

Charlesworth, D., and B. Charlesworth, 1987 Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. **18:** 237–268.

Charlesworth, B., D. Charlesworth and M. T. Morgan, 1990 Genetic loads and estimates of mutation rates in highly inbred plant populations. Nature **347:** 380–382.

Charlesworth, B., M. T. Morgan and D. Charlesworth, 1991 Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. Genet. Res. **57:** 177–194.

Cheng, R., A. Saito, Y. Takano and Y. Ukai, 1996 Estimation of the position and effect of a lethal factor locus on a molecular marker linkage map. Theor. Appl. Genet. **93:** 494–502.

Fowler, D. P., 1965 Effects of inbreeding in red pine, *Pinus resinosa* Ait. II. Pollination studies. Silvae Genet. **14:** 12–23.

Fowler, D. P., and Y. Park, 1983 Population studies of white spruce. I. Effect of self-pollination. Can. J. For. Res. **13:** 1133–1138.

Franklin, E. C., 1972 Genetic load in loblolly pine. Am. Nat. **106:** 262–265.

Fu, Y.-B., and K. Ritland, 1996 Marker-based inferences about epistasis for genes influencing inbreeding depression. Genetics **144:** 339–348.

Gill, D. E., L. Chao, S. L. Perkins and J. W. Wolf, 1995 Genetic mosaicism in plants and clonal animals. Annu. Rev. Ecol. Syst. **26:** 423–444.

Griffin, A. R., and D. Lindgren, 1985 Effect of inbreeding on production of filled seed in *Pinus radiata*—experimental results and a model of gene action. Theor. Appl. Genet. **71:** 334–343.

Hedrick, P. W., O. Savolainen and K. Kärkkäinen, 1999 Factors influencing the extent of inbreeding depression: an example from Scots pine. Heredity **82:** 441–450.

Husband, B. C., and D. W. Schemske, 1996 Evolution of the magnitude and timing of inbreeding depression in plants. Evolution **50:** 54–70.

Kao, C. H., Z. B. Zeng and R. D. Teasdale, 1999 Multiple interval mapping for quantitative trait loci. Genetics **152:** 1203–1216.

Kärkkäinen, K., V. Koski and O. Savolainen, 1996 Geographical variation in the inbreeding depression of Scots pine. Evolution **50:** 111–119.

Kondrashov, A. S., 1998 Measuring spontaneous deleterious mutation process. Genetica **102/103:** 183–197.

Kuang, H., T. E. Richardson, S. D. Carson and B. C. Bongarten, 1999 Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. II. Genetics of viability genes. Theor. Appl. Genet. **99:** 140–146.

Lande, R., and D. W. Schemske, 1985 The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution **39:** 24–40.

Lande, R., D. W. Schemske and S. T. Schultz, 1994 High inbreeding depression, selective interference among loci, and the threshold selfing rate for purging recessive lethal mutations. Evolution **48:** 965–978.

Lander, E. S., and D. Botstein, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

Liu, B.-H., 1998 *Statistical Genomics: Linkage, Mapping and QTL Analysis.* CRC Press, Boca Raton, FL.

Marques, C. M., J. A. Araujo, J. G. Ferreira, R. Whetten, D. M. O'Malley *et al.*, 1998 AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis.* Theor. Appl. Genet. **96:** 727–737.

McKeand, S. E., and F. E. Bridgwater, 1998 A strategy for the third breeding cycle of loblolly pine in the southeastern U.S. Silvae Genet. **47:** 223–234.

Mitchell-Olds, T., 1995 Interval mapping of viability loci causing heterosis in *Arabidopsis.* Genetics **140:** 1105–1109.

Morton, N. E., J. F. Crow and H. J. Muller, 1956 An estimate of the mutational damage in man from data on consanguineous marriages. Proc. Natl. Acad. Sci. USA **42:** 855–863.

Namkoong, G., and J. Bishir, 1987 The frequency of lethal alleles in forest tree populations. Evolution **41:** 1123–1127.

Park, Y. S., and D. P. Fowler, 1982 Effects of inbreeding and genetic variances in a natural population of tamarack (*Larix laricina* (Du Roi) K. Koch) in eastern Canada. Silvae Genet. **31:** 21–26.

Remington, D. L., R. W. Whetten, B.-H. Liu and D. M. O'Malley, 1999 Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda.* Theor. Appl. Genet. **98:** 1279–1292.

Ritland, K., 1996 Inferring the genetic basis of inbreeding depression in plants. Genome **39:** 1–8.

Rouppe van der Voort, J. N. A. M., P. van Zandvoort, H. J. van Eck, R. T. Folkertsma, R. C. B. Hutten *et al.*, 1997 Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol. Gen. Genet. **255:** 438–447.

SAS Institute, 1996 JMP Version 3.2.2. SAS Institute, Inc., Cary, NC.

Schemske, D. W., and R. Lande, 1985 The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. Evolution **39:** 41–52.

Simmons, M. J., and J. F. Crow, 1977 Mutations affecting fitness in *Drosophila* populations. Annu. Rev. Genet. **11:** 49–78.

Sorensen, F., 1969 Embryonic genetic load in coastal Douglas-fir, *Pseudotsuga menziesii* var. *menziesii.* Am. Nat. **103:** 389–398.

Stuber, C. W., S. E. Lincoln, D. W. Wolff, T. Helentjaris and E. S. Lander, 1992 Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics **132:** 823–839.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee *et al.*, 1995 AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23:** 4407–4414.

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 *IUFRO '97 Genetics of Radiata Pine.* FRI Bulletin No. 203, Rotorua, New Zealand.

Williams, C. G., and O. Savolainen, 1996 Inbreeding depression in conifers: implications for breeding strategy. For. Sci. **42:** 102–117.

Willis, J. H., 1993 Effects of different levels of inbreeding on fitness components in *Mimulus guttatus.* Evolution **47:** 864–876.

Woods, J. H., and J. C. Heaman, 1989 Effect of different inbreeding levels on filled seed production in Douglas-fir. Can. J. For. Res. **19:** 54–59.

Zeng, Z-B., 1994 Precision mapping of quantitative trait loci. Genetics **136:** 1457–1468.

Communicating editor: R. G. Shaw