Genetics of Postzygotic Isolation in Eucalyptus: Whole-Genome Analysis of Barriers to Introgression in a Wide Interspecific Cross of *Eucalyptus grandis* and *E. globulus*

Alexander A. Myburg,*,†,[1] Claus Vogl,‡ A. Rod Griffin,§ Ronald R. Sederoff*,† and Ross W. Whetten*

**Department of Forestry and †Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695, ‡Department of Computer Science, University of California, Riverside, California 92521 and §CRC for Sustainable Production Forestry and School of Plant Science, University of Tasmania, Hobart, 7001, Australia*

Manuscript received September 26, 2002
Accepted for publication November 10, 2003

ABSTRACT

The genetic architecture of hybrid fitness characters can provide valuable insights into the nature and evolution of postzygotic reproductive barriers in diverged species. We determined the genome-wide distribution of barriers to introgression in an F_1 hybrid of two Eucalyptus tree species, *Eucalyptus grandis* (W. Hill ex Maiden.) and *E. globulus* (Labill.). Two interspecific backcross families ($N = 186$) were used to construct comparative, single-tree, genetic linkage maps of an F_1 hybrid individual and two backcross parents. A total of 1354 testcross AFLP marker loci were evaluated in the three parental maps and a substantial proportion (27.7% average) exhibited transmission ratio distortion ($\alpha = 0.05$). The distorted markers were located in distinct regions of the parental maps and marker alleles within each region were all biased toward either of the two parental species. We used a Bayesian approach to estimate the position and effect of transmission ratio distorting loci (TRDLs) in the distorted regions of each parental linkage map. The relative viability of TRDL alleles ranged from 0.20 to 0.72. Contrary to expectation, heterospecific (donor) alleles of TRDLs were favored as often as recurrent alleles in both backcrosses, suggesting that positive and negative heterospecific interactions affect introgression rates in this wide interspecific pedigree.

THE nature of postzygotic reproductive barriers is a fundamental question in evolutionary genetics. More than 60 years after it was proposed that postzygotic barriers may be by-products of genetic differentiation among diverging species (Dobzhansky 1937; Muller 1942), the mechanisms involved in postzygotic isolation remain unknown for many groups of organisms, including most plant species. Two general mechanisms are thought to underlie postzygotic isolation. Deleterious interactions between heterospecific genes may lead to hybrid sterility and hybrid inviability, while chromosomal rearrangements in the parental species may result in abnormal meiotic products in their hybrids with negative effects on hybrid fertility (Burke and Arnold 2001). Both of these mechanisms have been shown to reduce gene flow in hybrids of plant species (Rieseberg *et al.* 1995; Quillet *et al.* 1995; Fishman and Willis 2001), but the relative importance of the two mechanisms (genic *vs.* chromosomal incompatibilities) is still questioned (Lynch and Force 2000; Rieseberg 2001). It has been proposed that chromosomal rearrangements may reduce hybrid fitness in plants primarily by suppressing recombination and thereby extending the effects of linked isolation genes (Rieseberg 2001). The sizes of chromosomal segments over which gene flow is restricted may therefore be determined by the combined effect of isolation genes and recombination modifiers in hybrid genomes (Rieseberg and Burke 2001).

The prevailing view emerging from studies of postzygotic barriers in plants and animals is that postzygotic isolation is caused mostly by extensive negative epistatic interactions in hybrid genomes (Burke and Arnold 2001). Simple two-locus incompatibilities such as those described by the Dobzhansky-Muller model predict the establishment of a postzygotic barrier when a gene from one species interacts negatively within the hybrid genetic background with a gene from another species (Dobzhansky 1937; Muller 1942). Variations of the Dobzhansky-Muller model can be used to predict patterns of reduced hybrid fitness observed in many different hybrids of animal species (Turelli and Orr 2000). However, except for a small number of case studies (Fishman and Willis 2001; Fishman *et al.* 2001), little experimental evidence has been obtained to support the importance of simple Dobzhansky-Muller incompatibilities in the reproductive isolation of plant species. Instead, the observation of multilocus interactions in several hybrid populations (Rieseberg *et al.* 1996; Burke *et al.* 1998; Jiang *et al.* 2000) has suggested the impor-

[1]*Corresponding author:* Department of Genetics, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, 0002, South Africa.
E-mail: zander.myburg@fabi.up.ac.za

tance of complex epistatic interactions in the postzygotic isolation of plant species. Interestingly, these interactions have not all been deleterious in nature. Burke *et al.* (1998) found both positive and negative heterospecific interactions in hybrids of iris, and Jiang *et al.* (2000) found that donor chromatin was favored by genetic interactions in some regions of hybrid cotton genomes. Selection of the same donor chromosome segments in independently generated introgression lines has also been reported in hybrids of sunflower (Rieseberg *et al.* 1996).

Over the past decade, the increased accessibility and automation of molecular marker systems such as amplified fragment length polymorphism (AFLP) markers (Vos *et al.* 1995; Myburg *et al.* 2001) have allowed the construction of nearly complete genetic linkage maps in many plant species (Bert *et al.* 1999; Remington *et al.* 1999; Lespinasse *et al.* 2000; Peng *et al.* 2000; Ramsay *et al.* 2000; Van Heusden *et al.* 2000). In studies where genetic mapping was performed in intra- or interspecific hybrids of plant species, these data were found useful to measure rates of introgression of parental alleles into hybrid genomes (Rieseberg *et al.* 1995; Jiang *et al.* 2000) and to follow the transmission of these alleles in advanced hybrid generations (Rieseberg *et al.* 1996). Complete genetic maps of hybrid genomes have also allowed whole-genome surveys of reproductive barriers in natural and experimental hybrid populations (Rieseberg *et al.* 1995; Whitkus 1998; Burow *et al.* 2001; Fishman *et al.* 2001; Harushima *et al.* 2001). These studies demonstrated the relative importance of chromosomal and genic factors as barriers to introgression in plant species (Rieseberg *et al.* 1995), the differential introgression of heterospecific alleles into hybrid genomes (Whitkus 1998; Burow *et al.* 2001), the relative contribution of gametic and zygotic selection to postzygotic isolation (Harushima *et al.* 2001), and the role of heterospecific gene interactions in the postzygotic isolation of plant species (Fishman and Willis 2001; Fishman *et al.* 2001).

With the exception of examples such as the *Ods* hybrid male sterility gene in Drosophila (Ting *et al.* 1998), very little information is available on the type of genes involved in the evolution of reproductive barriers in plants and animals. However, genetic mapping studies such as those described above often provide valuable information on the location and effects of isolation genes in hybrid genomes. The differential introgression of alleles at isolation loci usually causes transmission ratio distortion (TRD) in linked marker alleles, and this information can be used to estimate the position and effects of transmission ratio distorting loci (TRDLs) in parental and hybrid genomes (Fishman *et al.* 2001). Several methods have been proposed to characterize TRDLs. Initial single-marker methods (Hedrick and Muona 1990; Fu and Ritland 1994a; Karkkainen *et al.* 1999) were limited in power, because estimates of distortion effect were confounded by recombination between the markers and TRDLs. Maximum-likelihood interval mapping strategies (Fu and Ritland 1994b; Mitchell-Olds 1995; Cheng *et al.* 1996; Remington and O'Malley 2000) have provided more powerful approaches to locate TRDLs and to simultaneously estimate TRDL effects, but did not accommodate the possibility of multiple, linked TRDLs on a chromosome. Vogl and Xu (2000) proposed a Bayesian approach to locate multiple segregation distorting loci (SDL) per chromosome using molecular markers in a backcross design, while Harushima *et al.* (2001) proposed a multiresponse, nonlinear regression method to estimate the map position and intensity of multiple TRDLs per chromosome in F_2 intercross designs.

Here we present the results of a comparative, genome-wide analysis of postzygotic reproductive barriers in a wide interspecific backcross pedigree of two Eucalyptus tree species, *Eucalyptus grandis* (W. Hill ex Maiden.) and *E. globulus* spp. *globulus* (Labill.). These two well-diverged species belong to different sections of the subgenus Symphyomyrtus and are isolated by strong prezygotic and postzygotic barriers (Griffin *et al.* 2000). We previously used AFLP markers to construct comparative genetic linkage maps of an F_1 hybrid of *E. grandis* and *E. globulus* and of two pure-species parents in a double "pseudo-backcross" mapping pedigree (Myburg *et al.* 2003). Our mapping study revealed a large degree of colinearity between the genomes of the two species and their F_1 hybrid, suggesting that the high levels of postzygotic isolation observed in this cross were not the result of gross chromosomal rearrangements. However, a large proportion of AFLP markers exhibited distorted genotypic ratios in the two backcross families and mapped to defined regions of the parental linkage maps (Myburg *et al.* 2003). This finding suggested the action of strong genic incompatibilities in the *E. grandis* × *E. globulus* backcross pedigree.

In this article, we describe the use of the Bayesian approach developed by Vogl and Xu (2000) to characterize genetic factors (TRDLs) that may underlie postzygotic barriers in the F_1 hybrid of *E. grandis* and *E. globulus*. Our results reveal that positive and negative gene interactions affect the introgression of parental alleles into the genomes of *E. grandis* × *E. globulus* backcross progeny. This study also demonstrates the use of high-throughput AFLP marker analysis to rapidly assess the genetic architecture of postzygotic barriers in interspecific pedigrees of outcrossed plant species.

MATERIALS AND METHODS

Study system: Natural populations of *E. grandis* and *E. globulus* are completely allopatric. *E. globulus* populations range from ~38°S to 43°S and are restricted to eastern Tasmania and the southern extremes of Victoria, while *E. grandis* populations occur between 16°S and 33°S along the eastern coastline of

New South Wales and Queensland (Eldridge *et al.* 1993). High levels of hybrid inviability and hybrid abnormality occur in hybrids of *E. grandis* and *E. globulus* (Griffin *et al.* 2000), which is characteristic of other wide interspecific crosses in this genus (Griffin *et al.* 1988). These phenomena constitute strong postzygotic barriers to hybridization in Eucalyptus (Lopez *et al.* 2000). Structural and physiological incompatibilities in pollen tube growth also form important prezygotic isolation barriers in Eucalyptus. In crosses between species in different subgenera of Eucalyptus, pollen tube growth is arrested in the upper style, which completely prevents hybridization. The frequency of this physiological barrier increases with taxonomic distance between the parental species (Ellis *et al.* 1991). A unilateral structural barrier also exists where pollen tubes from small-flowered species such as *E. grandis* cannot grow the full length of the styles of large-flowered species such as *E. globulus* (Gore *et al.* 1990). In some cases, controlled pollination can be used to overcome prezygotic barriers and produce interspecific progeny (Harbard *et al.* 1999), but strong postzygotic barriers then continue to restrict gene flow. Variation in the strength of these isolating mechanisms may explain the different levels of TRD previously observed in interspecific mapping studies in Eucalyptus (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996; Marques *et al.* 1998).

Generation of mapping pedigree: Details of the mapping pedigree used in this study are presented elsewhere (Myburg *et al.* 2003) and are summarized in Figure 1. An F_1 hybrid individual (BBT01058, Forestal Oriental S.A., Uruguay) was selected for backcrossing from a large F_1 progeny set of *E. grandis* × *E. globulus*. This F_1 hybrid tree (referred to in this article as the "F_1 hybrid") was superior to other F_1 progeny in growth and volume and did not display any of the hybrid abnormalities present in many of its siblings. It also flowered early and produced large numbers of flowers. This allowed the crossing of this individual to unrelated *E. grandis* and *E. globulus* parents to obtain two pseudo-backcross families that were large enough for genetic mapping purposes. True backcrosses to the original parents of the F_1 hybrid were not used, because of the possibility of inbreeding depression in these families. Furthermore, the F_1 hybrid was used as a pollen parent in the backcross to *E. grandis* and seed parent in the backcross to *E. globulus* to avoid the unilateral crossing barrier between large- and small-flowered eucalypt species (Gore *et al.* 1990).

AFLP analysis and linkage map construction: AFLP marker analysis and the construction of framework linkage maps of the F_1 hybrid tree and the two backcross parents are described in detail in Myburg *et al.* (2003). In brief, we used a high-throughput AFLP analysis protocol (Remington *et al.* 1999; Myburg *et al.* 2001) to genotype 186 trees of each backcross family using 24 *Eco*RI/*Mse*I (+3/+3) selective AFLP primer combinations. The AFLP marker data of each backcross were subdivided into two testcross (1:1 segregation) data sets and an intercross (3:1 segregation) data set, according to the two-way pseudo-testcross approach proposed for dominant markers in outbred pedigrees (Grattapaglia and Sederoff 1994).

The testcross (1:1) markers segregating in the two parents of each backcross were used to construct four framework linkage maps. A paternal linkage map of the F_1 hybrid and maternal linkage map of the *E. grandis* backcross parent were constructed in the *E. grandis* backcross family, while a maternal linkage map of the F_1 hybrid and a paternal linkage map of the *E. globulus* backcross parent were constructed with data of the *E. globulus* backcross family. All four maps were aligned using testcross markers shared between the two maps of the F_1 hybrid and intercross markers shared between the maps of the F_1 hybrid and that of the backcross parents.

The species origin of markers in the maps of the F_1 hybrid was inferred by genotyping the original *E. grandis* seed parent of the F_1 hybrid (tree G50, Figure 1). Since the two linkage phases in the maps of the F_1 hybrid represent AFLP markers amplified from either the *E. grandis* or the *E. globulus* chromosome of each homologous pair, we were able to infer the species origin of all the testcross markers in the F_1 hybrid. This allowed us to determine the rate of transmission of donor (heterospecific) alleles from the F_1 hybrid to backcross progeny throughout all linkage groups.

Detection of TRD and candidate TRDL regions: Before analysis of TRD in the maps of the F_1 hybrid, all marker data were recoded to reflect the presence (1) or absence (0) of the *E. globulus* marker alleles. The phase of linkage was therefore consistent across all of the linkage groups in the maternal and paternal maps of the F_1 hybrid. This inference was not possible for the maps of the pure species parents, because the respective grandparents were not available for genotyping.

The observed genotypic ratio of each testcross AFLP fragment was compared to the expected Mendelian ratio of 1:1. Chi-square tests were performed for each of these markers at the 0.05 level of significance.

Distorted genotypic ratios of individual markers can arise purely by chance through sampling of a finite number of gametes or through linkage to a hybrid fitness factor. The identification of candidate regions containing hybrid fitness factors (or TRDLs) therefore necessitates the use of a genome-wide significance threshold. However, TRD of marker loci linked to viability or sterility factors should be limited by recombination to subchromosomal blocks of linked markers. A genome-wide significance threshold that is based on the effective number of independent tests should therefore be based on the number of independent genomic regions in each parental linkage map. In this study, the testcross markers were distributed across 11 linkage groups within each parental map. Assuming that each linkage group corresponds to one chromosome ($n = 11$ in Eucalyptus) and that each chromosome contains at least two independent regions (the average length of linkage groups was 125 cM; Myburg *et al.* 2003), at least 22 independent genomic regions are expected. A threshold of at least $0.05/22 \approx 0.002$ would therefore be required to obtain a genome-wide error rate of $\alpha = 0.05$. However, to ensure that regions with weak TRDLs are included and to evaluate epistatic interactions among as many putative TRDLs as possible, we applied a threshold of $\alpha = 0.01$ and recorded all distorted regions with two or more distorted markers.

Mapping and estimation of TRDL effects: The Bayesian Markov chain Monte Carlo (MCMC) method of Vogl and Xu (2000) was used for multipoint analysis of TRDLs in the four testcross marker sets. The model used for TRDL analysis assumed inbred parents and a backcross design. This model is compatible with the two-way pseudo-testcross approach (Grattapaglia and Sederoff 1994) used to analyze the testcross marker data in this outbred pedigree. However, it is important to note that all three parents were highly heterozygous and that up to four different alleles were segregating at each locus in the backcross families. Allelic effects on hybrid fitness at any particular locus were therefore averaged over the allelic effects of the other parent.

For the MCMC procedure (described in detail in Vogl and Xu 2000), we set a Poisson prior mean of one and a maximum of six TRDLs per linkage group. An error term of 2% was included in the model to account for possible scoring error in the marker data. The chain length was set to 3500 and every tenth cycle was recorded. This chain length was sufficiently accurate as determined by trial runs with longer chain lengths (data not shown). The number of TRDLs, TRDL positions, and TRDL effects reported for each linkage group were treated as samples from the joint posterior distribution.

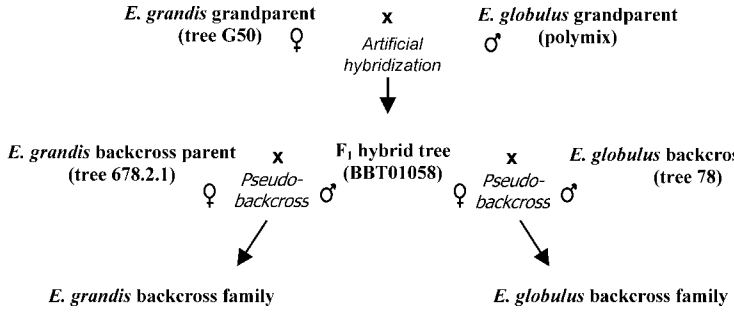

FIGURE 1.—"Pseudo-backcross" mating scheme used to generate the mapping populations. The direction of crosses (seed and female parents) is indicated by symbols underneath each parent.

Evaluation of epistatic interactions between TRDLs: To test for epistatic interactions between pairs of TRDLs, we selected the most distorted marker in each candidate region as representative of the putative TRDL. Contingency tests of all the two-locus TRDL allele combinations within each parental map were performed using Fisher's exact test as implemented in the JMP statistical software package (SAS Institute, Cary, NC). We also tested for interaction between distorted loci in the two parental maps of each backcross. Only unlinked marker pairs were considered in the tests for epistasis. An experimentwise significance threshold was set to account for the number of pairwise tests performed in each backcross family.

We evaluated the overall fitness effect of epistatic interactions among all putative TRDLs in each backcross family using a previously proposed regression method (FU and RITLAND 1996; REMINGTON and O'MALLEY 2000). Briefly, the nearest marker to each putative TRDL was selected and the presence or absence of the least frequent marker allele (representing the negative TRDL effect) was recorded for all backcross individuals and all TRDLs. We then counted the observed number of individuals (n_i) with i negative TRDL alleles at a total of m TRDLs in each backcross family. The expected frequency of individuals with i negative TRDLs was estimated from the binomial expectation without selection. The ratio of observed (n_i) to expected number of individuals with i out of m possible negative TRDL alleles is

$$f(i) = n_i \Big/ \binom{m}{i}\left(\frac{1}{2}\right)^i.$$

The log-transformed ratio of the observed/expected frequency $f(i)$ was then regressed on i. Two models for viability selection were compared within each backcross family. The simple linear model, $\ln f(i) = a + b_1 i$, was compared to $\ln f(i) = a + b_1 i + b_{11} i^2$, a model with a quadratic term to represent all possible pairwise interactions. In the absence of interaction between the TRDLs, a simple linear relationship is expected, while considerable negative interactions among TRDLs should result in a significant negative coefficient for the interaction term in the quadratic regression model (REMINGTON and O'MALLEY 2000).

RESULTS

Low fitness of *E. grandis* × *E. globulus* hybrids: The F_1 hybrid (tree BBT01058, Figure 1) selected for backcrossing was a rare superior individual in a large F_1 polymix progeny set. A total of 45 polymix families were produced by controlled pollination of selected *E. grandis* mother trees with two 10-tree pollen mixes of *E. globulus* parents. Only 4.4% of F_1 seed germinated and many weak and abnormal phenotypes were observed at the seedling stage. Only 3.2% of the surviving F_1 hybrid trees were sufficiently vigorous and normal to be advanced to clonal evaluation, and of these, only 9% met the selection criteria for commercial forest tree clones. Considerable variation in F_1 hybrid viability was observed among different combinations of *E. grandis* mother trees and *E. globulus* pollen mixes, which suggested that variation in hybrid combining ability existed within both species (GRIFFIN *et al.* 2000).

F_1 hybrid individual (BBT01058) was cross-fertile with both backcross parents, but the germination rate of the F_2 backcross seed was only ~50%. A greater proportion of normal plants were observed in the F_2 generation than in the F_1 generation. However, abnormal phenotypes such as dwarfs, abnormal leaf pigmentation, abnormal rooting, deformed stems, and deformed shoots were observed in many of the backcross individuals (data not shown).

Comparative genetic linkage maps and TRD of testcross AFLPs: Linkage analysis of >800 testcross and intercross AFLP fragments in each of the backcross families allowed us to construct comparative maps (Figure 2) of all 11 linkage groups expected in this mapping pedigree ($2n = 22$ in Eucalyptus). No crossover of marker positions or large clusters of markers, indicative of gross chromosomal rearrangements, were detected at the resolution of the four parental maps (Figure 2). In addition, recombination was generally not suppressed in the F_1 hybrid genome relative to that of the pure species as would be expected if major chromosomal incompatibilities existed (MYBURG *et al.* 2003).

A total of 704 testcross markers segregating in the *E. globulus* backcross family and 650 in the *E. grandis* backcross family were evaluated for TRD (Table 1). Approximately 27 and 30% of testcross markers were distorted at the 0.05 level of significance in the *E. globulus* and *E. grandis* backcross families, respectively. At the 0.002 level of significance (approximate genome-wide α of 0.05), ~13% of the testcross markers in each backcross family were significantly distorted. More testcross markers were distorted than could be explained by chance at both levels of significance.

Distribution of TRD in the comparative maps: The distribution of TRD in the comparative framework maps of the F_1 hybrid and the two backcross parents is shown in Figures 2 and 3. For the purpose of discussion, linkage

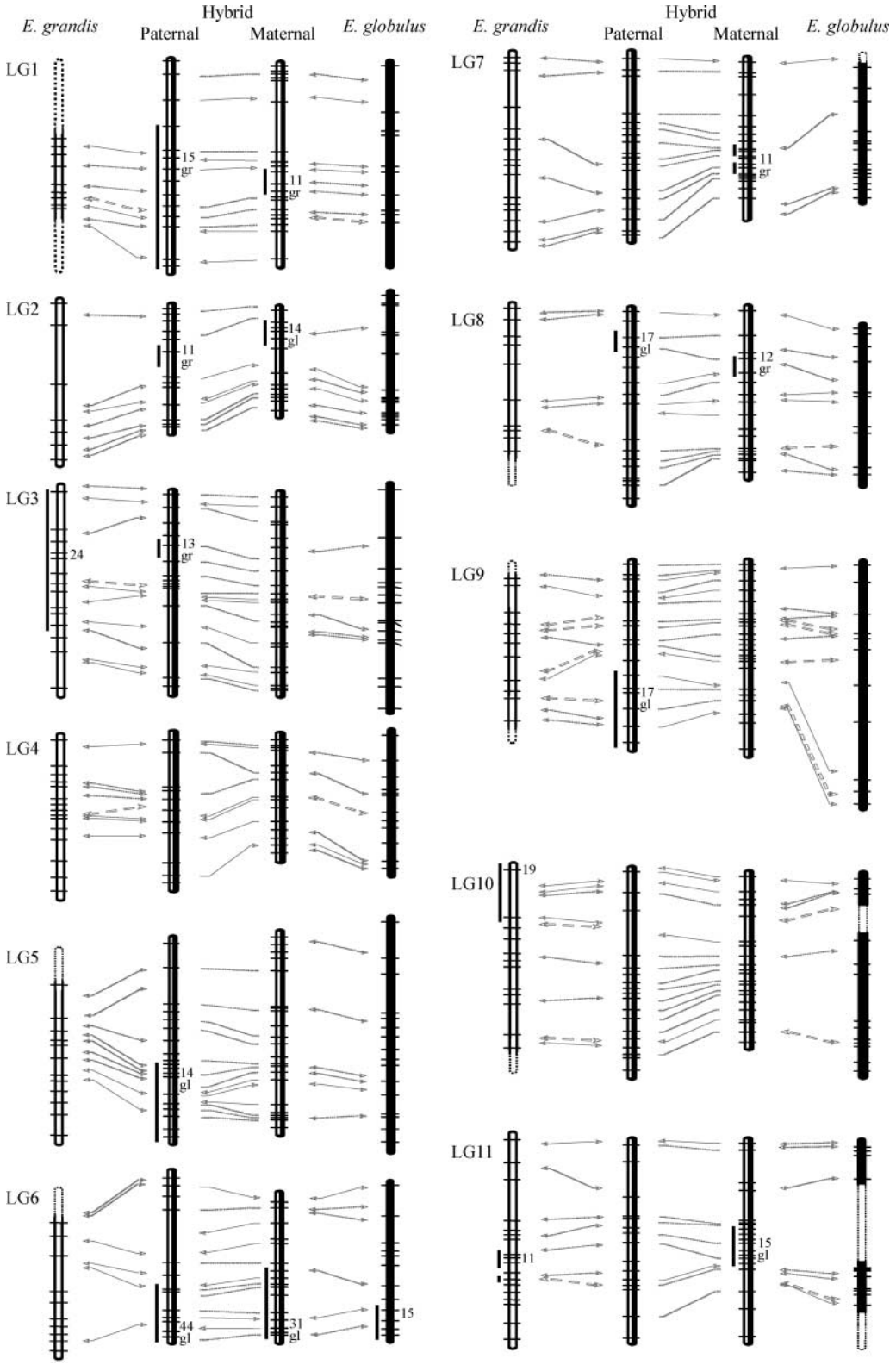

FIGURE 2.—Summary of comparative maps of *E. grandis*, *E. globulus*, and their F_1 hybrid showing the location of distorted map regions. *E. globulus* maps and linkage phases are shown in black and *E. grandis* maps and linkage phases in white. Vertical bars to the left of the linkage maps indicate regions that were distorted at the 0.01 level of significance (listed in Table 2). The value to the right of each distorted region shows the maximum percentage of deviation from a 1:1 genotypic ratio observed in that region. In the maps of the F_1 hybrid, the species origin of the favored marker allele is shown for each distorted region. The positions of shared testcross and intercross markers in the four framework AFLP maps are indicated with gray dotted lines and dotted arrows, respectively. Solid arrows and heavy dashed arrows indicate the positions of markers that segregated in both backcrosses in testcross/intercross or intercross/intercross configuration. Marker names and linkage phase details are reported elsewhere (MYBURG *et al.* 2003).

groups of the maternal and paternal linkage maps of the F_1 hybrid are indicated below by "m" and "p," while linkage groups of the *E. grandis* and *E. globulus* backcross parents are indicated by "gr" and "gl," respectively.

Regions of the parental maps that contained two or more distorted markers were all unidirectionally distorted with respect to linkage phase (Figure 3). These regions were all distorted with P values <0.01 and they were therefore recorded as candidate regions for TRDL analysis (Table 2). Although this pattern of TRD could have resulted from sampling bias in some regions, it is also possible that these regions contained weak hybrid

TABLE 1

Observed proportions of testcross AFLP loci with distorted genotypic ratios

	0.05 level[a]		0.002 level[a]	
Backcross family	No.	%	No.	%
E. globulus backcross				
Total no. of testcross AFLPs	704	100		
No. distorted in the maternal map of the F_1 hybrid	140/457	30.6	76/457	10.8
No. distorted in the *E. globulus* backcross parent	51/247	20.6	15/247	2.1
E. grandis backcross				
Total no. of testcross AFLPs	650	100		
No. distorted in the paternal map of the F_1 hybrid	104/365	28.5	37/365	5.7
No. distorted in the *E. grandis* backcross parent	88/285	30.9	46/285	7.1

[a] Level of significance (α) used to evaluate the departure from an expected 1:1 allelic transmission ratio of testcross markers.

fitness factors. The distortion of the same genomic regions in different genetic backgrounds (*e.g.*, LG2, LG3, and LG6; Figure 2) provided additional evidence for the presence of TRDLs in these regions. We excluded individually distorted markers at the ends of linkage groups, especially where the direction of allelic bias was opposite to that of neighboring markers (*e.g.*, LG8gl and LG7p, Figure 3), due to the less stringent criteria associated with the placement of framework markers at terminal positions of linkage groups.

In all of the distorted map regions, TRD was highly directional with respect to allelic origin. In some cases (*e.g.*, LG6gl, LG11m, LG1p, LG3gr, and LG11gr) directional distortion extended over almost the full length of the linkage group, while in other cases (*e.g.*, LG2m, LG2p, LG8m, LG8p, and LG10gr), it was restricted to a smaller region of the linkage group (Figure 3). Generally, marker loci that were not distorted at the 0.05 level of significance exhibited random fluctuation in the direction of deviation from a 1:1 segregation ratio (*e.g.*, LG10gl, Figure 3).

The most severely biased map region (LG6p and LG6m, Figure 3) exhibited TRD in both backcrosses. A marked excess of *E. globulus* marker alleles occurred in this region of the F_1 genome in both backcross families. Up to 80.6% of individuals in the *E. globulus* backcross and 94.2% of individuals in the *E. grandis* backcross carried *E. globulus* alleles in this region of LG6 (Table 2).

The ability to infer the species identity of the linkage phases of the maps of the F_1 hybrid allowed us to distinguish between donor (heterospecific) and recurrent (conspecific) alleles in both genetic backgrounds. On the basis of this inference, we found that donor alleles inherited via the F_1 hybrid were favored at approximately the same frequency as recurrent alleles in the two backcrosses. For example, in the paternal map of the F_1 hybrid donor (*E. globulus*) alleles were in excess on LG5p, LG6p, LG8p, and LG9p, while recurrent (*E. grandis*) alleles were favored on LG1p, LG2p, and LG3p. In the maternal map of the F_1 hybrid, three of the six distorted regions (LG1m, LG7m, and LG8m) were biased in favor of donor (*E. grandis*) alleles (Figure 3).

Multipoint comparative mapping of TRDLs: The Bayesian MCMC method of Vogl and Xu (2000) provided an efficient approach to estimate the positions and effects of putative TRDLs in distorted map regions (Table 2). It also allowed us to model random sources of distortion such as marker scoring error. The posterior distribution of TRDL positions and effects (Table 2), sampled through 3500 iterations of the MCMC procedure, provided estimates of the location and strength of TRDLs in the four linkage maps. In most cases, the distribution of posterior data points corresponded well with the magnitude and peaks of TRD in the raw marker data (Figure 3).

The posterior probability of TRDLs in each distorted map region was estimated by the frequency of TRDL detection in the map region during MCMC iterations (Vogl and Xu 2000). The posterior probabilities of TRDLs calculated in this way ranged from 0.11 to 1.0 (Table 2). Map regions that were distorted at $P < 0.0001$ contained TRDLs with posterior probabilities of >90% (Table 2). This provided an empirical genome-wide significance threshold for our data set. The five TRDLs with posterior probabilities of <0.5 (Table 2) are more likely to be the result of spurious TRD than the action of a TRDL. No TRDL (posterior probability <1%) was detected in the distorted region on LG11gr. This may be due to the low magnitude of distortion in this linkage group and the absence of a clearly defined peak in distortion (Figure 3).

The posterior means of TRDL effects ranged from 8.2 to 33.4% in the genomes of the three parents (Table 2). This corresponded to an excess or deficiency in individual TRDL alleles of 16.4–66.8%. In a backcross design, where allelic frequencies equal genotypic fre-

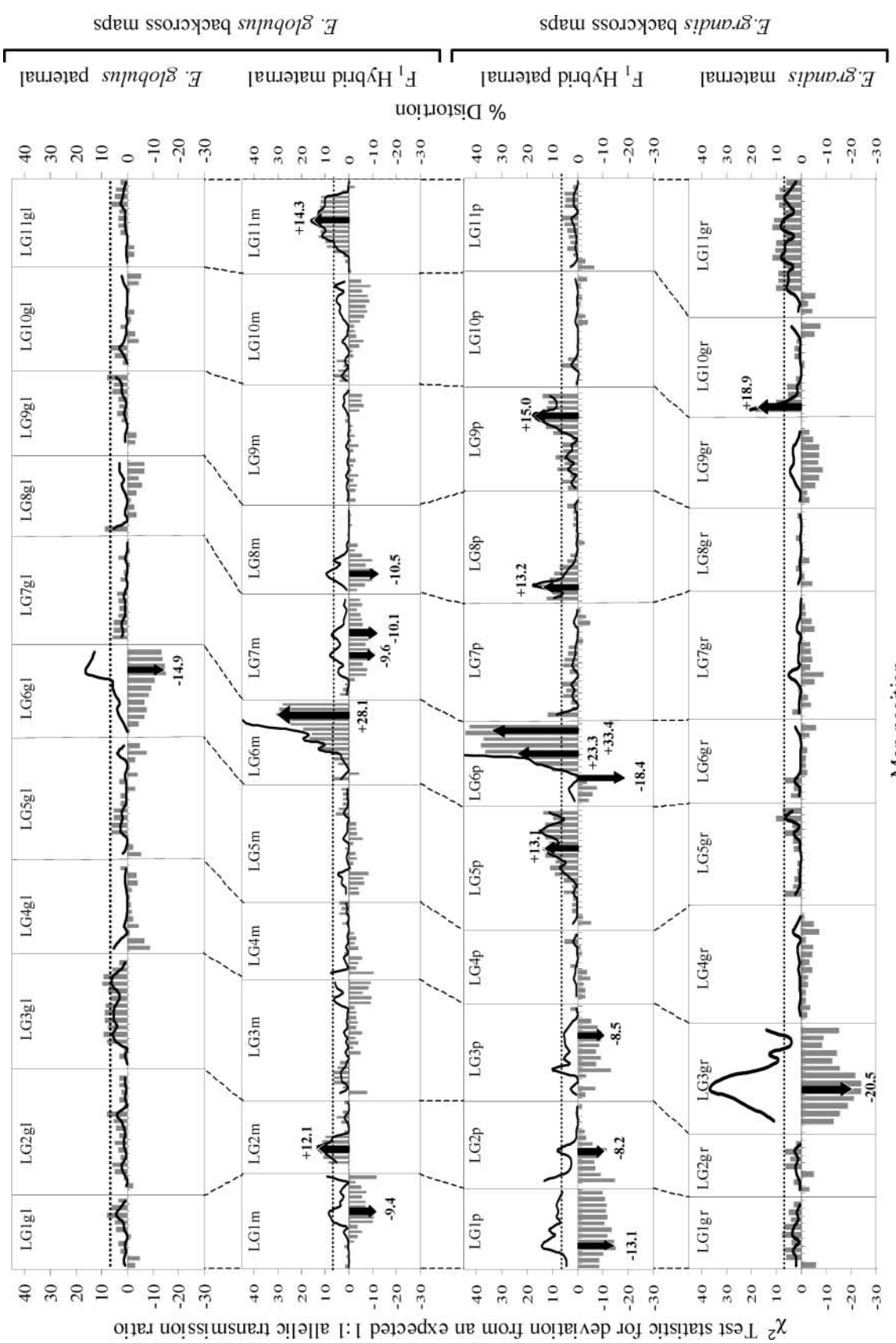

FIGURE 3.—Distribution, magnitude, and direction of transmission ratio distortion in the linkage maps of *E. grandis*, *E. globulus*, and the F_1 hybrid. The 11 linkage groups of each parental map are shown end to end. The solid lines plot the chi-square statistic for deviation from an expected 1:1 genotypic ratio for each testcross marker (primary *y*-axis). Horizontal dotted lines show the critical value (6.64) of the chi-square statistic at the 0.01 level of significance. The gray vertical bars represent the direction and percentage of distortion, *i.e.*, (allele frequency − 0.5) × 100%, at the marker locus (secondary *y*-axis). Phase assignment in the pure species maps is arbitrary from one linkage group to another. However, in the maps of the F_1 hybrid, marker data were recoded so that the direction of distortion represents that of the *E. globulus* allele at all marker loci. Solid arrows indicate the positions and average effects (indicated) of TRDLs estimated by the Bayesian MCMC procedure. Note that intermarker distances are not shown proportionally.

TABLE 2

Summary of distorted map regions and posterior distribution of putative TRDLs detected by the Bayesian Markov chain Monte Carlo procedure

Map	Marker[a]	Linkage group	Map position (cM)	*P* value[b]	Deviation from 1:1[c] (%)	TRDL no.	TRDL position (cM)		TRDL effect[d] (%)		Relative viability[e]	Posterior probability[f]
							Mean	STD	Mean	STD		
E. globulus backcross parent	ACA/CCT-354	6gl	94.4	6.16×10^{-05}	−15.1	1	97.7	4.3	14.9	7.9	0.54	1.00
F_1 hybrid maternal	AAG/CTG-048	1m	84.3	6.05×10^{-03}	+10.5	1	89.9	2.2	−9.4	4.9	0.68	0.11
F_1 hybrid maternal	ACA/CCA-164	2m	15.4	2.97×10^{-04}	−13.6	2	14.2	2.4	12.1	8.5	0.61	0.67
F_1 hybrid maternal	ATC/CCA-198	6m	91.4	6.04×10^{-16}	−30.6	3	91.2	1.4	28.1	5.7	0.28	1.00
						4	60.6	2.7	−9.6	4.9	0.68	0.22
F_1 hybrid maternal	AAA/CGA-114	7m	63.1	4.68×10^{-03}	−11.1	5	73.3	2.7	−10.0	4.6	0.67	0.13
F_1 hybrid maternal	AAC/CCC-397	8m	73.2	6.35×10^{-03}	−10.3	6	36.1	5.6	−10.5	5.2	0.65	0.56
F_1 hybrid maternal	ACG/CCA-271	11m	78.0	6.17×10^{-05}	−24.2	7	75.6	4.3	14.3	5.9	0.56	0.89
F_1 hybrid paternal	ACC/CCA-355	1p	57.5	2.10×10^{-04}	−14.9	1	58.9	6.1	−13.1	6.1	0.58	0.55
F_1 hybrid paternal	AAG/CGG-165	2p	0.0	2.98×10^{-04}	−14.9	2	29.2	5.1	−8.2	10.8	0.72	0.36
F_1 hybrid paternal	AAG/CCT-430	3p	32.3	1.54×10^{-03}	+13.2	3	70.8	16.2	−8.5	5.4	0.71	0.38
F_1 hybrid paternal	ACC/CCA-298	5p	103.8	1.10×10^{-04}	−15.6	4	94.8	7.1	13.1	7.1	0.58	0.90
F_1 hybrid paternal	AAA/CCG-279	6p	113.1	3.65×10^{-28}	+44.2	5	49.3	7.4	−18.4	6.7	0.46	0.90
						6	92.2	2.4	23.3	7.6	0.36	0.81
						7	115.0	2.1	33.4	9.5	0.20	1.00
F_1 hybrid paternal	ACA/CCA-087	8p	19.1	2.60×10^{-05}	+17.3	8	17.2	6.4	13.2	6.0	0.58	0.84
F_1 hybrid paternal	ACT/CCA-274	9p	91.5	3.14×10^{-05}	+16.7	9	91.5	5.3	15.0	5.8	0.54	0.85
E. grandis backcross parent	ACA/CCA-129	3gr	34.1	1.70×10^{-09}	+24.2	1	34.3	4.9	20.5	9.2	0.42	1.00
E. grandis backcross parent	ATC/CCA-077	10gr	0.0	6.62×10^{-06}	+18.5	2	7.7	5.3	18.9	6.9	0.45	1.00
E. grandis backcross parent	ACT/CCA-294	11gr	105.0	4.70×10^{-3}	+11.4	—	—	—	—	—	—	—

[a] The most distorted marker in each distorted map region. Only map regions with two or more distorted markers at $\alpha = 0.01$ are listed. Marker names indicate the (+3/+3) selective nucleotides of the *Eco*RI/*Mse*I primer combination and the size in base pairs of the AFLP fragment.

[b] *P* value of the calculated χ^2 statistic based on an expected transmission ratio of 1:1.

[c] Percentage of deviation from 1:1 ratio relative to the band present class of the testcross marker.

[d] The direction of TRDL effects in the backcross parents is arbitrary from one linkage group to the next. In the F_1 hybrid maps, all TRDL effects are relative to the *E. globulus* linkage phase.

[e] Ratio of the frequency of the less frequent TRDL allele to the more frequent TRDL allele.

[f] Posterior probability of a TRDL in the distorted map region, calculated as the frequency of TRDL detection in the distorted map region in 350 recorded iterations of the MCMC procedure.

TABLE 3

Pairwise epistatic interactions between TRDLs

Family	TRDL marker	LG	Map position (cM)	*P* (Fisher's exact test)	Deficient genotypic combinations[a]	Deviation from expected count (%)
E. globulus	AAA/CGA-114	7m	63.1	0.0007*	gr-gr	−30.0
Backcross	ACG/CCA-271	11m	78.0		gl-gl	−23.6
E. grandis	ACC/CCA-355	1p	57.5	0.0048	gl-r	−44.8
Backcross	ATC/CCA-077	10gr	0.0		gr-c	−11.2
E. grandis	ACC/CCA-355	1p	57.5	0.0332	gl-gl	−31.3
Backcross	AAG/CGG-165	2p	0.0		gr-gr	−8.8

* Significant at an experimentwise significance threshold of $\alpha = 0.05$ ($P = 0.0024$ and 0.0014 for the *E. globulus* and *E. grandis* backcross families, respectively).

[a] Two-locus genotype of the genotypic classes that contained fewer than expected individuals. The species origin of alleles at loci in the maps of the F_1 hybrid is indicated by "gl" for *E. globulus* and "gr" for *E. grandis*. Alleles in the maps of the backcross parents are denoted "c" for markers in coupling and "r" for markers in repulsion to the band present class.

quencies, TRDL effects can also be expressed as the relative viability (t) of gametes or zygotes with alternative genotypes and $0 < t < 1$ (Cheng *et al.* 1998). On the basis of this method, the relative viabilities of gametes or zygotes with alternative TRDL alleles ranged from 0.20 to 0.72 (Table 2).

The maximum number of TRDLs inferred on a single linkage group was three (for LG6p). Two separate TRDLs were inferred on LG6p at 93 and 166 cM with mean effects of 23.3 and 33.4%, respectively (Table 2). These TRDLs accounted for the severely distorted region at the lower end of this linkage group. A third TRDL of opposite effect was inferred at 51 cM. This TRDL may account for the abrupt change in the magnitude of distortion in the middle of LG6p. A more gradual change in distortion and more defined peak in distortion were observed in LG6m. A single TRDL with a mean effect of 28% was inferred in this linkage group (at 92 cM). The only TRDL detected in the map of the *E. globulus* backcross parent was located at a similar position (96 cM) on LG6gl (Figure 3).

Apparent homologous TRDLs were also detected on four other syntenic linkage groups. LG1m and LG1p were both biased in favor of *E. grandis* alleles (Figure 3). The same region of LG2m and LG2p contained an excess of alleles of the respective recurrent parents. Similar regions of LG8m and LG8p were distorted and in both cases the TRDL allele of the respective donor parent was favored (Figure 3). Finally, putative TRDLs were detected on LG3p and LG3gr in the two parental maps of the *E. grandis* backcross family.

Epistatic interactions among TRDLs: Contingency tests for interaction between TRDLs revealed one interaction in the *E. globulus* backcross and two interactions in the *E. grandis* backcross at the 0.05 level of significance (Table 3). Only one interaction, that between distorted markers on LG7m and LG11m, was significant ($P = 0.0007$) at the experimentwise significance level (Table 3). The frequency of *E. globulus* backcross progeny with the *E. grandis* donor allele at both marker loci, AAA/CGA-114 on LG7m and ACG/CCA-271 on LG11m, was 30% lower than expected, while 23.6% fewer than expected individuals carried *E. globulus* alleles at both loci.

We also tested the cumulative fitness effect of interactions among TRDLs segregating in each backcross family, using a modified version of the method of Fu and Ritland (1996). The frequency distribution of the observed number of individuals (n_i) with i out of m possible negative TRDL alleles (Figure 4A) showed that the majority of genotyped individuals of the *E. grandis* and *E. globulus* backcross families carried a combination of two, three, or four negative TRDL alleles. Only 2% of individuals in each backcross family had the favored allele at all TRDLs. In both backcrosses, the linear regression model provided very good fits with the observed data (model $P < 0.0001$, $R^2 > 0.96$; Figure 4B). Furthermore, the coefficient of the quadratic term was not significant in either backcross family ($P = 0.55$ and 0.46 in the *E. globulus* and *E. grandis* backcross families, respectively), suggesting that the overall fitness effect of deleterious epistatic interactions among TRDLs was very small.

DISCUSSION

Comparative genetic maps of F_1 hybrids and their parental species can be used as powerful tools to study postzygotic reproductive barriers in plant species (Rieseberg *et al.* 2000). These maps contain information on the genomic composition and relative viability of a wide array of F_2 hybrid genotypes produced when the two parental genomes are recombined in the F_1 hybrid. We have constructed the first detailed genetic linkage maps

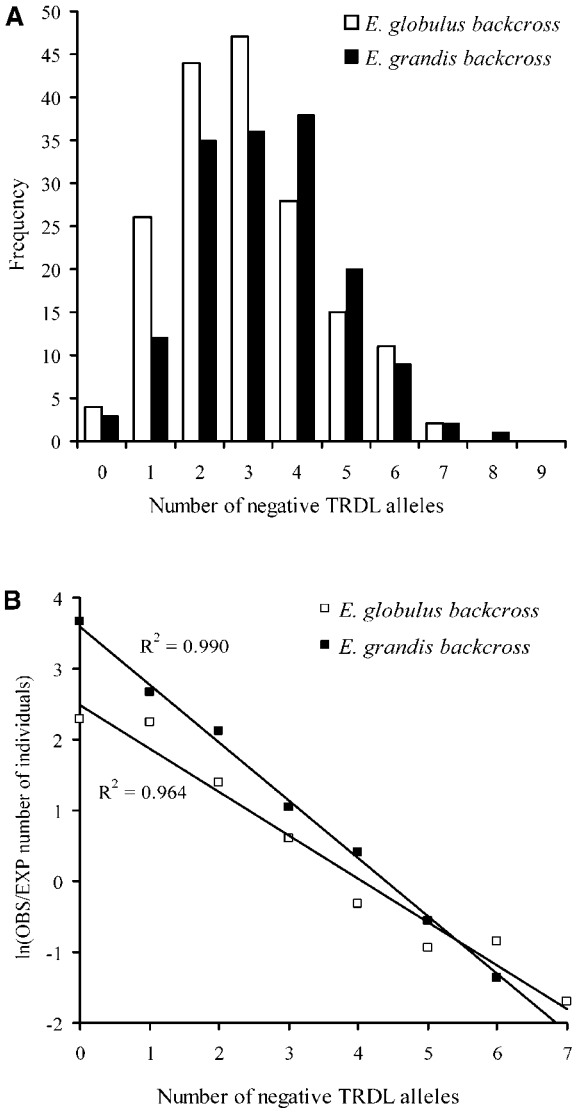

FIGURE 4.—Statistical test of the contribution of epistatic interactions of TRDLs to reduced hybrid fitness. (A) Frequency distribution of *i*, the number of negative TRDL alleles carried by genotyped individuals of the *E. grandis* and *E. globulus* backcross families. (B) Best-fit plots of the linear regression of the log-transformed ratio of observed *vs.* expected frequency of individuals with *i* negative TRDL alleles in each backcross family, on the number of negative TRDL alleles (*i*) observed in each backcross individual.

of a wide interspecific hybrid of two Eucalyptus tree species (MYBURG *et al.* 2003) and used this information to perform a whole-genome analysis of postzygotic barriers between *E. grandis* and *E. globulus*.

Genetic architecture of postzygotic barriers between *E. grandis* and *E. globulus*: Postzygotic barriers between *E. grandis* and *E. globulus* manifest in this interspecific backcross mapping pedigree as regions of TRD in the genetic maps of the F_1 hybrid and the two backcross parents. Allelic transmission ratios in these three trees may be altered from an expected 1:1 ratio (for testcross AFLPs) by selection on their gametes or by selection on the F_2 zygotes produced after fusion of these gametes. The fertilization ability of F_1 gametes and viability of F_2 zygotes may in turn be affected by genic incompatibilities between the genomes of *E. grandis* and *E. globulus* or by chromosomal rearrangements that result in abnormal meiotic products in the F_1 hybrid. The relative contribution of these two sources of hybrid incompatibility to postzygotic isolation of *E. grandis* and *E. globulus* is unknown and constitutes an important question addressed in this study.

The results of our comparative mapping effort (Figure 2 and MYBURG *et al.* 2003) generally support genic incompatibilities rather than chromosomal rearrangements as the major mechanism underlying postzygotic isolation of *E. grandis* and *E. globulus*. We did not detect any gross chromosomal rearrangements among the genomes of *E. grandis*, *E. globulus*, and their F_1 hybrid in the comparative AFLP marker maps of this pedigree. Furthermore, the directional pattern of TRD observed in the parental maps (Figure 3) was consistent with the presence of genic factors (TRDLs) altering the transmission ratio of linked marker alleles. This result is in agreement with recent evidence that supports the importance of genic incompatibilities in other hybrid plant genomes surveyed by molecular marker analysis (FISHMAN and WILLIS 2001; HARUSHIMA *et al.* 2001). It does not, however, rule out the possibility that smaller rearrangements, not detected by dominant AFLP marker analysis in our comparative mapping study, also contribute to postzygotic isolation of *E. grandis* and *E. globulus*.

Before interpreting the patterns of TRD observed in this study, it is important to be aware of potential nongenetic sources of TRD. These include sampling bias, reduced marker penetrance, comigration of DNA fragments during electrophoresis, and scoring error. Except for sampling bias, which affects groups of linked markers through random over- or underrepresentation of chromosome arms, all of these sources of error affect the genotypic ratios of individual markers in a random fashion and should therefore not have resulted in directional distortion of groups of linked markers in our linkage maps. In addition, we previously estimated scoring error to be <2% in our mapping data (MYBURG *et al.* 2001), which is well below the magnitude of error required to result in significant TRD in this mapping pedigree.

It is also important to be aware of the limitations of our experimental design for inferring the genetic mechanisms (*e.g.*, gametic *vs.* zygotic selection) underlying the reproductive barriers (TRDLs) observed in this study. The unidirectional crossing barrier between *E. grandis* and *E. globulus* prevented the construction of reciprocal backcrosses. The fertilization ability of the paternal and maternal gametes of the F_1 hybrid were consequently evaluated in two different hybrid genetic backgrounds, which precluded us from directly discriminating between sex-specific gametophyte selection and

zygotic selection in backcross progeny. The use of dominant AFLP markers and a backcross design also prevented us from distinguishing between these two sources of TRD. This issue can be addressed in the future by using codominant microsatellite markers and an outcrossing mapping model. A genome-wide set of highly informative microsatellite markers is currently under development in Eucalyptus and will be extremely useful for comparative analysis of postzygotic isolation barriers (Brondani *et al.* 2002). Despite these limitations, the use of high-throughput AFLP marker analysis and the pseudo-backcross design did allow us to rapidly achieve high marker coverage of the genomes of the F_1 hybrid and the backcross parents, align these maps through shared marker polymorphisms, and perform comparative, genome-wide scans for postzygotic reproductive barriers in the three genomes studied here.

To interpret the patterns of TRD observed in the F_1 hybrid, it is useful to consider (a) whether a map region was distorted in both backcrosses (LGs 1, 2, 6, and 8) or in only one backcross (LG3p, LG5p, LG7m, LG9p, and LG11m), (b) whether the distortion was in the same direction (LGs 1 and 6) or different directions (LGs 2 and 8) in the two backcrosses, and (c) whether the same region was also distorted in either of the backcross parents (LG3p or LG6m). These categories of TRDLs have to be matched to the patterns of segregation expected for different selective mechanisms. For example, genetic factors that caused differential fertility of F_1 hybrid pollen should segregate in the paternal map of the F_1 hybrid (*e.g.*, LG3p, LG5p, and LG9p; Figure 3), while factors that affected the fertilization of F_1 hybrid ovules should segregate in the maternal map (*e.g.*, LG7m and LG11m, Figure 3). Both types of gametic factors may exhibit interactions with loci in the backcross parent maps, if these loci affect interspecific gamete compatibility and they are heterozygous in both parents. In contrast, genetic factors that influenced the viability of F_2 zygotes, the germination rate of backcross seed, or early seedling survival may segregate in any of the parental maps. These factors may also be part of heterospecific interactions between *E. grandis* and *E. globulus* genes in the F_1 hybrid or "donor-recurrent" interactions between factors in the F_1 hybrid and the backcross parents. Clearly, some of the categories of TRDLs described above can be associated with more than one selective mechanism, depending on the mode of action and heterozygosity of the loci involved. For example, TRDLs with fitness effects in only one of the two backcrosses (*e.g.*, LG7m, LG11m, LG5p, and LG9p; Figure 3) may indicate male- or female-specific selection of F_1 gametes or may result from background-specific selection of TRDL alleles in F_2 zygotes. It is also possible that some of the TRDLs detected in the same regions of the paternal and maternal maps of the F_1 hybrid (*e.g.*, LG1, LG2, LG6, and LG8) represent nonhomologous genetic factors segregating in a background-specific manner.

Pollen tube growth characteristics may be a source of very strong gametic selection in this study. Genes that determine differences in pollen tube growth, a strong unilateral crossing barrier between these two species, should segregate in the F_1 hybrid and may cause severe TRD in the paternal map. *E. globulus* alleles of these genes may confer a selective advantage to their gametes due to the fact that pollen tubes of *E. globulus* grow farther and faster than those of the small-flowered *E. grandis*. Four TRDLs in the paternal map (LG5p, -6p, -8p, and -9p; Figure 3) were distorted in favor of *E. globulus* alleles. However, only the TRDLs on LG5p and LG9p were unique to the paternal map and fit the predicted pattern of male-specific gametophyte selection.

One of the most remarkable results of our study is the extent to which donor alleles in the F_1 hybrid were favored in the two backcross families. Selection against donor alleles in the "recurrent" genetic backgrounds (*e.g.*, LG2m and LG2p, Figure 3) was expected to be the dominant force shaping the genomic composition of *E. grandis* × *E. globulus* backcross progeny. However, we found that four out of seven distorted regions in the paternal map of the F_1 hybrid were biased toward the donor parent (*E. globulus*) in the backcross to *E. grandis*, while three of the six distorted regions in the maternal map of the F_1 hybrid were biased in favor of the donor parent (*E. grandis*) in the backcross to *E. globulus* (Figure 3). Selection for donor genetic material has also been reported in hybrids of other plant species such as iris (Burke *et al.* 1998), cotton (Wang *et al.* 1995; Jiang *et al.* 2000), and sunflower (Rieseberg *et al.* 1996). Rieseberg *et al.* (2000, p. 211) postulated that donor alleles favored in hybrid genetic backgrounds might represent examples of "selfish" genes, *i.e.*, "genes that enhance the success of gametes they inhabit even if they pose a significant fitness cost during the diploid phase of the life cycle." However, this phenomenon may also be explained by the alleviation of genetic load by donor alleles in the recurrent genetic backgrounds. Forest tree species carry high amounts of genetic load (Hardner and Potts 1997; Remington and O'Malley 2000) and it may not be unexpected to find that donor alleles confer increased fitness in hybrids of tree species. If these are indeed examples of genetic load exposed by the introgression of donor alleles, genetic mapping studies in other crosses of *E. grandis* and *E. globulus* should reveal whether these loci are fixed in *E. grandis* and *E. globulus*.

It is not clear how many of the reproductive barriers in the F_1 hybrid represent fixed genetic differences that have accumulated between the genomes of *E. grandis* and *E. globulus*. Fixed hybrid fitness loci should all be heterozygous in the F_1 hybrid, segregate in both backcrosses, and exhibit a deficiency of the respective donor alleles in at least one of the two backcrosses (consistent with the Dobzhansky-Muller model of negative hetero-

specific interactions). Only one genomic region in the F_1 hybrid (LG2m and LG2p, Figure 3) exhibited a deficiency of donor alleles in both backcrosses. Three other regions (LG1p, LG3p, and LG11m; Figure 3) were deficient in donor alleles in only one of the two backcrosses, which may indicate background specificity of the incompatible interactions or that these loci are not fixed between *E. grandis* and *E. globulus*. The two regions that were distorted in the F_1 hybrid and the respective backcross parent (LG6m-LG6gl and LG3p-LG3gr, Figure 3) may represent hybrid fitness loci that are segregating in one of the two species and fixed in the other. Loci that affect hybrid fitness may indeed segregate within the two species, providing that they do not have serious fitness effects within the parental species (Orr and Turelli 2001).

We found only one TRDL interaction that was indicative of a simple Dobzhansky-Muller incompatibility between a recurrent allele in the backcross parent and a donor allele in the F_1 hybrid. In the *E. grandis* backcross family, the frequency of individuals with the *E. globulus* donor allele on LG1p and one of the TRDL alleles on LG10gr was 45% lower than expected ($P = 0.0048$, Table 3). Such interactions can of course be detected only if the loci in the F_1 hybrid and the backcross parent are both heterozygous. If not, only one of the two interacting loci (most likely that in the F_1 hybrid) will be detected by TRD analysis, which may explain the presence of the recurrent-biased TRDLs on LG3p and LG11m that showed no deleterious interactions with loci in the respective backcross parents.

Finally, our evaluation of the overall fitness effect of two-locus epistatic interactions among TRDLs using the regression method of Fu and Ritland (1996) suggested that negative interactions among TRDLs contributed little overall to reduced hybrid fitness in the two backcross families (Figure 4). This finding is in contrast with that of other studies in plants that have documented extensive negative epistatic interactions in hybrid genetic backgrounds (Rieseberg *et al.* 1996; Burke *et al.* 1998; Jiang *et al.* 2000). Our inability to detect even moderate numbers of epistatic interactions in the two backcross families may be due to the limitations of our experimental design as discussed above (use of dominant markers in a backcross design and relatively small population size). The experimentwise significant interaction that we were able to detect in the *E. globulus* backcross family (LG7m and LG11m, Table 3) provided some evidence for the presence of negative epistatic interactions in this interspecific pedigree. It also provided another example of increased hybrid fitness conferred by the combination of a donor allele on one chromosome (LG7m) and a recurrent allele on another (LG11m), a common theme emerging from this study. We did not test for higher-order TRDL interactions due to the relatively low power to detect such interactions in the backcross families ($n = 186$).

Implications for natural and artificial hybridization in Eucalyptus: Viable, fertile F_1 progeny of *E. grandis* and *E. globulus* are completely absent in nature and very rare in artificial crosses. Even after prezygotic barriers between *E. grandis* × *E. globulus* crosses are overcome by artificial hybridization, <0.2% of F_1 hybrids are sufficiently viable and vigorous to produce offspring (Griffin *et al.* 2000). The F_1 hybrid individual selected here for backcrossing (tree BBT01058, Figure 1) was an exceptionally vigorous and fertile member of this elite group of F_1 individuals. The two gametes that make up its genome consequently represent a rare combination of *E. grandis* and *E. globulus* alleles that was apparently free of severe hybrid incompatibilities and allowed the expression of hybrid vigor in this individual. However, recombination of these same *E. grandis* and *E. globulus* homologs in the F_1 hybrid and segregation of the recombinant gametes in the two backcross families produced many new incompatible allele combinations as revealed by the high levels of TRD observed in the maps of the F_1 hybrid. Our results indicate that these novel incompatibilities expressed in the second hybrid generation may play an important role in restricting gene flow between well-diverged eucalypt species even when successful F_1 individuals are produced. Marker analysis of F_3 hybrid progeny in this same pedigree will reveal whether the reproductive barriers identified here continue to shape the genomes of advanced-generation hybrids of *E. grandis* and *E. globulus*.

Wu (2001) proposed a genic model of speciation in which four stages of differentiation can be distinguished. These four stages are defined by the proportion of the diverging genomes in which gene flow occurs upon secondary contact (Wu 2001, Figure 1). Although *E. grandis* and *E. globulus* clearly represent species with near-complete reproductive isolation in nature (stage IV differentiation), it is interesting to note that it was possible to generate artificial F_1 hybrids that allowed gene flow in a substantial proportion of their genomes (stage II–III differentiation). A large part of the F_1 hybrid genome was not affected by TRD and several regions were in fact strongly biased in favor of donor alleles (Figure 3). This result is not consistent with the view that well-diverged genomes act as cohesive units that resist the introgression of all donor genetic material (Wu 2001). However, the question remains whether the postzygotic barriers observed in this hybrid pedigree would be permeable enough to allow gene flow in natural stands, even in the presence of strong selection for donor genetic material. It is also not clear whether the selective advantage of these donor alleles will persist in advanced hybrid generations. The magnitude and number of reproductive barriers observed in this successful F_1 hybrid genome most likely represent lower estimates of those in "average" *E. grandis* × *E. globulus* hybrids.

Knowledge of the genetic basis of prezygotic and post-

zygotic hybridization barriers in Eucalyptus is a prerequisite for the development of effective strategies for hybrid breeding in this genus. Much has already been learned from the study of natural and artificial hybrids in Eucalyptus (Griffin *et al.* 1988), but the success of artificial hybridization and advanced-generation hybrid breeding will depend on more detailed analysis of reproductive barriers in known pedigrees. Genetic mapping studies of different F_1 hybrids of Eucalyptus species with highly informative markers such as microsatellites are required to determine common genetic components of postzygotic isolation. The genetic architecture of postzygotic barriers, and genetic diversity at the loci involved, will ultimately determine the feasibility of a directed-breeding approach to improve hybrid-combining ability in Eucalyptus. Our results suggest that genetic variation for hybrid fitness traits do exist in *E. grandis* and *E. globulus* and that parents with good hybrid combining ability may be identified if enough parents of each species are evaluated in interspecific crosses. Finally, the complexity of postzygotic barriers observed in this study suggests that very large natural and experimental populations of F_1 and F_2 hybrid progeny are required to recover individuals with desirable multilocus genotypes at hybrid fitness loci. In the future, ultrahigh-throughput marker genotyping technologies such as those recently developed on microarrays (Borevitz *et al.* 2003) will allow a much more detailed understanding of heterospecific genome interactions and may facilitate the identification of the genic components of hybrid incompatibility.

We express our gratitude to Pablo Santini and co-workers of Shell Uruguay Renewables S.A. for the maintenance of the backcross families and to Jane Harbard of Shell Forestry Technical Services, United Kingdom for the controlled pollinations that produced the backcross families. We thank David Remington and two anonymous reviewers for valuable insights and comments on this manuscript. This work was supported by funding from the North Carolina State University Forest Biotechnology Industrial Associates Consortium and by the National Institutes of Health (grant GM45344-06). A.A.M. was funded by the Fulbright Program and the National Research Foundation of South Africa. The plant materials used in this study were managed and provided by Shell Forestry, Forestal y Agricola Monte Aguila S.A., Chile, and Forestal Oriental S.A., Uruguay.

LITERATURE CITED

Bert, P. F., G. Charmet, P. Sourdille, M. D. Hayward and F. Balfourier, 1999 A high-density molecular map for ryegrass (*Lolium perenne*) using AFLP markers. Theor. Appl. Genet. **99:** 445–452.

Borevitz, J. O., D. Liang, D. Plouffe, H.-S. Chang, T. Zhu *et al.*, 2003 Large-scale identification of single-feature polymorphisms in complex genomes. Genome Res. **13:** 513–523.

Brondani, R. V. P., C. Brondani and D. Grattapaglia, 2002 Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. Mol. Genet. Genomics **267:** 338–347.

Burke, J. M., and M. L. Arnold, 2001 Genetics and the fitness of hybrids. Annu. Rev. Genet. **35:** 31–52.

Burke, J. M., T. J. Voss and M. L. Arnold, 1998 Genetic interactions and natural selection in Louisiana Iris hybrids. Evolution **52:** 1304–1310.

Burow, M. D., C. E. Simpson, J. L. Starr and A. H. Paterson, 2001 Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.): broadening the gene pool of a monophyletic polyploid species. Genetics **159:** 823–837.

Cheng, R., A. Saita, 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.

Cheng, R., A. Kleinhofs and Y. Ukai, 1998 Method for mapping a partial lethal factor locus on a molecular marker linkage map of a backcross and doubled-haploid population. Theor. Appl. Genet. **97:** 293–298.

Dobzhansky, T., 1937 *Genetics and the Origin of Species.* Columbia University Press, New York.

Eldridge, K., J. Davidson, C. Harwood and G. Van Wyk, 1993 *Eucalypt Domestication and Breeding.* Oxford University Press, Oxford.

Ellis, M. F., M. Sedgley and J. A. Gardner, 1991 Interspecific pollen pistil interaction in *Eucalyptus* L'Hér. (Myrtaceae)—the effect of taxonomic distance. Ann. Bot. **68:** 185–194.

Fishman, L., and J. H. Willis, 2001 Evidence for Dobzhansky-Muller incompatibilities contributing to the sterility of hybrids between *Mimulus guttatus* and *M. nasutus.* Evolution **55:** 1932–1942.

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

Fu, Y. B., and K. Ritland, 1994a Evidence for the partial dominance of viability genes contributing to inbreeding depression in *Mimulus guttatus.* Genetics **136:** 323–331.

Fu, Y. B., and K. Ritland, 1994b On estimating the linkage of marker genes to viability genes controlling inbreeding depression. Theor. Appl. Genet. **88:** 925–932.

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

Gore, P. L., B. M. Potts, P. W. Volker and J. Megalos, 1990 Unilateral cross-incompatibility in *Eucalyptus:* the case of hybridization between *Eucalyptus globulus* and *Eucalyptus nitens.* Aust. J. Bot. **38:** 383–394.

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

Griffin, A. R., I. P. Burgess and L. Wolf, 1988 Patterns of natural and manipulated hybridization in the genus *Eucalyptus* L'Herit—a review. Aust. J. Bot. **36:** 41–66.

Griffin, A. R., J. Harbard, C. Centurion and P. Santini, 2000 Breeding *Eucalyptus grandis* × *globulus* and other interspecific hybrids with high inviability—problem analysis and experience at Shell Forestry Projects in Uruguay and Chile, pp. 1–13 in *Hybrid Breeding and Genetics of Forest Trees: Proceedings of QFRI/CRC-SPF Symposium,* April 9–14, 2000, Noosa, Queensland, Australia, edited by H. S. Dungey, M. J. Dieters and D. G. Nikles. Department of Primary Industries, Brisbane, Australia.

Harbard, J. L., A. R. Griffin and J. Espejo, 1999 Mass controlled pollination of *Eucalyptus globulus*: a practical reality. Can. J. For. Res. **29:** 1457–1463.

Hardner, C. M., and B. M. Potts, 1997 Postdispersal selection following mixed mating in *Eucalyptus regnans.* Evolution **51:** 103–111.

Harushima, Y., M. Nakagahra, M. Yano, T. Sasaki and N. Kurata, 2001 A genome-wide survey of reproductive barriers in an intraspecific hybrid. Genetics **159:** 883–892.

Hedrick, P. W., and O. Muona, 1990 Linkage of viability genes to marker loci in selfing organisms. Heredity **64:** 67–72.

Jiang, C. X., P. W. Chee, X. Draye, P. L. Morrell, C. W. Smith *et al.*, 2000 Multilocus interactions restrict gene introgression in interspecific populations of polyploid *Gossypium* (cotton). Evolution **54:** 798–814.

Karkkainen, K., H. Kuittinen, R. Van Treuren, C. Vogl, S. Oikarinen *et al.*, 1999 Genetic basis of inbreeding depression in *Arabis petraea.* Evolution **53:** 1354–1365.

Lespinasse, D., M. Rodier-Goud, L. Grivet, A. Leconte, H. Legnate *et al.*, 2000 A saturated genetic linkage map of rubber tree

(*Hevea* spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. Theor. Appl. Genet. **100:** 127–138.

Lopez, G. A., B. M. Potts and P. A. Tilyard, 2000 F_1 hybrid inviability in *Eucalyptus*: the case of *E. ovata* x *E. globulus*. Heredity **85:** 242–250.

Lynch, M., and A. G. Force, 2000 The origin of interspecific genomic incompatibility via gene duplication. Am. Nat. **156:** 590–605.

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.

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

Muller, H. J., 1942 Isolating mechanisms, evolution and temperature. Biol. Symp. **6:** 71–125.

Myburg, A. A., D. L. Remington, D. M. O'Malley, R. R. Sederoff and R. W. Whetten, 2001 High-throughput AFLP analysis using infrared dye-labeled primers and an automated DNA sequencer. BioTechniques **30:** 348–357.

Myburg, A. A., A. R. Griffin, R. R. Sederoff and R. W. Whetten, 2003 Comparative genetic linkage maps of *E. grandis*, *E. globulus* and their F_1 hybrid based on a double pseudo-backcross mapping approach. Theor. Appl. Genet. **107:** 1028–1042.

Orr, H. A., and M. Turelli, 2001 The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution **55:** 1085–1094.

Peng, J., A. B. Korol, T. Fahima, M. S. Roder, Y. I. Ronin *et al.*, 2000 Molecular genetic maps in wild emmer wheat, *Triticum dicoccoides*: genome-wide coverage, massive negative interference, and putative quasi-linkage. Genome Res. **10:** 1509–1531.

Quillet, M. C., N. Madjidian, Y. Griveau, H. Serieys, M. Tersac *et al.*, 1995 Mapping genetic factors controlling pollen viability in an interspecific cross in *Helianthus* sect. *Helianthus*. Theor. Appl. Genet. **91:** 1195–1202.

Ramsay, L., M. Macaulay, S. D. Ivanissevich, K. Maclean, L. Cardle *et al.*, 2000 A simple sequence repeat-based linkage map of barley. Genetics **156:** 1997–2005.

Remington, D. L., and D. M. O'Malley, 2000 Whole-genome characterization of embryonic stage inbreeding depression in a selfed loblolly pine family. Genetics **155:** 337–348.

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.

Rieseberg, L. H., 2001 Chromosomal rearrangements and speciation. Trends Ecol. Evol. **16:** 351–358.

Rieseberg, L. H., and J. M. Burke, 2001 A genic view of species integration—commentary. J. Evol. Biol. **14:** 883–886.

Rieseberg, L. H., C. R. Linder and G. J. Seiler, 1995 Chromosomal and genic barriers to introgression in Helianthus. Genetics **141:** 1163–1171.

Rieseberg, L. H., B. Sinervo, C. R. Linder, M. C. Ungerer and D. M. Arias, 1996 Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. Science **272:** 741–745.

Rieseberg, L. H., S. J. E. Baird and K. A. Gardner, 2000 Hybridization, introgression, and linkage evolution. Plant Mol. Biol. **42:** 205–224.

Ting, C. T., S. C. Tsaur, M. L. Wu and C. I. Wu, 1998 A rapidly evolving homeobox at the site of a hybrid sterility gene. Science **282:** 1501–1504.

Turelli, M., and H. A. Orr, 2000 Dominance, epistasis and the genetics of postzygotic isolation. Genetics **154:** 1663–1679.

Van Heusden, A. W., J. W. Van Ooijen, R. Vrielink-Van Ginkel, W. H. J. Verbeek, W. A. Wietsma *et al.*, 2000 A genetic map of an interspecific cross in *Allium* based on amplified fragment length polymorphism (AFLP $^{(TM)}$) markers. Theor. Appl. Genet. **100:** 118–126.

Verhaegen, D., and C. Plomion, 1996 Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. Genome **39:** 1051–1061.

Vogl, C., and S. Xu, 2000 Multipoint mapping of viability and segregation distorting loci using molecular markers. Genetics **155:** 1439–1447.

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.

Wang, G. L., J. M. Dong and A. H. Paterson, 1995 The distribution of *Gossypium hirsutum* chromatin in *Gossypium barbadense* germplasm—molecular analysis of introgressive plant breeding. Theor. Appl. Genet. **91:** 1153–1161.

Whitkus, R., 1998 Genetics of adaptive radiation in Hawaiian and Cook Islands species of *Tetramolopium* (Asteraceae). II. Genetic linkage map and its implications for interspecific breeding barriers. Genetics **150:** 1209–1216.

Wu, C.-I, 2001 The genic view of the process of speciation. J. Evol. Biol. **14:** 851–865.

Communicating editor: O. Savolainen