GENETIC MAPPING IN LOBLOLLY PINE

M. M. Sewell and D. B. Nealel

Abstract.--A consensus map for loblolly pine (Pinus taeda L.) was constructed using restriction fragment length polymorphisms (RFLPs) as genetic markers. This map is based on segregation data from two unrelated three-generation pedigrees (Devey et al., 1994; Groover et al., 1994) and was assembled using the linkage computer program JoinMap (Stam, 1993). The merger of individual genetic maps into a consensus map for loblolly pine allows for the integration of genetic information from independent sources onto a single map and facilitates the consolidation of linkage groups to represent the 12 chromosomes of loblolly pine. This consensus map contains many known and characterized genes, and serves as the foundation for present and future genetic studies in loblolly pine and for studies of genome organization and evolution in conifers.

Keywords: Pinus taeda L., genetic mapping, RFLP molecular markers, genome organization.

INTRODUCTION

An increasing number of genetic linkage maps are being constructed for forest tree species. These maps are commonly used for mapping Mendelian traits such as disease resistance and quantitative traits of agronomic importance. Therefore, multiple maps for a single species are often constructed from individual pedigrees that segregate for a specific trait of interest. Synthesis of these individual maps into a single consensus map to represent each species will be a valuable resource for breeders and evolutionary biologists alike. Breeders can ascertain the map positions of genes controlling the synthesis of important traits as well as the relative position of those genes to one another. Evolutionary biologist can determine the copy number of a gene found within the genome (and the relative map position of genes within these gene families), the frequency of these multigene families within the genome, etc. As more genetic information accumulates for a species and is incorporated into a consensus map, hypotheses regarding genome organization and evolution in conifers can be further explored.

- Institute of Forest Genetics, USDA Forest Service, Pacific Southwest Research Station, Placerville, CA 95667

For loblolly pine (Pinus taeda L.), an independent linkage map, based primarily on restriction fragment length polymorphism (RFLP) markers, was previously constructed from each of two outbred three-generation pedigrees (Devey et al., 1994; Groover et al., 1994). Each map contains unique genetic information; Devey et al. (1994) mapped complementary DNA (cDNA) and isozyme loci, while Groover et al. (1994) mapped additional cDNAs and quantitative trait loci for wood specific gravity.

The objective **of** this study is to integrate these two loblolly pine maps, derived from genetically independent pedigrees, into a single consensus map. This objective will serve two purposes:

1) to synthesize the available genetic information into a single consensus map, which will serve as a foundation for genetic study in loblolly pine and for studies of genome organization and evolution in conifers, and

2) to consolidate linkage groups to represent each of the 12 chromosomes of loblolly pine.

Our strategy is to first saturate each independent linkage map with as many markers as are readily available, thereby consolidating linkage groups *within* each pedigree. Secondly, we will map loci that are common to each pedigree. These "common" loci then serve as anchor points to integrate the linkage data from each pedigree and will further consolidate linkage groups *among* pedigrees.

MATERIALS AND METHODS

Mapping Populations

Genetic linkage maps for each of two independent three-generation outbred pedigrees were recently completed. The first map (here-to-fore called the *base* map) was constructed from F2 segregation data for 90 RFLP and six isozyme markers from 95 progeny (Devey et al., 1994). This pedigree was constructed and maintained by the North Carolina State University Cooperative Tree Improvement Program (NCSU Coop). The second map (called the *gt1* map) was constructed from F2 segregation data for 142 RFLP markers from 175 progeny (Groover et al., 1994). This pedigree was constructed and maintained by the NCSU Coop and Weyerhaeuser Company and was selected based on extreme-high and -low values for wood specific gravity among grandparental pairs.

Source of Probes

We chose three sources of genetic markers for mapping: RFLPs (Devey et al., 1991), isozymes (Conkle, 1981) and

random amplified polymorphic DNA (RAPDs) (Kiehne and Neale, 1995). We use RFLP and isozyme markers because they are 1) codominant and multiallelic, 2) highly repeatable and syntenic across genetic backgrounds (Conkle, 1981; Ahuja et al., 1994) and 3) RFLPs are the most efficient way to map cDNAs. Kinlaw et al. (1995) has initiated "single-pass" sequencing of these cDNAs which are then compared to known genes in nucleotide sequence databases. Consequently, many of these markers are of known genes and they can also be used as orthologous markers among different species of pines. We use RAPDs because they are an efficient source of a high number of markers that are putatively found at random from throughout the genome.

Strategy for Map Integration

A. Integration of linkage data within a pedigree.

Since recombination occurs independently during the production of maternal and paternal gametes, the genetic segregation observed in the progeny represents both sources of recombination. Therefore, we arrange the segregation data into independent maternal and paternal datasets.

In an outbred pedigree, four informative mating types are possible for any given locus (see below). Each mating type reflects which parent is heterozygous and therefore produces alternate alleles that segregate among the progeny.

Mating Type	Cross
Maternally Informative (MI)	HxA
Paternally Informative (PI)	AxH
Fully Informative (FI)	HixHj
Both-Informative (BI)	HixHi

where, A = homozygote
H = heterozygote
(Hi, Hj are different heterozygotes)

For MI and PI mating types, only one parent is heterozygous. However, for FI mating types, each parent is heterozygous for a different pair of alleles. Therefore the segregation data for a locus of an FI mating type can be recoded once as MI and again as PI (i.e., HixHj is recoded as HixA and AxHj). Parents of BI mating types are also heterozygous and can be treated in a similar manner. However, since each parent is heterozygous for the same pair of alleles, some segregation data is ambiguous (i.e., it is difficult to determine which parent contributed the alleles of the heterozygous progeny class).

The loci from FI and BI mating types, after being recoded and placed into the appropriate maternal or paternal

dataset, serve as common loci among these datasets. Utilizing these FI and BI loci as anchor-points, these independent datasets are integrated using the linkage computer program JoinMap (Starr, 1993), which uses a modified least squares procedure for estimating map distances from independent and weighted joint estimates of pairwise recombination frequencies.

B. Integration of linkage data among pedigrees.

The integration of linkage data among independent pedigrees is performed in a manner similar to the integration of maternal and paternal linkage data within a pedigree. Instead of uniting maternal and paternal linkage data via loci from FI and BI mating types, linkage data from each pedigree is integrated via loci that are "common" to each pedigree dataset.

To determine commonality of loci among pedigree datasets, comparison of molecular markers are first made by inspection of migration distances of RFLP bands on autoradiograms. Secondly, map distances between putatively common loci pairs are compared for collinearity. Loci that meet both of these criteria can then be used to integrate linkage data from independent pedigrees.

RESULTS AND DISCUSSION

Map Integration

Figure 1 illustrates the integration of independent linkage data from the *base* and *qtl* pedigrees into a single *consensus* linkage group. In this example, two BI and three FI loci (represented by dashed arrows) serve as anchor-points to unite the maternal and paternal linkage groups within the *qtl* map; one BI and four FI loci unite the maternal and paternal linkage groups within the *base* map. Six loci serve as "common" loci that unite linkage groups among pedigrees (represented by solid arrows). This example demonstrates that more than one linkage group from any given dataset can be united by this integration process (e.g., two maternal linkage groups from the *qtl* map are brought together via integration with the *base* map). Note that at least two anchor-points between linkage groups are necessary for unambiguous orientation of individual linkage groups.

In an attempt to saturate the linkage groups within each pedigree, we have added approximately 150 new genetic markers to the existing *base* and *qtl* maps. Forty-one markers from the *base* map and 49 markers from the *qtl* map serve as anchor-points to unite the maternal and paternal linkage groups within each pedigree and 42 "common" markers serve as anchor-points to unite linkage groups among each pedigree. Although

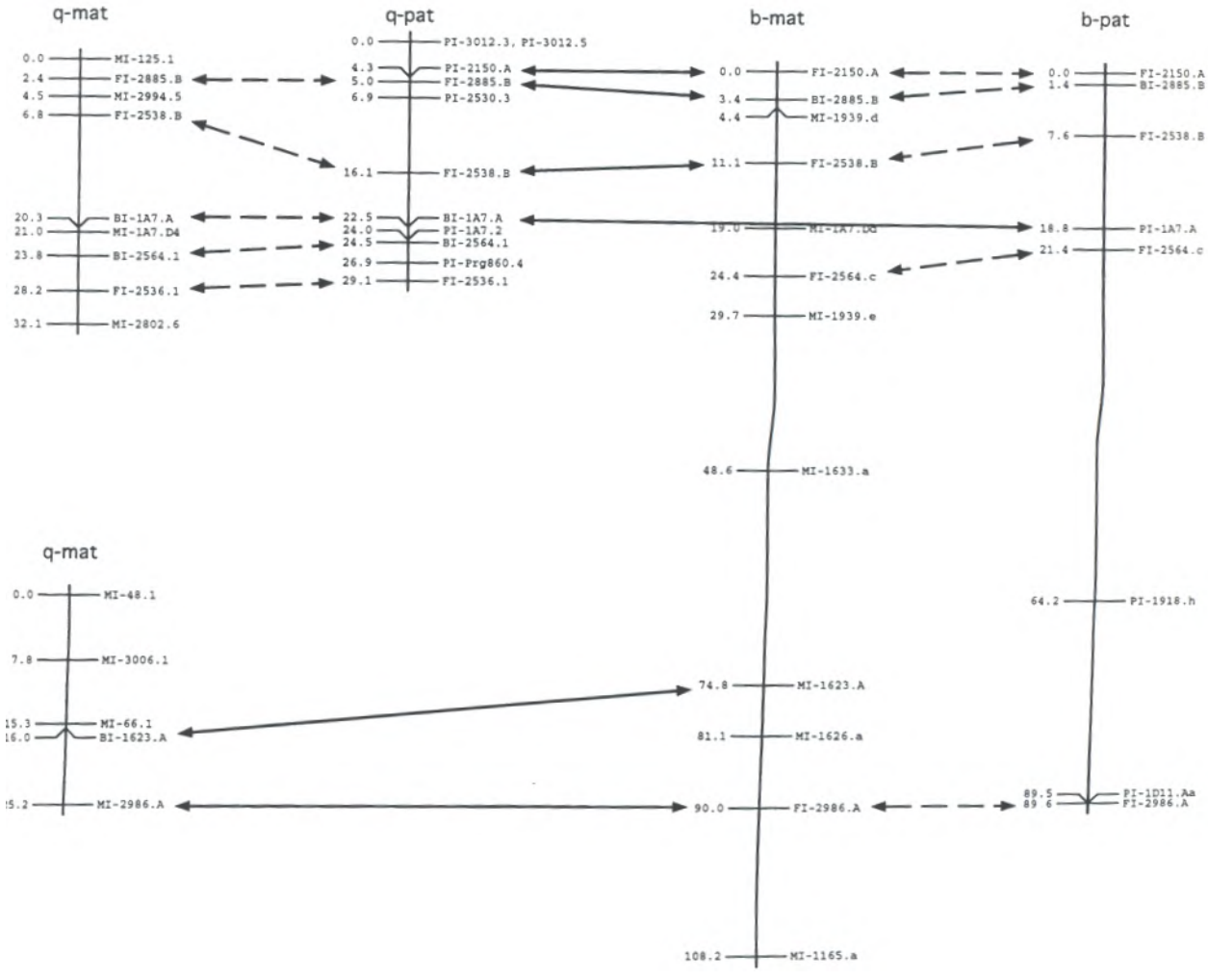

Figure 1. Integration of linkage data from within and among pedigrees. Using *consensus* linkage group 2 as an example, maternal (mat) and paternal (pat) linkage data from the *base* (b) and *qtl* (q) pedigrees are integrated into a single linkage group. Within a pedigree, FI and BI loci serve as anchor-points (dashed arrows); among pedigrees, "common" loci serve as anchor-points (solid arrows). See text for more details.

the number of markers and the degree of integration of linkage groups on the *consensus* map is reasonably high, we have yet to achieve our goal of 12 discrete linkage groups that represent the 12 chromosomes of loblolly pine (Table 1).

Table 1. Saturation of genetic markers on the *base, qtl* and *consensus* maps.

	base	*qtl*	*consensus*
Total unique loci	126	236	320
MI loci	45	96	
PI loci	41	91	
FI loci	33	36	
BI loci	8	13	
Common loci			42
No. linkage groups	9	17	18
Total distance (cM)	575	908	936
Average distance (cM)	6.7	4.4	3.5

Mapped Genes of Known Function

To identify functions for the molecular markers used in our genetic mapping project, Kinlaw et al. (1995) has determined partial DNA sequences for more than 200 loblolly pine cDNAs. Approximately 44% of these sequences have matched known genes based on database searches. Most sequence similarities are to genes from other plant species and include many enzymes involved in cellular metabolism and photosynthesis. This data is being used to study conifer genome organization and evolution. Sequencing efforts are being expanded to a variety of clones from specific conifer tissues. In addition, various collaborators have contributed cDNAs of characterized genes from loblolly pine (J. Cairney; D. Harry; C. Loopstra; D. O'Malley) and Scotts pine (P. sylvestris) (Jansson and Gustafsson; Karpinski et al.). Data for 18 isozyme loci is also included. Approximately 200 of these genes of known function are placed on the *consensus* map.

Future Directions

A logical extension from integrating independent maps from loblolly pine is comparative mapping among species from the genus Pinus. In collaboration with other laboratories, we have begun comparisons among loblolly, Monterey and slash pine. We will extend this effort to representative species from subsections of Pinus.

Comparative genetic mapping has become a powerful technique for investigating the mode and tempo of chromosomal evolution (Whitkus et al., 1992). However, for our goals of

mapping important traits in conifers, a more important aspect of comparative mapping is the potential to establish collinearity of linkage groups among pine genomes. If pine genomes are collinear, then it is likely that the location of a gene in one species can be used to predict the location of the homologous gene in another species (Jena et al., 1994).

The practical application of collinearity is that map locations for genes of important traits in one species will lead to a directed search for homologous genes in another species. For example, our success in mapping quantitative trait loci (QTLs) for wood specific gravity in loblolly pine (Groover et al., 1994) may help uncover QTLs for wood specific gravity in other pine species. Consequently, the emphasis of forest genetics research is not on a single species, but instead can focus on many species, each with regional importance and concerns. For this reason, establishing the extent of collinearity among pine genomes is fundamental to understanding genome organization and function in pines.

ACKNOWLEDGMENTS

Many people have contributed to this mapping project. Leon Burris gladly fulfilled our germplasm requests; Paul Hodgskiss contributed to the isozyme analysis; Phillip Wilcox and Ron Sederoff provided orientation to the RAPD analysis; Brad Sherman, Jennifer Lee and Stacey Harrington provided technical support. We are grateful for financial support from the USDA Competitive Grants Program, the USDA/ARS Plant Genome Project, Weyerhaeuser Company and the USDA Forest Service.

LITERATURE CITED

Ahuja, M. R., M. E. Devey, A. T. Groover, K. D. Jermstad, and D. B. Neale. 1994. Mapped DNA probes from loblolly pine can be used for restriction fragment length polymorphism mapping in other conifers. Theor Appl Genet 88:279-282.

Conkle, M. T. 1981. Isozyme variation and linkage in six conifer species. In: Proc. Symp. Isozymes of N. Amer. Forest Trees and Forest Insects. USDA For. Ser. Gen. Tech. Rep., PSW-48, pp 11-17.

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

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

Groover, A. T., M. E. Devey, T. A. Fiddler, J. M. Lee, R. A. Megraw, T. Mitchell-Olds, B. K. Sherman, S. L. Vujcic, C. G. Williams, and D. B. Neale. 1994. Identification of quantitative trait loci influencing wood specific

gravity in an outbred pedigree of loblolly pine. Genetics 138:1293-1300.
Jena, K. K., G. S. Khush, and G. Kochert. 1994. Comparative RFLP mapping of a wild rice, Oryza officinalis, and cultivated rice, 0. sativa. Genome 37:382-389.
Kiehne and Neale. 1995. Fine structure mapping of quantitative trait loci influencing wood specific gravity in loblolly pine. In preparation.
Kinlaw, C. S., C. Baysdorfer, E. E. Gladstone, D. E. Harry, M. M. Sewell, and D. B. Neale. 1995. Loblolly pine cDNA sequence analysis. Poster abstract P124, Plant Genome III, p. 53.
Starr, P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. The Plant Journal 3:739-744.
Whitkus, R., J. Doelbly, and M. Lee. 1992. Comparative genome mapping of sorghum and maize. Genetics 132:1119-1130.