GENOMIC MAPPING OF *PINUS SYLVESTRIS* (L.) USING RANDOM AMPLIFIED POLYMORPHIC DNA MARKERS

Reza Yazdani[1], Francis C. Yeh[2] & Janis Rimsha[3]

[1)] Department of Forest Genetics, Swedish University of Agricultural Sciences, Box 7027, S–750 07, Uppsala, Sweden
[2)] Department of Renewable Resources, University of Alberta, Edmonton, Canada, T6G 2H1
[3)] Latvian Forestry Research Institute, 111 Riga Str, Salaspils, LAT–229021 Latvia

Received March 6, 1995; accepted November 8, 1995

ABSTRACT

This paper reports the construction of a high-density Random Amplified Polymorphic DNA (RAPD) map in *Pinus sylvestris* (L.). Genomic DNA of haploid megagametophytes from seeds originating from a single tree were amplified with 220 random decamer oligonucleotide primers by the polymerase chain reaction. Two hundred and ninety-eight RAPDs with fragment sizes that ranged between 250 and 2500 base pairs were found segregating at 104 random decamer oligonucleotide primers. Of these, 16 (5%) were excluded from mapping because they did not conform to 1:1 Mendelian segregation. Two hundred and sixty-one of the remaining 282 RAPDs formed 14 linkage groups for a total distance of 2,638.6 cM. Four of the 21 RAPDs we did not map could not be placed with the main linkage groups. The remaining 17 RAPDs could be mapped, but they were loosely linked at recombinant frequency near 50% to warrant serious consideration. The availability of a genetic linkage map in *Pinus sylvestris* with a large number of RAPD markers should facilitate the identification of quantitative trait loci, contribute towards marker-assisted selection and allow the choice of unlinked RAPD markers for population genetic studies.

Key words: *Pinus sylvestris*, RAPDs, linkage

INTRODUCTION

Molecular genetic linkage maps of conifers have been constructed only recently (NEALE & WILLIAMS 1991; TULSIERAM *et al.* 1992; NELSON *et al.* 1993; AHUJA *et al.* 1994; BINELLI & BUCCI 1994). Some of the earlier genetic maps in conifers related to linkage relationships among a small number of isozyme loci (*e.g.*, GURIES *et al.* 1978; ADAMS & JOLY 1980; EKERT *et al.* 1981; NEALE & ADAMS 1981; EL-KASSABY *et al.* 1982; KING & DANCIK 1983; CHELIAK & PITEL 1985; MUONA *et al.* 1987; NA'IEM *et al.* 1993). They showed that linkage blocks in conifers were highly conserved and large chromosomal rearrangements likely did not occur during evolution (CONKLE 1981). Recently, AHUJA *et al.* (1994) showed that nearly all *Pinus taeda* probes cross-hybridized to DNA from species within the genus *Pinus*. Consequently, the genetic linkage map of a given conifer probably could be extrapolated to that of related species. NEALE and WILLIAMS (1991) estimated that the genomes of *Pinus* species are around 2,500 cM each. A Random Amplified Polymorphic DNA (RAPD) map of approximately 2,160 cM accounted for 64-75% of the total *Pinus elliottii* genome (NELSON *et al.* 1993) and in *Pinus taeda,* 191 RAPDs were mapped to 12 linkage groups with a distance of 1,687 cM (O'MALLEY *et al.* unpublished data). BINELLI and BUCCI (1994) placed 185 RAPDs to 17 major linkage groups covering 3,584 cM in *Picea abies.* Hence, the size of conifer genome, in terms of map units, is not expected to be much greater than that of the crop species (MCCOUCH *et al.* 1988; TANKSLEY *et al.* 1992; VELDBOOM *et al.* 1994).

The number of genes governing economic traits, their spatial organizations and the degree they vary among different trees, populations and species remain largely unknown in conifers. Genomic maps will enable us to identify and locate these genes, thus, assist tree breeders select seedlings with combinations of desirable genes. This type of plant selection, known as marker assisted selection (MAS), has the potential to become one of the most useful tree breeding tools ever developed. Genomic maps will also allow the choice of unlinked markers, important to unbiased estimate of genetic parameters for natural populations.

We report the construction of molecular genetic linkage map of *Pinus sylvestris* (L.) using RAPD markers that rely on the amplification of specific DNA fragments from total genomic DNA. The advantages of RAPD markers to mapping over the Restriction Fragment Length Polymorphism DNA markers are their rapidity, simplicity and the need for very small amounts of genomic DNA (INNIS *et al.* 1990). The availability of a large number of RAPD markers should facilitate the identification of quantitative trait loci, contribute towards marker-assisted selection and allow the choice

of unlinked RAPD markers for population genetic studies.

MATERIALS AND METHODS

Plant Materials

Open-pollinated seeds from clone 49-2, which originated from the Eiche's population trial in Sweden, were collected from the clonal archive at Bogesund near Stockholm during 1993. Screening of RAPD markers was carried out on genomic DNA extracted from each of 70 megagametophytes.

Laboratory Protocols

Genomic DNA were isolated from individual megagametophyte of germinating seed by a CTAB procedure modified from DOYLE and DOYLE (1990). Proteinase K (1 $mg \cdot ml^{-1}$) was added before incubation at 60 °C. This method yielded between 1,000–2,000 ng of DNA per megagametophyte. Our tests indicated that the optimal amount of DNA template used for RAPD PCR reaction was between 1–2 ng. RAPD PCR reactions were assembled in 96-well microtiter plates using random decamer oligonucleotide primers purchased from Operon Technologies (Alameda CA).

RAPD PCR reaction volume was 13 µl in 96-well plates. Each well contained 1.3 µl Promega PCR buffer, 1.06 µl of a 10 mM dNTP for each nucleotide, 2 ng template DNA in 3 µl of double-distilled water, 15 ng of a 10-mer from the collections of random decamer oligonucleotide primers suspended in 1.5 µl of double-distilled water, 0.7 unit of Taq DNA polymerase and water to a total volume of 13 µl. We used Tag DNA polymerase purchased from Promega in 209 of the 222 RAPD PCR reactions. The remaining 13 RAPD PCR reactions, we used Tag DNA polymerase purchased from Dynozyme. RAPD PCR reactions were carried out using a MJ Research 96-well thermocycler for 41 cycles consisting of denaturation for 1 minute at 92 °C, primer annealing for 1 minute at 37 °C and primer extension for 2 minutes at 72 °C. RAPD products were subjected to electrophoresis on 1.5% (w / v) agarose gels in tris-borate EDTA buffer at 100 V for 4 h. Gels were stained in 0.5 $\mu g \cdot \mu l^{-1}$ ethidium bromide solution and photographed under illumination with UV light.

Primer Selection and RAPD Analysis

Two hundred and twenty-two random decamer oligonucleotide primers that consistently revealed scorable and reproducible RAPD fragments over several independent runs were chosen for this mapping study. Random primers could be used individually and in pairwise combination to generate new genomic fingerprints (WELSH & MCCLELLAND 1991). We studied five pairwise combination of eight random decamer oligonucleotide primers and found that less than half of the amplified fragments were the same as those produced by one of the two primers alone, while the other half of the amplified fragments were new variants. Thus, a few random decamer oligonucleotide primers could potentially be used in a large number of single and pairwise combinations, each producing distinct PCR fingerprints.

Most RAPD markers are dominant but in haploid megagametophytes there could be either presence "+", or absence " –" of a band (fragment). Molecular weights for the RAPD fragments were determined by comparison with the molecular weight of a reference marker produced by a standard 1 kb-ladder from BRL/Gibco. We have used standard nomenclature to assign different RAPD fragment with primers identity and fragment size that ranged between 250–2,500 base pairs.

Statistical Methods

Chi-square analysis of segregating RAPD markers at the 0.05 significance level tested the goodness of fit to a 1:1 Mendelian ratio. RAPD markers that did not conform to 1:1 Mendelian segregation were excluded from the linkage analysis. The assignment of RAPD markers to linkage groups used MAPMAKER Macintosh V2.0 (LANDER *et al.* 1987), an interactive computer package for constructing primary genetic linkage maps. Raw data was prepared as an F2 backcross data file for the MAPMAKER program. We first performed a two-point analysis of our entire data set with a LOD score of at least 3.0 and a recombination fraction of at most 0.40 to assign RAPD markers to possible linkage groups. Then, we used three-point analysis to determine the likely orders of these linkage groups. Finally, we used multipoint analysis to resolve any discrepancies not resolved by the three-point analysis and to display the maximum likelihood map distances for the indicated map orders.

RESULTS AND DISCUSSION

Two hundred and ninety-eight RAPDs with fragment sizes that ranged between 250 and 2,500 base pairs were found segregating at 104 of 222 random decamer oligonucleotide primers. Figs.1 and 2 are representative samples of the stained gels from screening and mapping of RAPDs with different random decamer oligonucleotide primers on genomic DNA extracted from megagametophytes. Fig.1 shows two fragments, 1 and 2, for primer A01. In Fig. 2 there are six fragments (1, 2, 3, 4, 5 and 6) for primer G13.

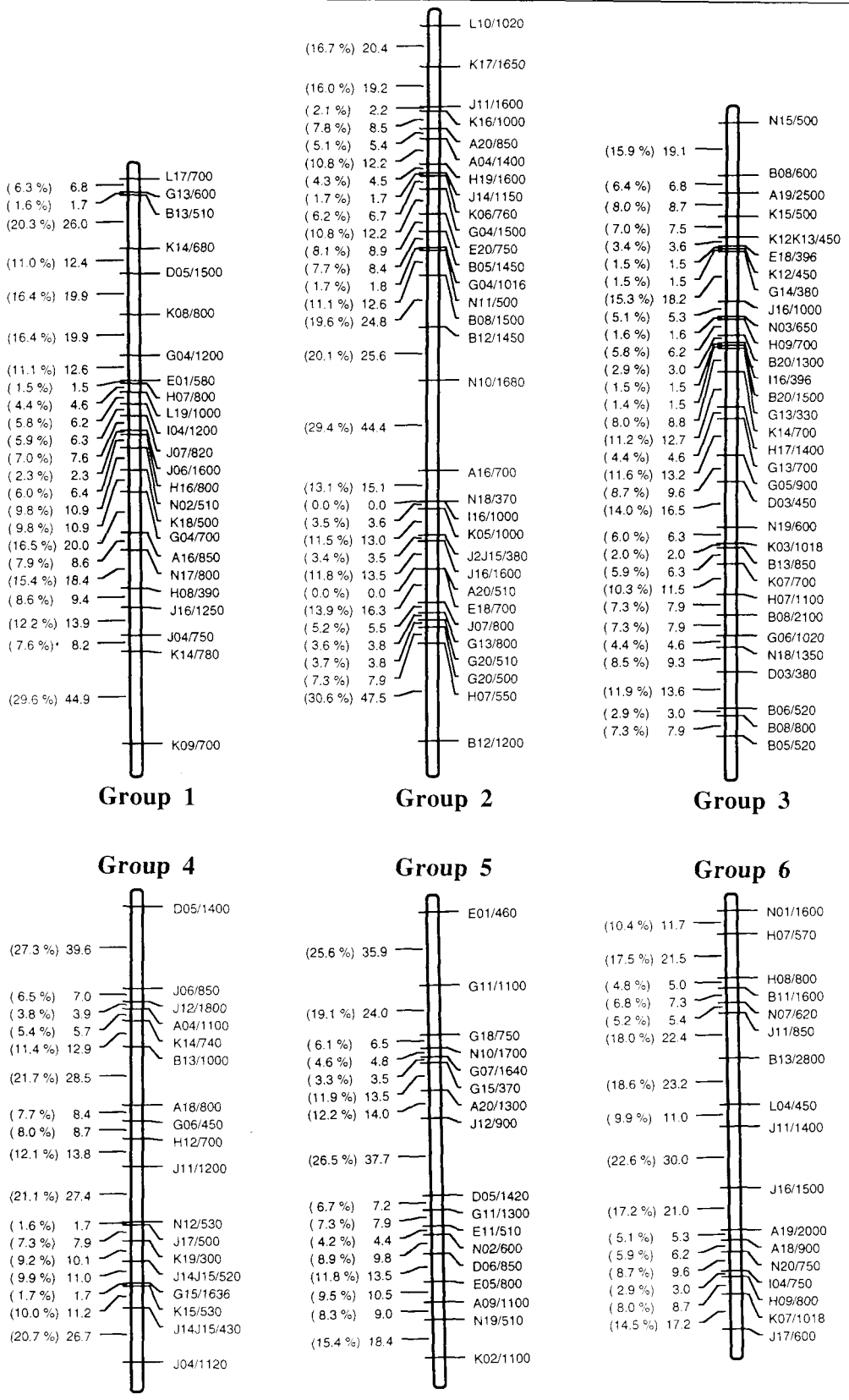

Figure 3 Linkage map for *Pinus sylvestris* based on 261 RAPD markers. Marker names containing the primer ID and molecular weights for RAPD fragments are given to the right of the linkage groups and recombination frequency (in parenthesis) and Haldane centimorgan distances (HALDANE 1919) are given to the left. Marker names containing a mixture of two primers are

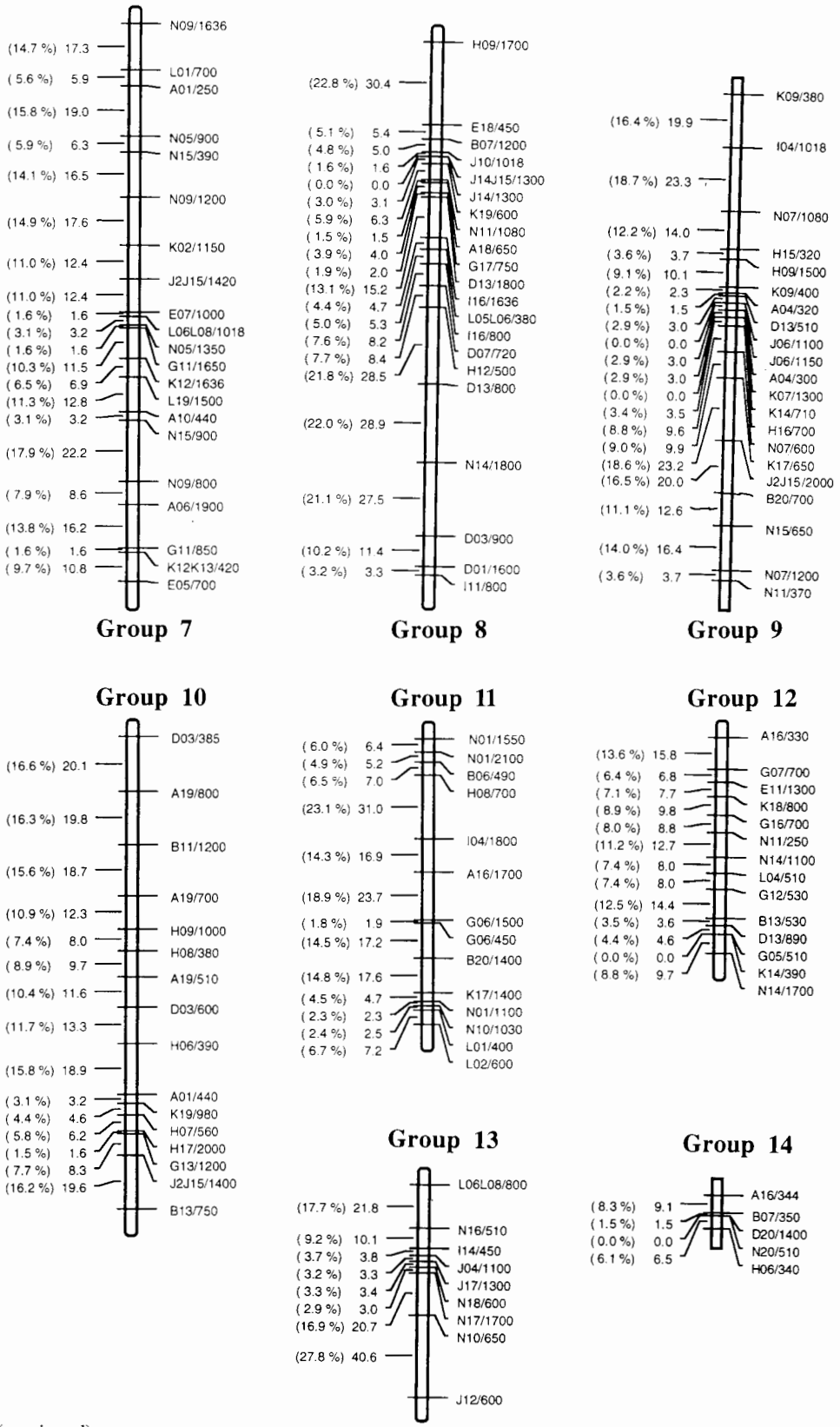

Figure 3 (continued)
designated with two primer ID on the right-hand side of the linkage groups. The map for linkage group 2 is scaled at 20 cM per cm while the remaining 13 linkage groups were mapped at 15 cM per cm.

cal features like the centromeres, telomeres and heterochromatin (LINDAHL 1991). The later explanation is suggested by data from yeast and mice where certain sequences enhanced recombination by a factor of 5 or more (COLEMAN *et al.* 1986; LINDAHL 1991). There are some reports concerning recombination differences as a result of influence by different genetic background or age (ANDERSSON & SANDBERG 1983).

The number of map units per linkage group may correspond to the cytological length of chromosomes in *Pinus sylvestris.* In tomato and potato genomes, TANKSLEY *et al.* (1992) showed that map units per chromosome tightly correlated with the number of markers per chromosome and pachytene length (r = 0.84 and 0.89, respectively).

Gaps have been found in some regions of our linkage map. Some of the longest gaps correspond to 47.5 cM distance or a 30.6 % recombination and 44.4 cM or 29.4 % recombination on linkage group 2, and 44.9 cM or 29.6 % recombination on linkage group 1. There is a total of 18 gaps that exceed 25 cM distance or 20 % recombination frequency (Figure 3). Mapping additional RAPD markers may eventually fill the larger gaps in the map for *Pinus sylvestris.* It should also be stressed that estimates of map distances are very sensitive to sampling errors. Thus, the 70 megagametophytes analysed in this study probably did not represent a reliable sample. The small size of the sampled megagametophytes might be one reason why we could not established with absolute certainly the orders of the RAPD markers at the ends for some of the linkage groups.

One important application of molecular maps is to improve our understanding of genomic organization (WELSH & MCCLELLAND 1990). For example, it is of interest to compare the consistency of the RAPD linkage map based on recombination frequencies, as is done in this study, with the RAPD maps of other individuals within a population and species. This kind of comparison will give us information about how the genome of *Pinus sylvestris* is organised. It is also informative to identify the presence of chromosomal rearrangement such as inversions, duplications, translocations and so on by comparing maps of different tree species (LUNDIN 1993). Phylogenetically related forest tree species may have conserved gene orders. When many putative RAPD loci are assigned to each of several linkage groups, as in this study, the genetic linkage map can infer about chromosome rearrangement, the order of loci, and the recombination frequency between adjacent loci in different species. If the loci are linked in more than one species this suggests the presence of a conserved chromosomal segment. First, patterns of conserved regions can address the evolutionary history of chromosomes and thus, shed some light on the genomic organization of ancestral species (LUNDIN 1993). Second, conserved segments are extremely useful when seeking for candidate genes based on gene arrangements in other species. Whenever a trait is localised to a particular chromosomal region, the map can be compared with those of the other organisms (JOHANSSON *et al.* 1992).

Quantitative traits in forest trees are assumed to be under the control of many genes each with small effect. Recent advances in molecular mapping allow us to test if such assumptions are correct. In F_2 progeny originating from crosses between *Populus trichocarpa* and *Populus deltoides,* BRADSHAW *et al.* (1994) used segregants for a variety of phenotypic traits to construct a linkage map and for mapping quantitative trait loci (QTL). The *Populus* genome was estimated to have a genetic length of 2600 cM and 4 QTL controlled 86% of the genetic variance and 49% of phenotypic variance in two-year field stem volume growth. In *Eucalyptus grandis,* a genomic region that simultaneously increased volume growth and wood specific gravity was identified, suggesting pleiotropic gene action (GRATTAPAGLIA 1994). A major dominant gene for resistance to *Cronartium ribicola* in *Pinus lambertiana* was rapidly mapped (DEVEY 1994) using RAPDs and bulk segregant analysis advanced by MICHELMORE *et al.* (1991).

The map developed by us has RAPD markers, on average, every 10.1 cM. This density of RAPDs makes it very likely that any target gene will be located within a few map units of at least one of the RAPD marker. Obvious quantitative traits to be analyzed for future mapping study of *Pinus sylvestris* will include wood density, fibre quality, and adaptive traits like bud flushing, bud set, bud dormancy, frost hardiness and growth capacity. Once a QTL has been detected, the linkage information can be applied in breeding using marker assisted selection (TUSKAN 1991; TAUER *et al.* 1992; WILLIAMS & NEALE 1992). Furthermore, following detection of a QTL, the next phase of gene identification and isolating of an unknown gene product via genetic linkage map will start (ORKIN 1986).

ACKNOWLEDGEMENT

We thank Ron Sederoff, David O'Malley and Ben Liu for suggestions on data analysis, and three anonymous reviewers for many valuable comments. This study was supported in part by grants from the Swedish Council for Forestry and Agricultural Research to R. Y. and fromAssociation for Forest Tree Breeding and the Swedish University of Agricultural Sciences and the University of Alberta to F. C. Y.

REFERENCES

ADAMS, W. T. & JOLY, R. J. 1980: Linkage relationships among twelve allozyme loci in loblolly pine. *J. Hered.*

71:199–202.

AHUJA, M. R., DEVEY, M. E., GROOVER, A. T., JERMSTAD, K. D. & NEALE, D. B. 1994: Mapped DNA probes from loblolly pine can be used for restriction fragment length polymorphism mapping in other conifers. *Can. J. For. Sci.* **88**:279–282.

ANDERSSON, L. & SANDBERG, K. 1983: Genetic linkage in the horse. II. Distribution of male recombination estimates and the influence of age, breed and sex on recombination. *Genetics* **106**:109–122.

BINELLI, G. & BUCCI, G. 1994: A genetic linkage map of *Picea abies* Karst. based on RAPD markers as a tool in population genetics. *Theor. Appl. Genet.* **88**:283–288.

BRADSHAW, H. D., VILLAR, M., WATSON, B. D., OTTO, K. G., STEWART, S. & STETTLER, R. F. 1994: Molecular genetics of growth and development in *Populus*. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. *Theor. Appl. Genet.* **89**:167–178.

BURNHAM, C. R. 1962: Discussion in cytogenetics. GR Burnham Press, St. Paul.

CHANG, C., BOWMAN, J. L., DEJOHN, A. W., LANDER, E. S. & MEYEROWITZ, E. M. 1988: Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **85**:6856–6860.

CHELIAK, W. M. & PITEL, J. A. 1985: Inheritance and linkage of allozymes in *Larix laricinia*. *Silvae Genet.* **34**:142–148

COLEMAN, K. G., STEENSMA, H.V., KABACK, D. B. & PRINGLE, J. R. 1986: Molecular cloning of chromosome 1 DNA from *Saccharomyces cerevisiae:* isolation and characterization of the CDC24 gene and adjacent regions of the chromosome. *Mol. Cell. Biol.* **6**:4516–4525.

CONKLE, M. T. 1981: Isozyme variation and linkage in six conifer species. *In*: Proceeding of the Symposium on Isozymes of North American Forest Trees and Forest Insects. (eds. M.T. Conkle). pp. 11–17. USDA For. Serv. Gen. Tech. Rep. PSW–48.

DEVEY, M. 1994: Super 80055 radiata pine, subject of research in Austarlia and New Zealand. *Dendrome* **1**(2): 1–3.

DOYLE, J. J. & DOYLE, J. L. 1990: Isolation of plant DNA from fresh tissue. *BRL Focus* **12**:13–15.

EKERT, R. T, JOLY, R. J. & NEALE, D. B. 1981: Genetics of isozyme variants and linkage relationships among allozyme loci in 35 eastern white pine clones. *Can. J. For. Res.* **11**:575–579.

EL-KASSABY, Y. A., SZIKLAI, O. & YEH, F. C. 1982: Linkage relationships among nineteen polymorphic allozyme loci in coastal Douglas-fir (*Pseudotsuga menziesii*). *Can. J. Genet. Cytol.* **24**:101–108.

GRATTAPAGLIA, D. 1994: Genetic mapping of quantitatively inherited economically important traits in *Eucalyptus*. PhD thesis. Forest Biotechnology Group and Departments of Genetics and Forestry, North Carolina State University, Raleigh, NC.

GURIES, R. P., FRIEDMAN, S. T. & LEDIG, F. T. 1978: A megagametophyte analysis of genetic linkage in pitch pine (*Pinus rigida* Mill.). *Heredity* **40**:309–314.

HALDANE, J. B. S. 1919: The combination of linkage values, and the calculation of distance between the loci of linked factors. *J. Genet.* **8**:299–309.

INNIS, M. A., GELFAND, G. H., SNINSKY, J. J. & WHITE, T. J. 1990: PCR Protocols: A Guide to Methods and Applications. Academic Press, San Diego.

JOHANSSON, M., ELLEGREN, H., MARKLUND, L., GUSTAVSSON, U., RINGMAR–CEDERBERG, E., ANDERSSON, K., EDFORS–LILJA, I. & ANDERSSON, L. 1992: The gene for dominant white color in the pig is closely linked to ALB and PDGFRA on chromosome 8. *Genomics* **14**:965–969.

KING, J. N. & DANCIK, B. P. 1983: Inheritance and linkage of isozymes in white spruce (*Picea glauca*). *Can. J. Genet. Cytol.* **25**:430–436.

LANDER, E. S., GREEN, P., ABRAHAMSON, J., BARLOW, A., DALY, M. J., LINCOLN, S. & NEWBURG, L. 1987: Mapmaker: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**:174–181.

LINDAHL, K. F. 1991: His and hers recombinational hotspots. *Trends In Genet.* **7**:273–276.

LUNDIN, L. G. 1993: Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics* **16**:1–19.

MCCOUCH, S. R., KOCHERT, G., YU, Z.-H., WANG, Z.-Y., KHUSH, G. S., COFFMAN, W. R. & TANKSLEY, S. D. 1988: Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* **76**:815–829.

MICHELMORE, R. W., PARAN, I. & KESSELI, R. V. 1991: Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci USA* **88**:9828–9832.

MUONA, O., YAZDANI, R. & LINDQVIST, G. 1987: Analysis of linkage in *Picea abies*. *Hereditas* **106**:31–36.

NA'IAM, M., TSUMURA, Y., UCHIDA, K., NAKAMURA, T. & OHBA, K. 1993: Linkage of allozyme loci in Japanese red pine (*Pinus densiflora*). *Can. J. For. Res.* **23**:680–687.

NEALE, D. B. & ADAMS, W. T. 1981: Inheritance of isozyme variants in seed tissues of balsam fir (*Abies balsamea*). *Can. J. Bot.* **59**:1289– 1291.

NEALE, D. B. & WILLIAMS, C. G. 1991: Restriction fragment length polymorphism mapping in conifers and applications to forest genetics and tree improvement. *Can. J. For. Res.* **21**:543–554.

NELSON, C. D., NANCE, W. L. & DOUDRICK, 1993: A partial genetic linkage map of slash pine (*Pinus elliottii* Engelm. var. *elliottii*) based on random amplified polymorphic DNAs. *Theor. Appl . Genet.* **87**:145–151.

ORKIN, S. H. 1986: Reverse genetics and human disease. *Cell* **47**:845–850.

TANKSLEY, S. D., GANAL, M. W., PRINCE, J. P., DE VICENTE, M. C., BONIERBALE, M. W. & BROUN, P. 1992: High density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**:1141–1160.

TAUER, C. G., HALLGREN, S. W. & MARTIN, B. 1992: Using marker-aided selection to improve tree growth response to abiotic stress. *Can. J. For. Res.* **22**:1018–1030.

TULSIERAM, L. K., GLAUBITZ, J. C., KISS, G. & CARLSON, J. E. 1992: Single-tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. *Bio/Technology* **10**:686–690.

TUSKAN, G. A. 1991: Marker-aided selection: A tool for the improvement of forest tree species. A workshop held in

Gatlinburg, Tennessee 13–14 June 1991.

VELDBOOM, L. R., LEE, M. & WOODMAN, W. L. 1994: Molecular marker-facilitated studies in an elite maize population: I. Linkage analysis and determination of QTL for morphological traits. *Theor. Appl. Genet.* **88**:7–16.

WELSH, J. & MCCLELLAND, M. 1990: Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **18**:213–218.

WELSH, J. & MCCLELLAND, M. 1991: Genomic fingerprinting using arbitrary primed PCR and a matrix of pairwise combinations of primers. *Nucleic Acids Res.* **19**:5275–5279.

WILLIAMS, C. G. & NEALE, D. B. 1992: Conifer wood quality and marker-aided selection: a case study. *Can. J. For. Res.* **22**:1009–1017.