Heredity **74** (1995) 661–668

Received 14 September 1994

Genomic mapping in *Pinus pinaster* (maritime pine) using RAPD and protein markers

C. PLOMION†‡*, N. BAHRMAN†, C.-E. DUREL† & D. M. O'MALLEY‡

†INRA, Laboratoire de Génétique et Amélioration des arbres forestiers, BP45, F-33610 Cestas, France and ‡Forest Biotechnology Group, Department of Forestry, North Carolina State University, Box 8008 Raleigh, NC 27695, U.S.A.

A detailed genomic map was constructed for one F_1 individual of maritime pine, using randomly amplified polymorphic DNA (RAPD) and protein markers scored on megagametophytes of germinated seeds. Proteins allowed the localization of exclusively coding DNA in the large genome of this *Pinus* species, mapped with RAPD markers that essentially fall within repetitive (i.e. mostly noncoding) DNA. Dot blots experiments of 53 RAPD fragments showed that 89 per cent amplified from highly repetitive chromosomal regions. The map comprised 463 loci, including 436 RAPDs amplified from 142 10-mer oligonucleotide primers and 27 protein loci. Twelve major and one minor linkage groups were identified using a LOD score $\leqslant 5$ and a recombination fraction $\Theta \leq 0.30$. A framework map was ordered with an interval support ≥ 4, covering 1860 cM which provided almost complete coverage of the maritime pine genome. The average distance between two framework markers was 8.3 cM; only one interval was larger than 30 cM. Protein loci were well distributed throughout the map. Their potential use as anchor points to join RAPD-based maps is discussed. Finally, the genomic maps of *Arabidopsis* and maritime pine were compared. Linkage groups were shown to have similar total map lengths on a chromosomal basis, despite a 57-fold difference in DNA content.

Keywords: 2-D electrophoresis, linkage map, *Pinus pinaster*, protein, RAPDs.

Introduction

The first linkage studies of *Pinus* were based on segregation of isozymes extracted from megagametophytes. More than 10 species were studied for about 15 loci (reviewed by Tulsieram *et al.*, 1992). Conkle (1981) located more loci but the number of markers that were resolved and analysed was still too low for applications that required a broader genome coverage (e.g. quantitative trait dissection studies). Two-dimensional electrophoresis (2D-PAGE) of megagametophyte proteins identified a larger number of loci. Bahrman & Damerval (1989) reported linkage analysis for 119 loci and Gerber *et al.* (1993) reported a 65 loci linkage map covering 530 cM of the maritime pine genome. Both isozymes and proteins correspond to coding DNA. Devey *et al.* (1994) presented linkage groups in loblolly pine for 80 RFLPs detected using cDNA probes. RFLPs typically sample genetic variation in coding regions or directly adjacent to coding regions of the genome, and use low copy probes. The RAPD method (Williams *et al.*, 1990) permits identification of a large number of polymorphic DNA markers distributed throughout the genome, including both coding and noncoding regions (Williams *et al.*, 1990). RAPDs have been used for genomic mapping in several conifer species (Neale & Sederoff, 1991; Tulsieram *et al.*, 1992; Nelson *et al.*, 1993; Binelli & Bucci, 1994). *Pinus* species have a large genome (Ohri & Khoshoo, 1986; Wakamiya *et al.*, 1993) characterized by a high proportion of repetitive DNA (Miksche & Hotta, 1973; Rake *et al.*, 1980; Kriebel, 1985). Therefore, RAPDs and markers that are based on coding sequences could provide different coverage of the genome in pine.

Each type of marker has advantages and limitations and many factors can influence choice of marker systems for a given purpose. Marker based technologies are being used for linkage map construction, quantitative traits dissection experiments, germplasm evaluation, genetic fingerprinting and manipulation of genes. Genomic maps can also be used to locate and clone genes of interest. In addition, they may provide information for understanding genome structure and evolution (Neale & Williams, 1991). RFLP methods

*Correspondence: INRA, Laboratoire de Génétique et Amélioration des arbres forestiers, BP45, F-3360 Cestas, France.

are well suited for species maps because the same hybridization probes can be used in comparisons among species. Ahuja *et al.* (1994) showed that mapped DNA probes from loblolly pine can be used to construct RFLP maps for other members of Pinaceae, thus presenting an opportunity to compare the genome in related pine species. Forest trees exhibit generally high levels of genetic diversity and are highly outcrossed. As a result, linkage disequilibrium should be low. Thus, alleles at quantitative trait loci (QTLs) and alleles at marker loci should be randomly associated in different genotypes (Strauss *et al.*, 1992). Therefore the ability to create maps for individual trees, and to assess marker genotypes on hundreds of progenies, is essential for breeding experiments that aim to use marker-assisted selection. The RAPD technology appears to be well suited for developing single-tree maps. The major advantage of this technique is the rapidity of screening for polymorphisms, the identification of a large number of markers and its potential automation (Sobral & Honeycutt, 1993). However, synteny of linkage groups with other species could be difficult to established with RAPD markers, as noticed by Torres *et al.* (1993).

We located 2-D protein polymorphisms that represent coding DNA on a RAPD-based framework map of maritime pine (*Pinus pinaster* Ait., $2x = 2n = 24$). Protein markers were used because (i) they only reveal gene products, (ii) they have been well studied in that species, and (iii) they provide interesting tools, as candidate genes, for genetic dissection of vigour and adaptative traits in the frame of the maritime pine breeding programme. The objectives of this contribution are threefold: (i) genetic map construction based on markers of coding and noncoding chromosomal regions, (ii) estimation of genome size of maritime pine, and (iii) characterization of RAPD fragments' internal sequences for copy number in the genome. The map and markers will be used for QTL mapping of height growth and adaptative traits in an F_2 progeny of the mapped tree.

Materials and methods

Plant material

The material used was 124 megagametophytes from seeds of one inter-racial hybrid (parent F_1: Corsican × Landes) of maritime pine. Individual megagametophytes were harvested after 2 weeks of germination. Each megagametophyte was cut in two parts, with one-quarter freeze-dried for RAPD assay, and three-quarters stored at −80°C and devoted to protein analysis.

DNA extraction and RAPD procedure

DNA samples were prepared from needles of the Corsican and Landes grandparents (accessions C10 and L146, respectively), and the F_1 hybrid parent (accession H12), as well as from the megagametophytes from H12. DNA was extracted using the CTAB procedure described by Bousquet *et al.* (1990). The DNA extracted from needle samples was purified by centrifugation in a CsCl-ethidium bromide density gradient. The pine DNA was diluted to a working concentration of approximately 1 ng/μL, by comparison with the fluorescence of lambda DNA concentration standards on ethidium bromide-stained agarose gel. RAPD reactions were performed using the method of Williams *et al.* (1990) with 5–10 ng template; 20 ng of ten-base primers from Operon Technologies (Alameda, CA), kits A–Z; 8 μg/μL non-acetylated bovine serum albumin and 1 unit of Taq DNA polymerase. The mixture was covered with 50 μL of mineral oil and amplifications were carried out in 96-well microtitre plates using an MJ Research PT-100 thermal cycler (MJ Research, Watertown, MA). Amplification products were separated by electrophoresis on 2 per cent agarose gels, detected by staining with ethidium bromide, and the gels were photographed. RAPD fragments were sampled from the agarose gel by stabbing the fluorescing band with a pipette tip and rinsing the tip into 100 μL of 20 per cent TE buffer (10 mM Tris-HCl, 0.2 mM EDTA), and were stored at −20°C until required for reamplification.

Characterization of genomic sequence complexity of RAPD markers

To characterize the copy number of RAPDs internal sequences, 53 RAPD fragments were labelled and used as non-radioactive probes on dot blots (strip blots with four dots containing 20 μg, 2 μg and 0.2 μg of pine DNA and 20 μg of herring sperm DNA as a negative control), as described by Grattapaglia & Sederoff (1994).

Protein extraction and electrophoresis

A sample of 34 megagametophytes was individually crushed in 6μL/mg UKS buffer (9.5 M urea, 5 mM K_2CO_3, 1.25 per cent SDS, 0.5 per cent dithiothreitol, 2 per cent pharmalyte pH 3–10 and 6 per cent Triton X-100). Thirty-five microlitres of the supernatant was submitted to electrofocusing in the first dimension, followed by a second dimension electrophoresis (Bahrman & Thiellement, 1987). The gels were silver-stained according to Damerval *et al.* (1987) in the

apparatus described by Granier & de Vienne (1986) and dried.

Scoring and nomenclature of RAPD and protein markers

Segregation of RAPD markers was recorded in four sets of 31 different megagametophytes from H12. DNA extraction, reaction mixture preparations, gel analysis and genotype scoring were performed independently for each set. This replicated design provided a control that aimed to retain RAPD markers that amplified consistently in the studied population. RAPD fragments were named by the OPERON primer code, followed by their molecular size in base pairs. Because protein analysis requires elaborate laboratory techniques, protein markers were recorded for only 34 randomly chosen megagametophytes among the mapping sample of 124 megagametophytes. The dried gels were visually scored on an illuminated box by superimposition. Three kinds of variations were scored in 2-D protein patterns: position (V), presence/absence (P) and staining intensity (I) variations. The name of each marker included the grandparental origin ('+' denoted markers inherited from the Corsican grandparent, '−' denoted markers inherited from the Landes grandparent).

Linkage analysis

A total of 463 genetic markers were tested for departure from the 1:1 Mendelian ratio of presence:absence of band. The linkage relationships of the markers were analysed with the Macintosh MAPMAKER v2.0 computer program (Lander *et al.*, 1987). Markers were considered to be linked when their LOD score was ≥ 5.0 and recombination fraction $\Theta \leq 0.30$. A subset of markers that could be ordered with an interval support ≥ 4 (i.e. difference in log likelihood between the best and alternative orders ≥ 4), provided a framework map. Accessory markers that could not meet this ordering criterion were located to the closest framework markers. Recombination distances of accessory markers to the nearest framework markers were incorporated in the marker names. Recombination fractions were converted to map distances using Kosambi's mapping function.

Results

Identification of polymorphic markers

We scored 35 protein markers in 2-D protein patterns obtained from germinated megagametophytes. Twenty spots belonging to 10 polypeptides corresponded to allelic products of structural genes varying in position, 22 spots concerned presence/absence variations and three spots involved staining intensity variations (see Bahrman & Damerval, 1989; Gerber *et al.*, 1993, for genetic analysis of each variation). They all segregated in a Mendelian fashion ($\alpha = 0.01$). From 520 OPERON primers screened for polymorphisms, 142 showed amplification and segregating RAPD markers present in one grandparent as well as in the hybrid parent and absent in the other grandparent. Out of the 142 primers, 113 and 29 were used to produce RAPD markers on four and two replicates of 31 megagametophytes, respectively. Segregation of 470 RAPD markers was scored in the whole experiment (Fig. 1). A total of 437 RAPDs were repeatable among the mapping replications. On average, one primer produced three polymorphic RAPD fragments. Fragment sizes ranged from 194 to 2627 base pairs. They all conformed to Mendelian segregation ($\alpha = 0.01$). The nonreproducible RAPD bands were discarded from further analysis. They were typically very faint, and often had a molecular weight > 2000 bp or < 200 bp.

Construction of the genomic map

Grouping and initial ordering of markers were carried out at LOD ≥ 5.0 and $\Theta \leq 0.30$. Out of 471 markers (436 RAPDs and 35 proteins), 463 loci (436 RAPDs and 27 proteins) were assigned to 13 linkage groups. Eight of the protein markers were not linked to any other locus when lowering the statistical stringency. However, lowering the LOD score to 3.0 and keeping Θ to 0.30 would result in the merging of linkage group 1 and 13 as indicated by a faint line between markers R10_767/− and O18_1207/+ (Fig. 2). Few additional mapped markers would be needed to fill the gap between these two groups. Local mapping based techniques (Reiter *et al.*, 1992) should facilitate this objective. The number of major groups corresponded to the 12 expected based on the known karyotype of maritime pine (Saylor, 1964).

A framework map was established using the RIPPLE command, to identify a subset of loci that could be locally ordered with an interval support ≥ 4. Approximately 53 per cent of the markers were placed on the framework map defining a total of 244 loci and 1860 cM of map distance. The size of linkage groups ranged from 177.9 cM to 16.6 cM. The average distance between two framework markers was 8.3 cM, with only a few gaps exceeding 20 cM. Only one interval between two markers located in linkage group 5 was larger than 30 cM. However, the LOD score for this interval was above the threshold 5. The majority of the intervals (72 per cent) were < 10 cM. Most of the accessory markers were placed within 5 cM of the nearest framework

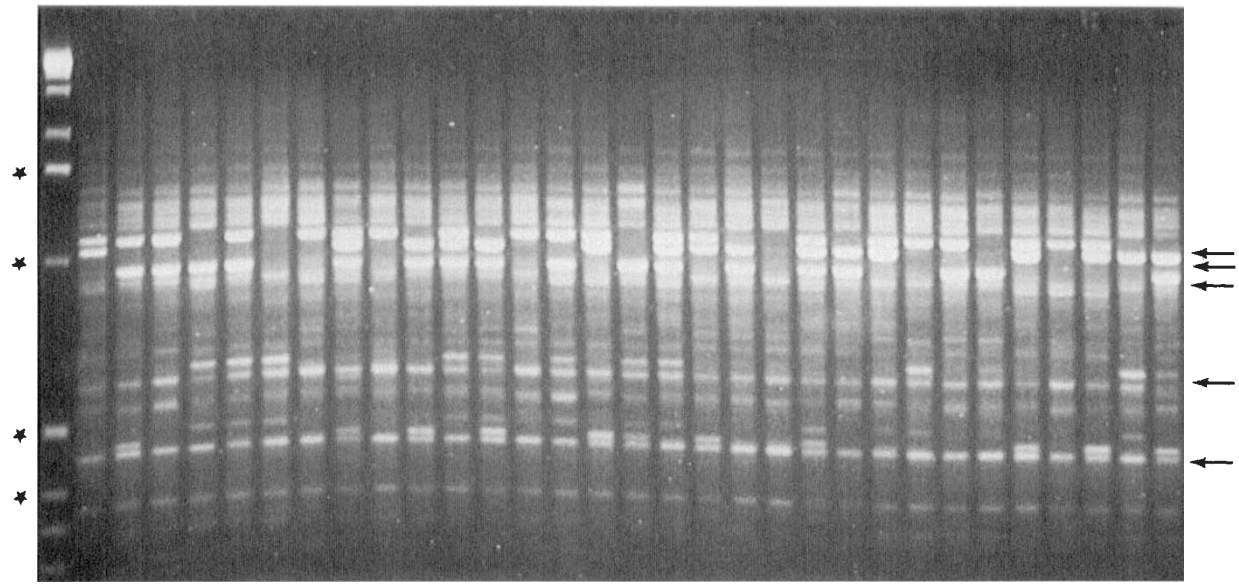

Fig. 1 Segregation of RAPD markers. The first lane is a molecular weight size standard (stars from the top correspond to 1636, 1018, 510 and 396 base pairs). Other lanes show the separation of five RAPD fragments amplified from 31 genomic DNA samples, using primer J16. Segregating RAPD markers are indicated by arrows (arrows from the top correspond to RAPD markers J16_1118/+, J16_1056/+, J16_965/−, J16_643/+, J16_479/−).

marker. The limited number of megagametophytes did not allow a precise estimate of the recombination distances and ordering of tightly linked markers. In a sample of 62 or 124 megagametophytes, approximately 95 per cent of the gametes will show no recombination in a 5 cM interval and provide no information on order. Assuming the same error rate in genotyping, we also observed that the number of accessory RAPD markers placed at distances 6 cM from the nearest framework marker was not much higher for 62 (two sets of 31 megagametophytes) than for 124 megagametophytes (four sets): 22 per cent and 18 per cent amplified from samples of 62 and 124 megagametophytes, respectively. Thus, there was little increase in precision on the relative position of accessory markers when genotyping 124 individuals instead of 62.

Assuming that the order of framework markers was correct, we used the SHOW RAW command to identify double recombinants that involved flanking loci. True double recombination events should be rare, and an excess of double recombination could indicate potential scoring errors. Double crossovers were systematicaly re-examined on gel photos. When dubious data points were found they were treated as missing data. Then the ordering analysis was performed again. Errors in genotyping were mostly from weak amplification of a specific band, smearing problems or artefacts from the loading of wells in the gel. This data quality control, in combination with the framework map construction procedure, should provide a high confidence for map length and loci order.

The 27 mapped protein loci were well distributed throughout the genomic map (Fig. 2). A total of four, three, two and one protein loci were mapped in two, two, five and three linkage groups, respectively. The three types of proteins (see Materials and methods) were represented on the map. Protein loci were genotyped on only 34 megagametophytes, which did not allow a precise estimation of two-point recombination fractions. This could explain the high proportion of unlinked protein markers (23 per cent) and the fact that almost all proteins could not meet the local order criterion used for framework map construction (interval support ≥ 4). Therefore, most protein loci were placed as accessory markers.

Copy number of RAPD fragments

Out of 53 RAPD fragments, 11 per cent did not show detectable hybridization or gave a faint signal in the 20 μg dilution. They were classified as amplifying from low-copy to moderately repetitive chromosomal regions (Fig. 3a,b). Twenty per cent gave a signal in the 20 μg and 2 μg dilutions and were classified as amplifying from highly repeated regions (Fig. 3c); 69 per cent gave a signal in all dilutions and were classified as amplifying from very highly repeated regions (Fig. 3d,e). So, 89 per cent of the RAPD fragments were

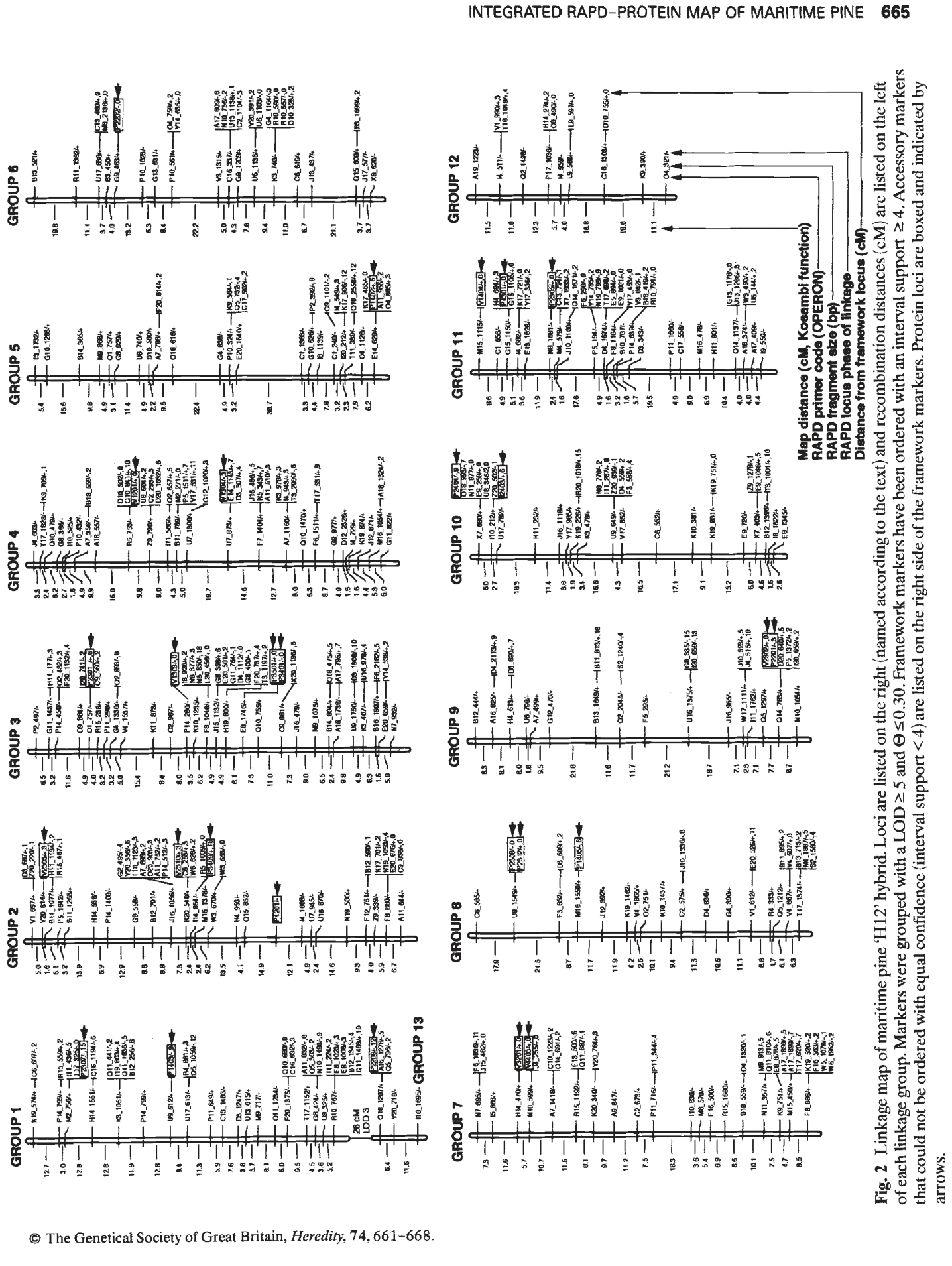

Fig. 2 Linkage map of maritime pine 'H12' hybrid. Loci are listed on the right (named according to the text) and recombination distances (cM) are listed on the left of each linkage group. Markers were grouped with a LOD ≥ 5 and $\Theta \leq 0.30$. Framework markers have been ordered with an interval support ≥ 4. Accessory markers that could not be ordered with equal confidence (interval support < 4) are listed on the right side of the framework markers. Protein loci are boxed and indicated by arrows.

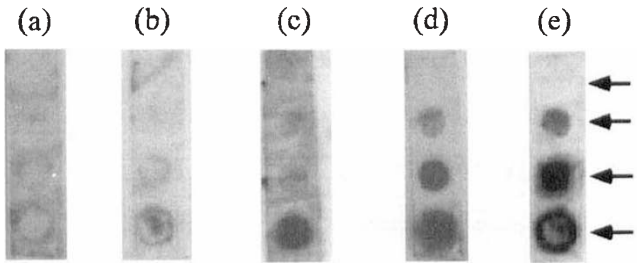

Fig. 3 Dot blot analysis of five RAPD fragments using a chemiluminescence detection assay. Arrows from the top correspond to 20 μg herring sperm DNA as a negative control, and a serial dilution of 0.2 μg, 2 μg and 20 μg of pine genomic DNA, immobilized on nylon membrane. RAPD fragments containing low-copy to moderately repetitive sequences (a,b), highly repetitive sequences (c) and very highly repetitive sequences (d,e) are shown.

presumed to contain at least some high-copy sequences, and were likely to amplify from highly repetitive chromosomal regions. Thus, RAPD fragments should be poor probes for hybridization experiments (e.g. RFLP assay) in maritime pine.

Discussion

Pine genome organization

Gymnosperm species are characterized by: (i) their antiquity (conifers appeared 140 millions years before the first angiosperm) and their longevity, (ii) the absence of ploidy level and chromosome number evolution (reviewed by Neale & Williams, 1991), and (iii) the very high and consistent amount of DNA per nucleus (Ohri & Khoshoo, 1986; Wakamiya *et al.*, 1993). The technique of DNA reassociation kinetics applied to *Pinus* species (Miksche & Hotta, 1973; Rake *et al.*, 1980; Kriebel, 1985) showed that 25 per cent of total DNA is low- to single-copy, 75 per cent being middle to highly repetitive. Thus, the vast majority of the DNA in the pine genomes is arranged in repeated sequence families. Most of this repeated DNA does not encode proteins (Thompson & Murray, 1981). If 60 000 genes are expressed during the life cycle of a plant (Kamalay & Goldberg, 1980), given an average size of a gene of 2000 bp (exons only) and the size of maritime pine genome as 24×10^6 kbp, 0.5 per cent of the genome is likely to be coding DNA. This could be an underestimate because a significant number of genes occur in multigene families (Kinlaw & Gerttula, 1993; Ahuja *et al.*, 1994; Devey *et al.*, 1994). Thus, coding DNA may not represent more than a few per cent of the pine genome. This result agrees with the estimated fraction of coding regions in plant species (Goldberg *et al.*, 1978; Thompson & Murray, 1981). The RAPD primers used for mapping consisted of random sequences that should not discriminate coding and noncoding chromosomal regions. Therefore, and at least in conifer species, most RAPD loci are likely to fall within noncoding DNA. The characterization of the internal sequence of 53 RAPD fragments for copy number in the maritime pine genome showed that, although RAPD fragments mapped to unique genomic sites, most of them contained highly repeated sequences. Conversely, protein markers sample regions of coding DNA. Our results showed that mapped protein markers were well distributed throughout the genome of maritime pine.

Single-tree map vs. species map

A species consensus map of markers and traits could be difficult to use for breeding applications in allogamous species with a wide genetic base, such as forest trees (Grattapaglia & Sederoff, 1994). Marker:trait associations are likely to be in linkage equilibrium in early generations of the breeding population and will probably have to be established for each cross independently. Mapping of individual trees using markers specific to only one cross provides a powerful approach to genetic analysis of quantitative and complex qualitative traits within families. However, genomic maps of individuals using RAPD markers can not readily be combined to make a concensus species map because the migration distance of a RAPD fragment is not sufficient information to identify uniquely a specific locus across a species. Similar problems exist for RFLP probes that recognize several bands (e.g. Tanksley *et al.*, 1988; Song *et al.*, 1991; Devey *et al.*, 1994), a problem addressed by using probes that yield only one band (e.g. Beavis & Grant, 1991). Furthermore, many individuals could be homozygous and the marker would not be available for mapping in many crosses, depending on gene frequency. The criteria for establishing synteny using RAPD markers, or multiple band RFLP markers must be more stringent, perhaps requiring parallel linkage groups having several markers in the same order in different individuals. The identity of some allozyme or protein markers (Gerber *et al.*, 1993) should be useful for establishing the correspondence of linkage groups in RAPD maps from different trees. The distribution of the 27 mapped protein loci throughout the genomic map of maritime pine is encouraging for that objective. A further advantage to using proteins as genetic markers for the mapping of quantitative or qualitative traits is that the polymorphism of a specific gene product could potentially be responsible for the mapped quantitative effect (Damerval *et al.*, 1994). Alternatively, a small number of hypervariable microsatellite markers could be

assayed in each cross to establish the correspondence of linkage groups.

Genome size of maritime pine

The protein data produced by Bahrman & Damerval (1989) and Gerber *et al.* (1993) suggested a genome size of approximately 2000 cM for maritime pine (Gerber & Rodolphe, 1994). However, framework map procedures had not been used to construct these two protein-based maps. This may lead to overestimates for genetic distances and total map length. Maritime pine has 12 metacentric chromosomes of approximately equal size (Saylor, 1964). Linkage groups 1–12 had approximately the same length (about 155 cM) and therefore should provide almost complete coverage of the genome. In addition, the genome size of the presented framework map (1860 cM) agreed with what has been found for other dense linkage maps of loblolly pine, constructed with approximately 400 RAPD markers (H. Amerson & P. Wilcox, personal communication). The relationship between recombination rates and genome size has been a matter of speculation for many years. Grant (1958) predicted that plants with long generation times, such as pine, will have genetic systems that promote recombination. Short-lived annual plants, such as *Arabidopsis*, should have genetic systems that restrict recombination. One mechanism to promote recombination could be an increased number of chromosomes. Grant (1958) also speculated that long-lived organisms such as pine might have a higher chiasma frequency to promote recombination. Our data for pine, however, do not support this idea. The total map distance per chromosome was approximately 1.55 Morgans for pine and 1.30 Morgans for *Arabidopsis* (Reiter *et al.*, 1992). Maritime pine and *Arabidopsis* have approximately 2 pg and 0.03 pg of DNA per chromosome, respectively (Ohri & Khoshoo, 1986; Arumuganathan & Earle, 1991). Thus, on a chromosomal basis, maritime pine has approximately 57-fold more DNA per cM than *Arabidopsis*. However, the number of crossovers per chromosome was almost equivalent and did not seem to be really affected by the DNA content and the proportion of coding DNA. Although large and small genomes could differ in the organization and structure of genomic DNA (John & King, 1980; Flavell *et al.*, 1985; Brown & Sundaresan, 1991), the mechanism of crossing-over must be highly conserved on a chromosomal basis and independent of physical map size and the fraction of coding DNA. This observation is consistent with other results showing that recombination per chromosome was approximately constant despite large differences in DNA amount (Rees & Durrant, 1986; Tanksley *et al.*, 1988).

Acknowledgements

The RAPD assay was completed in the Forest Biotechnology Group at North Carolina State University. We thank Dario Grattapaglia for assistance with the characterization of genomic sequence complexity of RAPD markers and spending much time discussing genetics. We thank Antoine Kremer for critical reading of the manuscript. This work was supported by a grant from the Institut National de la Recherche Agronomique (AIP Génome des plantes cultivées 1994) and from the USDA Forest Service (Coop Agreement No 29-1009).

References

AHUJA, M. R., DEVEY, M. E., GROOVER, A. T., JERMSTAD, K. D. AND NEALE, D. B. 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.

ARUMUGANATHAN, K. AND EARLE, E. D. 1991. Nuclear DNA content of some important plant species. *Plant. Mol. Biol. Rep.*, **9**, 208–218.

BAHRMAN, N. AND DAMERVAL, C. 1989. Linkage relationships of loci controlling protein amounts in maritime pine (*Pinus pinaster* Ait.). *Heredity*, **63**, 267–274.

BAHRMAN, N. AND THIELLEMENT, H. 1987. Parental genome expression in synthetic wheats (*Triticum turgidum* sp. × *T. tauschii* sp.) revealed by two-dimensional electrophoresis of seedling proteins. *Theor. Appl. Genet.*, **74**, 218–223.

BEAVIS, W. D. AND GRANT, D. 1991. A linkage map based on information from four F_2 populations of maize, *Zea mays* L.. *Theor. Appl. Genet.*, **82**, 636–644.

BINELLI, G. AND 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.

BOUSQUET, J., SIMON, L. AND LALONDE, M. 1990. DNA amplification from vegetative and sexual tissues of trees using polymerase chain reaction. *Can. J. Forest Res.*, **20**, 254–257.

BROWN, J. AND SUNDARESAN, V. 1991. A recombination hotspot in the maize A1 intragenic region. *Theor. Appl. Genet.*, **81**, 185–188.

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

DAMERVAL, C., LEGUILLOUX, M., BLAISONNEAU, J. AND de VIENNE, D. 1987. Simplification of the Heukeshoven and Dernick's silver staining of proteins. *Electrophoresis*, **8**, 158–159.

DAMERVAL, C., MAURICE, A., JOSSE, J. M. AND de VIENNE, D. 1994. Quantitative trait loci underlying gene product variation: a novel perspective for analysing regulation of genome expression. *Genetics*, **137**, 289–301.

DEVEY, M. E., FIDDLER, T. A., LIU, B. H., KNAPP, S. J. AND NEALE, B. D. 1994. An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. *Theor. Appl. Genet.*, **88**, 273–278.

FLAVELL, R. B., O'DELL, M., SMITH, D. B. AND THOMPSON, W. F. 1985. Chromosome architecture: the distribution of recombination sites, the structure of ribosomal DNA loci and the multiplicity of sequences containing inverted repeats. In: van Vloten-Doting, L., Groot, G. S. S. P. and Hall, T. C. (eds) *Molecular Form and Function of the Plant Genome. NATO ASI*, **83**, pp. 1–4. Plenum Press, New York.

GERBER, S. AND RODOLPHE, F. 1994. An estimation of the genome length of maritime pine (*Pinus pinaster* Ait.). *Theor. Appl. Genet.*, **88**, 289–292.

GERBER, S., RODOLPHE, F., BAHRMAN, N. AND BARADAT, P. 1993. Seed-protein variation in maritime pine (*Pinus pinaster* Ait.) revealed by two-dimensional electrophoresis: genetic determinism and construction of a linkage map. *Theor. Appl. Genet.*, **85**, 521–528.

GOLDBERG, R. B., HOSCHEK, G., KAMALAY, J. C. AND TIMBERLAKE, W. E. 1978. Sequence complexity of nuclear and polysomal RNA in leaves of the tobacco plant. *Cell*, **14**, 123–131.

GRANIER, F. AND de VIENNE, D. 1986. Silver staining of proteins: standardized procedure for two-dimensional gels bound to polyester sheets. *Analyt. Biochem.*, **155**, 45–50.

GRANT, V. 1958. The regulation of recombination in plants. *Cold Spring Harbor Symp. Quant. Biol.*, **23**, 337–363.

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

JOHN, B. AND KING, M. 1980. Heterochromatin variation in *Cryptobothrus chrysophorus*. 3. Synthetic hybrids. *Chromosoma*, **78**, 165–186.

KAMALAY, J. C. AND GOLDBERG, R. B. 1980. Regulation of structural gene expression in tobacco. *Cell*, **19**, 935–946.

KINLAW, C. S. AND GERTTULA, S. M. 1993. Complex pine gene families. *Proceedings of the 22nd Southern Forest Tree Improvement Conference*, June 14–17, 1993, Atlanta, Georgia, pp. 275–283. Sponsored Publication No. 44 of the Southern Forest Tree Improvement Committee.

KRIEBEL, H. B. 1985. DNA sequence components of the *Pinus strobus* nuclear genome. *Can. J. Forest Res.*, **15**, 1–4.

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

MIKSCHE, J. P. AND HOTTA, Y. 1973. DNA base composition and repetitious DNA in several conifers. *Chromosoma*, **41**, 29–36.

NEALE, D. AND SEDEROFF, R. 1991. Genome mapping in pine takes shape. *Probe*, **1**, 1–3.

NEALE, D. B. AND WILLIAMS, C. G. 1991. Restriction fragment length polymorphism mapping in conifers and applications to forest genetics and tree improvement. *Can. J. Forest Res.*, **21**, 545–554.

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

OHRI, D. AND KHOSHOO, T. N. 1986. Genome size in Gymnosperms. *Pl. Syst. Evol.*, **153**, 119–131.

RAKE, A. V., MIKSHE, J. P., HALL, R. B. AND HANSEN, K. M. 1980. DNA reassociation kinetics of four conifers. *Can. J. Genet. Cytol.*, **22**, 69–79.

REES, H. AND DURRANT, A. 1986. Recombination and genome size. *Theor. Appl. Genet.*, **73**, 72–76.

REITER, R. S., WILLIAMS, J., FELDMAN, K., RAFALSKI, J. A., TINGEY, S. V. AND SCOLNIK, P. A. 1992. Global and local genome mapping in *Arabidopsis thaliana* recombinant inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 1477–1481.

SAYLOR, L. C. 1964. Karyotype analysis of Pinus-Group Lariciones. *Silvae Genet.*, **13**, 165–170.

SOBRAL, B. W. S. AND HONEYCUTT, R. J. 1993. High output genetic mapping in polyploids using PCR-generated markers. *Theor. Appl. Genet.*, **86**, 105–112.

SONG, K. M., SUZUKI, J. Y., SLOCUM, M. K., WILLIAMS, P. H. AND OSBORN, T. C. 1991. A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci. *Theor. Appl. Genet.*, **82**, 296–304.

STRAUSS, S. H., LANDE, R. AND NAMKOONG, G. 1992. Limitations of molecular-marker-aided selection in forest tree breeding. *Can. J. Forest Res.*, **22**, 1050–1061.

TANKSLEY, S. D., BERNATZKY, R., LAPITAN, N. L. AND PRINCE, J. P. 1988. Conservation of gene repertoire but not gene order in pepper and tomato. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 6419–6423.

THOMPSON, W.F AND MURRAY, M. G. 1981. The nuclear genome: structure and function. In: Stumpf, P. K. and Conn, E. E. (eds) *The Biochemistry of Plants*, vol. 6, *Proteins and Nucleic Acids*, pp. 1–81. Academic Press, New York.

TORRES, A. M., WEEDEN, N. F. AND MARTIN, A. 1993. Linkage among isozyme, RFLP and RAPD markers in *Vicia faba*. *Theor. Appl. Genet.*, **85**, 937–945.

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

WAKAMIYA, I., NEWTON, R. J., JOHNSTON, J. S. AND PRICE, H. J. 1993. Genome size and environmental factors in the genus *Pinus*. *Am. J. Bot.*, **80**, 1235–1241.

WILLIAMS, J. G. K., KUBELIK, A. R., LIVAK, K. J. AND RAFALSKI, J. A. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acids Res.*, **18**, 6531–6535.