Mol Gen Genet (1998) 258: 466–478

ORIGINAL PAPER

G. P. Paglia · A. M. Olivieri · M. Morgante

Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.)

Received: 5 November 1997 / Accepted: 16 March 1998

Abstract Genetic linkage maps have been produced for a wide range of organisms during the last decade, thanks to the increasing availability of molecular markers. The use of microsatellites (or Simple Sequence Repeats, SSRs) as genetic markers has led to the construction of "second-generation" genetic maps for humans, mouse and other organisms of major importance. We constructed a second-generation single-tree genetic linkage map of Norway spruce (*Picea abies* K.) using a panel of 72 haploid megagametophytes with a total of 447 segregating bands [366 Amplified Fragment Length Polymorphisms (AFLPs), 20 Selective Amplification of Microsatellite Polymorphic Loci (SAMPLs) and 61 SSRs, each single band being treated initially as a dominant marker]. Four hundred and thirteen markers were mapped in 29 linkage groups (including triplets and doublets) covering a genetic length of 2198.3 cM, which represents 77.4% of the estimated genome length of *Picea abies* (approximately 2839 cM). The map is still far from coalescing into the expected 12 chromosomal linkage groups of Norway spruce (2n = 2x = 24). A possible explanation for this comes from the observed non-random distribution of markers in the framework map. Thirty-eight SSR marker loci could be mapped onto 19 linkage groups. This set of highly informative Sequence Tagged Sites (STSs) can be used in many aspects of genetic analysis of forest trees, such as marker-assisted selection, QTL mapping, positional cloning, gene flow analysis, mating system analysis and genetic diversity studies.

Key words *Picea abies* · Linkage map · Sequence-tagged sites · AFLP · SAMPL

Communicated by H. Saedler

G. P. Paglia · A. M. Olivieri · M. Morgante (✉)
Dipartimento di Produzione Vegetale e Tecnologie Agrarie, Università degli Studi di Udine, Via delle Scienze 208, I-33100 Udine, Italy
Fax. +39-432-558603; e-mail: Morgante@DPVTA.Uniud.it

Introduction

Norway spruce (*Picea abies* K.) is an ecologically and economically important conifer species in Europe. There is a general lack of detailed genetic studies in this species as well as in other gymnosperms, when compared to angiosperms. Genetic maps allow the association of genome segments with phenotypes and they are a prerequisite for the positional cloning of genes of interest and the mapping of loci contributing to complex traits. They have proven to be very efficient and useful basic tools for genetic analysis in conifers (Nelson et al. 1993; Devey et al. 1994, 1996; Plomion et al. 1995a, 1996; Echt and Nelson 1997). Marker-assisted selection strategies are currently utilised in crop and animal breeding, and they also promise to be useful in studies on forest trees that are directed towards obtaining faster genetic improvement in timber quality, growth rate, stress and disease tolerance (Grattapaglia and Sederoff 1994; Plomion et al. 1996). However, for all of these applications high-density linkage maps are highly desirable. In long-lived, undomesticated and highly outbred species like conifers, however, map construction is complicated by the lack of suitable pedigrees and controlled genetic crosses due to deleterious effect of high genetic load and the long generation time. To overcome these problems the construction of single-tree linkage maps using haploid megagametophytes, which in conifers represent female meiotic segregation products, has been proposed (Tulseriam et al. 1992).

Several genetic maps have been constructed so far in conifers (Nelson et al. 1993, 1994; Binelli and Bucci 1994; Devey et al. 1994, 1996; Plomion et al. 1995a; Echt and Nelson 1997). Most of them are based either on a relatively small number of markers, or on time-consuming and technically demanding marker techniques (Restriction Fragment Length Polymorphisms, RFLPs) or on the use of markers that are not highly repeatable and transferable (Random Amplified Polymorphic DNAs, RAPDs) (Jones et al. 1997). The only genetic

map so far reported in Norway spruce was obtained using RAPD markers (Binelli and Bucci 1994).

Norway spruce, like other conifers, has a very large genome ($3–4 \times 10^{10}$ bp; Govindaraju and Cullis 1991), which is known to be largely composed of repetitive sequences (Miksche and Hotta 1973; Rake et al. 1980). This can represent a problem in the development of markers (Pfeiffer et al. 1997). The choice of appropriate molecular markers is therefore critical when constructing a genetic linkage map in such species. The use of arbitrary multilocus PCR-based molecular marker systems can increase mapping efficiency even in species for which no previous DNA sequence information is available. This applies to fingerprinting techniques such as RAPDs, but is particularly true for AFLPs (Vos et al. 1995) and SAMPLs (Morgante and Vogel 1994), which have a higher effective multiplex ratio, i.e. a higher number of polymorphic loci that are detectable in a single assay/gel lane (Powell et al. 1996), and higher repeatability when compared with RAPDs (Karp et al. 1997). The modification of the original AFLP and SAMPL protocols to provide a procedure that targets hypomethylated regions of the genome has made it possible to extend the applicability of these techniques to species with very large genomes, such as *Picea abies* (Paglia and Morgante 1998). High-throughput PCR-based multiplex fingerprinting techniques therefore offer the chance of rapidly generating genetic linkage maps which can serve as a skeleton on which other classes of markers can be mapped.

Microsatellite or Simple Sequence Repeat (SSR) markers have been generally recognised to be an excellent marker system. Besides having the advantage of being STSs (Sequence Tagged Sites, Olson et al. 1989), they have proven to be highly repeatable, widely and uniformly distributed, often codominant (although null alleles are sometimes observed, Morgante et al. 1994), suitable for automated detection, and, above all, are the most informative markers because of their hypervariability (Goodfellow 1992, 1993; Powell et al. 1996). The initial development of SSRs is quite an expensive and time-consuming task; however, their ease of use and low cost compensate the for primary effort (Rafalski and Tingey 1993). They have been extensively used in animal species for genetic and physical mapping, as well as for studies of population structure, analysis of parentage, and gene flow (Bruford and Wayne 1993). Considerable progress has been achieved in mammalian genome mapping as a result of the transition from RFLP-based maps to maps based on microsatellites (SSRs), also defined as second-generation linkage maps (Weissenbach et al. 1992; Goodfellow 1993; Dib et al. 1996; Dietrich et al. 1996).

More recently the use of SSRs has been extended to plant species such as soybean (Akkaya et al. 1992; Morgante and Olivieri 1993), maize (Senior and Heun 1993; Taramino and Tingey 1996), rice (Wu and Tanksley 1993; Zhao and Kochert 1993), and conifers (Smith and Devey 1994; Kostia et al. 1995; Devey et al. 1996; van der Ven and McNicol 1996; Echt and Nelson 1997; Pfeiffer et al. 1997). The identification of SSR markers in species with large genomes, such as conifers and wheat, is made more difficult by the high proportion of primer pairs that amplify multiple bands (Smith and Devey 1994; Kostia et al. 1995; Röder et al. 1995; Pfeiffer et al. 1997). Fully informative, multiallelic markers, such as SSRs, which can unambiguously identify all the alleles transmitted from the parents to the offspring (up to 4 alleles in a two-generation pedigree of a diploid species such as Norway spruce) are, however, especially desirable (Grattapaglia and Sederoff 1994) in conifers, due to the difficulty – in some instances the impossibility – of carrying out suitable genetic crosses.

In this work we have tried to achieve four main objectives: (i) the construction of a genetic linkage map of *P. abies* based on markers obtained from hypomethylated sequences (*Pst*I-AFLPs and SAMPLs); (ii) the genetic mapping of a significant set of SSRs which represents the first attempt to advance towards a second-generation (SSR-based) linkage map in conifers, as has been done in mammals and in some crop species; (iii) estimation of genome size and degree of coverage, and the distribution of markers achieved using hypomethylated sequences and SSRs; and (iv) we wished to show that, through extensive use of automation, a high throughput can be achieved using PCR-based markers and that such markers are therefore useful for the construction of genetic maps and in marker-assisted selection in conifer species.

Materials and methods

Plant material and DNA isolation

The segregating population for map construction was composed of 72 haploid megagametophytes from tree V23. Megagametophytes were isolated by removing the outer and inner seed coats and the embryo with a scalpel. Norway spruce individual V23 from the Italian population of Val Meledrio was used as mother tree due its high heterozygosity (Pfeiffer et al. 1997). A population size of 72 was chosen to take full advantage of the available automation facilities. DNA was extracted following the procedure of Doyle and Doyle (1990) with modifications (an extraction step using phenol:chloroform:isoamyl alcohol 25:24:1 v:v was introduced prior to the chloroform:isoamyl alcohol 24:1 v:v extraction indicated in the original protocol).

Microsatellite isolation and segregation analysis

Three genomic libraries were constructed – two non-enriched libraries (Pfeiffer et al. 1997) from which 46 primer pairs were selected, and one enriched for the presence of (AC/GT)n sequences, from which 50 primers were selected (Paglia, Magni and Morgante, in preparation). These 96 primer pairs were used to amplify the DNA of the 72 megagametophytes of the mapping panel. For each of the 96 primer pairs, conditions of amplification (i.e. annealing temperature) and product separation were tested on a subset of six individual megagametophytes chosen from the panel of 72. Products which could not be well resolved on agarose gels were run on polyacrylamide gels, following radioactive labelling of one of the two primers. In both cases PCR reactions were set up in 25-μl reactions containing 2.5 ng of genomic DNA, 200 μM of each

dNTP, 50 mM KCl, 10 mM TRIS-HCl (pH 8.3), 1.5 or 2.5 mM $MgCl_2$ and 0.8 units of AmpliTaq Gold Taq polymerase (Perkin Elmer). In the case of radioactive labelling, 2.5 pmol of one of the two primers (here called primer A) were combined with 1.5 μCi of [γ^{33}P]ATP and 0.04 units of T4 polynucleotide kinase in 1 × kinase buffer and incubated for 45′ at 37° C. The enzyme was deactivated by heating for 10 min at 70° C. The labelling reaction was then combined with the PCR mix containing 0.2 μM of unlabelled primer B and 0.1 μM of unlabelled primer A. For non-radioactive PCR 0.2 μM of each primer were used. PCRs were performed in a Perkin-Elmer 9600 thermal cycler or a Hybaid Touch Down thermal cycler according to the following profile: (i) one cycle at 95° C for 10 min; (ii) 40 cycles of 94° C for 45 s, 53–60° C (depending on the primers used) for 45 s, 72° C for 45 s; and (iii) a final cycle at 72° C for 10 min. Non-radioactive amplification products were separated on 3.5% Metaphor (FMC BioProducts) agarose gels in 1 × TBE (1 × TBE is 89 mM TRIS-borate, 2 mM EDTA pH 8.3) stained with 1 × SYBR Green (FMC BioProducts). Gels were photographed with a Polaroid film on a UV transilluminator using the SYBR Green gel stain filter (FMC BioProducts). Radioactive amplification products were added to an equal amount of loading buffer and 4.5 μl of this solution was loaded on a 6% Hydrolink Long Ranger gel containing 7 M urea and 1.2 × TBE. The gel was dried and exposed to X-ray film (Kodak BioMax) using standard autoradiographic procedures.

AFLP and SAMPL analysis

For the AFLP and SAMPL assay, 500 ng of genomic DNA from Norway spruce was used following the protocols described by Paglia and Morgante (1998). In all, 29 AFLP and 2 SAMPL primer combinations were used.

Streamlining of genotype analysis

The procedures were automated where possible. PCR reaction mixes were set up in 96-well plates using in most cases a Biomek 2000 Workstation (Beckman Instruments). Amplified samples were then loaded onto agarose gels in a fully automated fashion: a Biomek 1000 Workstation was modified (DuPont) to load 72 samples into custom-made electrophoresis cells, so that up to 720 samples could be run by a single person in a day.

Linkage analysis of markers

All segregating bands were scored for presence or absence in each of the 72 individual megagametophytes and tested for 1:1 segregation ratio using the chi-squared (χ^2) test. The scores were first entered on a spreadsheet, considering all markers as dominant; then data were converted into a text-file which was used as input file for the mapping software MAPMAKER 3.0 (Lander et al. 1987). Marker phenotypes were reciprocally coded for each locus so that both coupling- and repulsion-phase linkages could be detected from the arbitrary linkage-phase data using the MAPMAKER backcross data type. Preliminary grouping of marker loci (two point analysis) was performed at LOD 4.3 and recombination fraction (θ) ≤ 0.35. Linkage groups obtained from this doubled data-set of 894 markers were then divided into two symmetrical sets of groups, and markers were chosen from one of the two sets identified to make a new data set of 447 markers. A subset of highly informative, evenly spaced markers was subsequently ordered using full multipoint analysis at an initial LOD score threshold of 3.0, to obtain the framework map. Additional markers were then added to the map by lowering the multipoint LOD threshold to 2.0. Markers that could not be ordered with equal confidence were indicated as accessory markers at an already specified locus on the map and do not contribute to the overall genetic distance of the framework map. Once the framework map was obtained, the command "Genotype" of MAPMAKER 3.0 (Lander et al. 1987) was used to identify double crossovers. Since double recombination events should be rare, they were systematically re-examined on gel photos or X-ray films. When dubious data points were found, they were treated as missing data. After this data quality control, the ordering analysis was carried out again to give the final version of the map. The Kosambi (1944) function was used to calculate genetic distances between linked markers.

Estimation of genome length

The estimate of genome length was done following the method of Hulbert et al. (1988) and method No.3 of Chakravarti et al. (1991), using only framework markers to avoid an overestimate of genome coverage (Grattapaglia and Sederoff 1994).

Genomic distribution of markers

The comparison between theoretical and observed map coverage was performed by estimating the minimum number of randomly distributed markers required to cover a proportion, p, of a genome of size k with a maximum distance ($2c$) between markers. This is given by the following equation of Lange and Boehnke (1982): $N = 1.25\ ((\log(1-p))/\log(1-(2c)/k))$.

The occurrence of clusters of crossovers and clusters of markers in the map was examined according to Dietrich et al. (1994). Each linkage group was transformed into a string of the type "*mmmcccmmmmccmcm*....." where each *m* denotes the occurrence of a marker and each *c* denotes the occurrence of a crossover. In such a form the position of each marker relative to that of each crossover can be determined and the hypothesis that markers are randomly distributed with respect to crossovers can be tested by comparing the observed distribution with the distribution expected for a biased coin tossed n times (where n is the number of markers or crossovers) with probability Pm of being "*m*" and probability Pc of being "*c*", where $Pm = M/(M+C)$ and $Pc = 1 - Pm$; M = total number of markers, C = total number of crossovers. If R_n denotes the longest run of consecutive heads when the coin is flipped n times, the expected value of R_n is $\mu = \log_{1/p}((n-1)(1-p)+1)$, where p is the probability of heads, and the distribution of R_n is given approximately by $\text{Prob}(R-\mu > 1) = 1 - \exp(-p^t)$. This analysis was performed on framework markers only (284 markers).

Results

Molecular markers

We were able to score 366 segregating AFLP bands using 29 primer combinations, yielding an average of 12.6 bands (s.d. = 5.8), a maximum of 27 and a minimum of 3 bands per assay. Twenty polymorphic bands (9 and 11 bands, respectively) were obtained with two different SAMPL primer combinations. Only 34 microsatellite primer pairs out of the 96 tested produced at least one segregating band (Table 1, Fig. 1). The total number of segregating SSR bands (each being treated initially as a dominant marker) was 61. The total number of markers used for linkage analysis was 447. Segregation ratios that departed from the Mendelian expectation of 1:1 at $\alpha = 0.05$ were detected for a total of 25 bands (22 AFLPs, 2 SSRs and 1 SAMPL markers), which is very close to the value expected (23.5) to occur by chance alone with a probability of 5%. No clustering of markers showing distorted segregation was apparent.

Table 1 SSR polymorphic bands used in linkage analysis

Name[a]	Size (bp)	Type of segregation[b]	Linkage group	Framework marker[c]
EAC1D10a	230	Codominant	DO7	Yes
EAC1D10b	136	Codominant	DO7	Yes
EAC1F4a	351	Codominant	K	Yes
EAC1F4b	233	Codominant	K	Yes
EAC1G3a	101	Codominant	DO6	Yes
EAC1G3b	75	Codominant	DO6	Yes
EAC1G5a	280	Codominant	T	Yes
EAC1G5b	234	Codominant	T	Yes
EAC6A6a	117	Codominant	A	Yes
EAC6A6b	99	Codominant	A	Yes
EAC6B2a	191	Codominant	F	Yes
EAC6B2b	135	Codominant	F	Yes
EAC6C10a	70	Codominant	D	Yes
EAC6C10b	64	Codominant	D	Yes
EAC6D11	139	Dominant	S	Yes
EAC6E2a	126	Codominant	A	No
EAC6E2b	100	Codominant	A	No
EAC6E2c	120	Dominant	M	Yes
EAC6E2d	90	Dominant	G	Yes
EAC6E9a	180	Codominant	L	Yes
EAC6E9b	120	Codominant	L	Yes
EAC6F4a	115	Dominant	C	Yes
EAC6F4b	104	Dominant	I	Yes
EAC6G3a	192	Dominant	D	Yes
EAC6G3b	185	Dominant	J	Yes
EAC6G7	141	–	Unlinked	–
EAC6G8	159	Dominant	J	Yes
EAC7A7	225	Dominant	M	No
EAC7B9	90	Dominant	E	Yes
EAC7D10a	148	Codominant	V	Yes
EAC7D10b	124	Codominant	V	Yes
EAC7F10a	136	Codominant	G	Yes
EAC7F10b	124	Codominant	G	Yes
EAC7F6	230	Dominant	A	Yes
EAC7F8a	238	Codominant	H	Yes
EAC7F8b	218	Codominant	H	Yes
EAC7G7	414	Dominant	D	Yes
EAC7H7a	130	Codominant	DO4	Yes
EAC7H7b	104	Codominant	DO4	Yes
L2AGH1	150	Dominant	A	Yes
NACB6a	190	Codominant	F	Yes
NACB6b	160	Codominant	F	Yes
NACG7a	190	Codominant	E	Yes
NACG7b	180	Codominant	E	Yes
SpAC1B8a	260	Dominant	D	No
SpAC1B8b	180	Dominant	D	Yes
SpAC1H8a	135	Codominant	S	Yes
SpAC1H8b	119	Codominant	S	Yes
SpAGC1a	111	Codominant	DO7	Yes
SpAGC1b	103	Codominant	DO7	Yes
SpAGC2	126	Codominant	C	Yes
SpAGC2b	104	Codominant	C	Yes
SpAGD1a	185	Codominant	B	Yes
SpAGD1b	147	Codominant	B	Yes
SpAGG3a	150	Codominant	B	Yes
SpAGG3b	136	Codominant	B	Yes
SpAGH1a	141	Codominant	J	Yes
SpAGH1b	129	Codominant	J	Yes
SpL3AG1A4	90	Dominant	T	Yes
SpL3AG1H4a	204	Codominant	C	Yes
SpL3AG1H4b	180	Codominant	C	Yes

[a] SSR primer names are followed by *lower case letters* when two or more polymorphic scorable bands were detected and used for linkage analysis

[b] Those bands which were amplified by the same primer pair and cosegregated in repulsion phase were considered codominant markers after linkage analysis

[c] Only those markers which could be ordered with a LOD score ≥2.0 were considered framework markers (see Materials and methods)

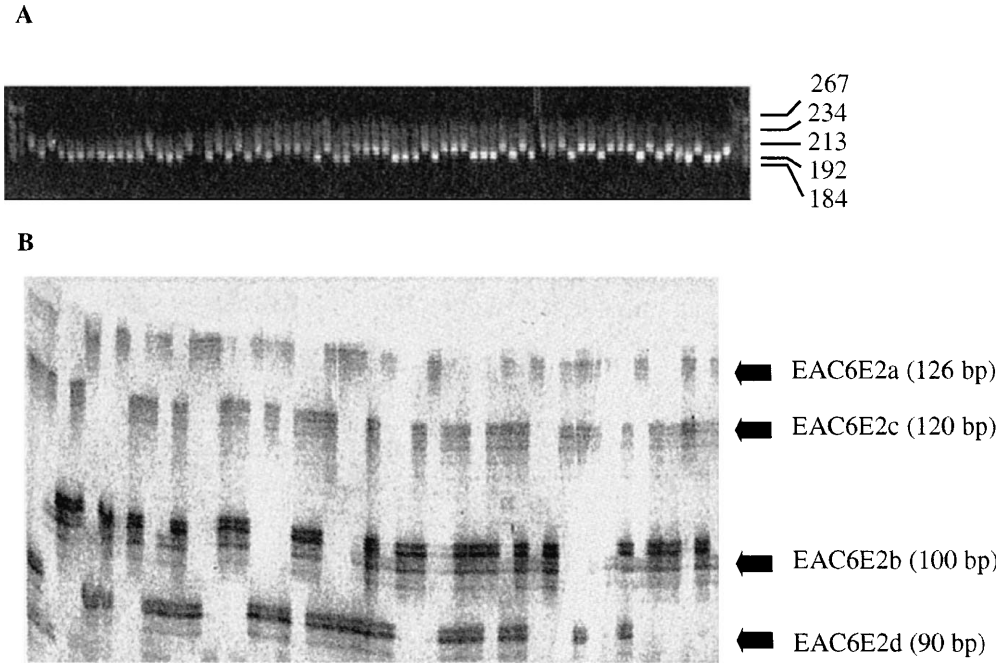


Fig. 1A, B Segregation of SSR markers in megagametophytes of Norway spruce (*Picea abies* K.). **A** Agarose gel showing the segregation of the codominant SSR marker SpAG1H4 in the mapping panel of 72 individual megagametophytes. The agarose gel was automatically loaded using a modified Beckman 1000 workstation and stained with SYBR Green (FMC BioProducts). The first two lanes on the left and the last two lanes on the right were loaded with DNA molecular weight marker V (Boehringer Mannheim). The 72 lanes in between represent amplifications from the mapping panel used for map construction. Sizes in bp are indicated on the right. **B** Polyacrylamide gel showing the segregation of marker fragments amplified with the SSR primer EAC6E2 in 48 megagametophytes representing a subset of the mapping panel of 72. *Arrows* indicate the four polymorphic bands which were amplified by the SSR primer. Each band was assigned a letter following the marker name. Sizes (in bp) of the segregating bands are indicated in *parentheses*

Linkage map statistics, genome coverage and length estimation

Linkage relationships of the 447 segregating markers were established using two-point analysis at a LOD score of 4.3 (maximum $\theta = 0.35$). Under these conditions 29 linkage groups (21 major linkage groups, 1 triplet and 7 doublets) containing 413 markers were established, while 34 markers remained unlinked (Fig. 2). Linkages were robust at a LOD score range from 4.0 to 5.0. However, LOD scores lower than 4.3 occasionally produced spurious linkages resulting from the agglomeration of the biggest linkage group with a smaller one, while increasing LOD scores up to 5.0 did not change the overall linkage group frame, barring a few more markers (14) that remained unlinked. The map is still far from coalescing into the twelve chromosomal linkage groups, despite the over 400 markers used, indicating that there are many large gaps with few markers. The 21 linkage groups featuring more than three unique positions in the framework map were assigned alphabetical designations from A to V from the larger to the smaller in size (genetic distance), while triplets and doublets were designated with TR and DO, respectively, followed by numbers. The maximum number of markers in a single linkage group was 59, the average length of linkage groups (in map units) was 75.8 cM (s.d. = 59.8 cM). Fifty-two AFLP bands (14% of the total AFLP bands) resulted in 26 codominant marker loci after linkage analysis (to be considered codominant the two bands had to be amplified by the same primer pair, cosegregate in repulsion phase and differ only slightly in size). None of the SAMPL bands produced codominant markers.

Twenty-one of the microsatellite primer pairs gave two codominant bands, one gave one codominant and two dominant bands, nine gave a single dominant band and three gave two dominant bands. All SSRs, with the exception of a codominant one that remained unlinked, were mapped to 38 loci spanning 19 linkage groups. Interestingly, we found two SSRs (DO7) that could not be separated because they cosegregated perfectly in our mapping population of 72 individuals. The two SSRs however differ in sequence [SpAGC1 amplifies an (AG) microsatellite while EAC1D10 amplifies an (AC) microsatellite].

Two hundred and eighty-four markers (distributed in 237 unique positions or bins) constituted the framework map, covering a genetic length of 2198.3 cM. The other 129 markers that could not be placed in the framework map do not contribute to the overall genetic distance and were placed as ancillary markers. The average genetic distance between two bins in the framework map is 9.3 cM.

The total genome size of *P. abies* was estimated following the method described by Hulbert et al. (1987)

and method No. 3 by Chakravarti et al. (1991) (Table 2) and taking only the 284 framework markers into account. The expected genome size ranged between 2783 cM and 2899 cM, with an average of 2839 cM. The framework map covers 77.4% of the total genetic distance according to this estimate.

Genomic distribution of markers

Estimates of genome coverage (Lange and Boehnke 1982) indicate that 409 randomly distributed markers should cover 95% of the Norway spruce genome (Table 3) and that a coverage similar to that obtained by us should have been reached with approximately 220 markers. In our case the use of 447 markers did not result in the predicted coverage and the map did not coalesce into the 12 linkage groups corresponding to the 12 *P. abies* chromosomes. It has been hypothesised in similar cases (Kesseli et al. 1994; Echt and Nelson 1997) that the assumption of randomness of marker distribution could have been violated. To examine this possibility further, we analysed our marker distribution with reference to the occurrence of clusters of crossovers and/or clusters of markers in the map, following the method used by Dietrich et al. (1994).

We observed (Table 4) an excess of large clusters of crossovers between consecutive markers (classes between 8 and 12 crossovers per interval) and a deficiency of small clusters (classes between one and two crossovers per interval) with respect to the expectation (Fig. 3a). The number of markers occurring between consecutive crossovers fits the expectation quite well, except that there are more clusters containing four markers than expected (Fig. 3b).

Discussion

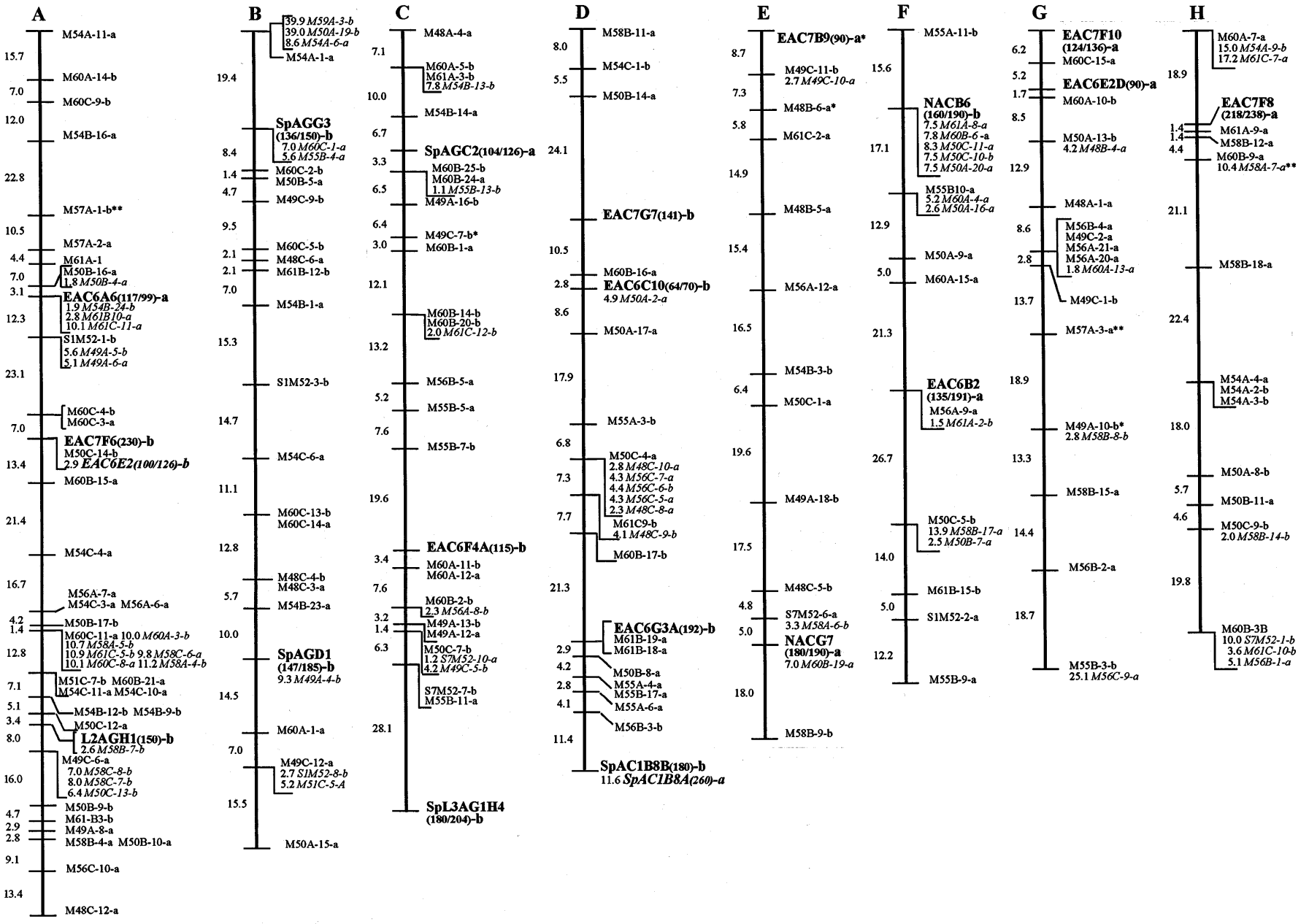
Genetic linkage maps have been produced in the past ten years for humans, for model species (mouse, fruitfly) and for many of the major crop species. Efforts have been made to construct genetic maps of tree species, which, owing to their characteristically long lifespan, in natural populations, are not ideal experimental organisms. In the recent past the need for more informative genetic maps has been recognised, i.e. maps that not only contain more markers, but that contain markers with a high information content (Goodfellow 1992). These so-called second-generation maps have been rapidly constructed using microsatellites (SSRs) (Rohrer et al. 1994; Dib et al. 1996; Dietrich et al. 1996), which may be considered to be ideal markers because they are randomly distributed, very frequent, PCR-based (and therefore amenable to automated assay) and highly informative. In the latest release of the human Généthon genetic map, Dib et al. (1996) reported that the 5246 mapped SSRs have a mean heterozygosity of 70%, and can therefore be useful for mapping rare monofactorial diseases and mapping of loci contributing to complex traits. A high information content means that the marker will segregate and thus will be useful in most of the mapping populations. This is also important for mapping in trees, where it is often difficult to produce and utilise appropriate crosses. Traditional mapping strategies such as three-generation outbred pedigrees (Devey et al. 1994), and more recent ones, such as the pseudo-testcross mapping strategy (Grattapaglia and Sederoff, 1994), and haploid megagametophyte mapping (Conkle 1981; Tulsieram et al. 1992) have the potential to overcome these limitations, especially when coupled with the use of highly polymorphic, fully informative, multiallelic markers such as SSRs that can detect all four allelic variants in a mating configuration. Pfeiffer et al. (1997) found from 6 to 22 alleles per locus and an average expected heterozygosity of 79% in a panel of 18 trees representing different populations of Norway spruce, using 7 of the SSRs which were also used in the present work.

Strategies for map construction in conifers

This is the first time that AFLPs and SAMPLs have been used for mapping purposes in conifers. These classes of markers have now been adapted to a species – Norway spruce – with a large genome (Paglia and Morgante 1998) and were chosen because of their high effective multiplex ratio, i.e. the number of polymorphic loci that are detectable in a single assay/gel lane (Powell et al. 1996), and their higher repeatability when compared to RAPDs (Karp et al. 1997). Furthermore, an inherent characteristic of AFLP and SAMPL is that of offering the possibility to produce a very large number of markers starting from a very limited amount of genomic DNA (250–500 ng). Thanks to their efficiency, these PCR-based multiplex DNA fingerprinting techniques rapidly provided a frame onto which SSRs could be

►

Fig. 2 Linkage map of *P. abies* K. Genetic distances (in cM, Kosambi) are listed on the left-hand side and loci are listed on the right-hand side of the linkage group. AFLP and SAMPL loci are designated by the primer combination with which they were detected (Paglia and Morgante 1998), followed by a number indicating the relative position of the marker in the gel lane, starting from the bands showing higher molecular weights. SSR markers appear in the map in *bold type* and are designated by the primer pairs, followed by letters when more than one polymorphic band was produced (Table 1, Fig. 1). SSR marker sizes (in bp) are shown in *parentheses* beside the name of the segregating band (only one size is shown for dominant SSRs, two sizes corresponding to those of the two allelic bands are shown for codominant SSRs). All markers in the reading phase are designated with an "a" following the marker identification code in the map, while recoded markers are indicated with a "b". Framework markers were ordered with an interval support of at least ≥2. Accessory markers that could not be ordered with equal confidence are shown in *italics* to the right of the nearest upper framework marker, together with the multipoint distance in cM from it. Markers showing segregation distortion are indicated by *one* or *two asterisks*, depending on the level of significance ($P < 0.05$ or $P < 0.01$, respectively) of the χ^2 test for 1:1 segregation


A
M54A-11-a
15.7
M60A-14-b
7.0
M60C-9-b
12.0
M54B-16-a
22.8
M57A-1-b**
10.5
M57A-2-a
4.4
M61A-1
7.0
M50B-16-a
1.8 M50B-4-a
3.1
EAC6A6(117/99)-a
1.9 M54B-24-b
2.8 M61B10-a
10.1 M61C-11-a
12.3
S1M52-1-b
5.6 M49A-5-b
5.1 M49A-6-a
23.1
M60C-4-b
M60C-3-a
7.0
EAC7F6(230)-b
M50C-14-b
2.9 EAC6E2(100/126)-b
13.4
M60B-15-a
21.4
M54C-4-a
16.7
M56A-7-a
M54C-3-a M56A-6-a
4.2
M50B-17-b
1.4
M60C-11-a 10.0 M60A-3-b
10.7 M58A-5-b
12.8
10.9 M61C-5-b 9.8 M58C-6-a
10.1 M60C-8-a 11.2 M58A-4-b
7.1
M51C-7-b M60B-21-a
M54C-11-a M54C-10-a
5.1
M54B-12-b M54B-9-b
3.4
M50C-12-a
8.0
L2AGH1(150)-b
2.6 M58B-7-b
16.0
M49C-6-a
7.0 M58C-8-b
8.0 M58C-7-b
6.4 M50C-13-b
4.7
M50B-9-b
2.9
M61-B3-b
2.8
M49A-8-a
M58B-4-a M50B-10-a
9.1
M56C-10-a
13.4
M48C-12-a
(this linkage group is 65% smaller than the actual size)
B
39.9 M59A-3-b
39.0 M50A-19-b
8.6 M54A-6-a
M54A-1-a
19.4
SpAGG3
(136/150)-b
7.0 M60C-1-a
8.4
5.6 M55B-4-a
1.4
M60C-2-b
M50B-5-a
4.7
M49C-9-b
9.5
M60C-5-b
2.1
M48C-6-a
2.1
M61B-12-b
7.0
M54B-1-a
15.3
S1M52-3-b
14.7
M54C-6-a
11.1
M60C-13-b
M60C-14-a
12.8
M48C-4-b
M48C-3-a
5.7
M54B-23-a
10.0
SpAGD1
(147/185)-b
9.3 M49A-4-b
14.5
M60A-1-a
7.0
M49C-12-a
2.7 S1M52-8-b
5.2 M51C-5-A
15.5
M50A-15-a
C
M48A-4-a
7.1
M60A-5-b
M61A-3-b
7.8 M54B-13-b
10.0
M54B-14-a
6.7
SpAGC2(104/126)-a
3.3
M60B-25-b
M60B-24-a
1.1 M55B-13-b
6.5
M49A-16-b
6.4
M49C-7-b*
3.0
M60B-1-a
12.1
M60B-14-b
M60B-20-b
2.0 M61C-12-b
13.2
M56B-5-a
5.2
M55B-5-a
7.6
M55B-7-b
19.6
EAC6F4A(115)-b
3.4
M60A-11-b
M60A-12-a
7.6
M60B-2-b
2.3 M56A-8-b
3.2
M49A-13-b
1.4
M49A-12-a
6.3
M50C-7-b
1.2 S7M52-10-a
4.2 M49C-5-b
S7M52-7-b
M55B-11-a
28.1
SpL3AG1H4
(180/204)-b
D
M58B-11-a
8.0
M54C-1-b
5.5
M50B-14-a
24.1
EAC7G7(141)-b
10.5
M60B-16-a
2.8
EAC6C10(64/70)-b
4.9 M50A-2-a
8.6
M50A-17-a
17.9
M55A-3-b
6.8
M50C-4-a
2.8 M48C-10-a
4.3 M56C-7-a
7.3
4.4 M56C-6-b
4.3 M56C-5-a
2.3 M48C-8-a
7.7
M61C9-b
4.1 M48C-9-b
M60B-17-b
21.3
EAC6G3A(192)-b
M61B-19-a
2.9
M61B-18-a
4.2
M50B-8-a
2.8
M55A-4-a
M55B-17-a
4.1
M55A-6-a
M56B-3-b
11.4
SpAC1B8B(180)-b
11.6 SpAC1B8A(260)-a
E
EAC7B9(90)-a*
8.7
M49C-11-b
2.7 M49C-10-a
7.3
M48B-6-a*
5.8
M61C-2-a
14.9
M48B-5-a
15.4
M56A-12-a
16.5
M54B-3-b
6.4
M50C-1-a
19.6
M49A-18-b
17.5
M48C-5-b
4.8
S7M52-6-a
3.3 M58A-6-b
5.0
NACG7
(180/190)-a
7.0 M60B-19-a
18.0
M58B-9-b
F
M55A-11-b
15.6
NACB6
(160/190)-b
7.5 M61A-8-a
7.8 M60B-6-a
8.3 M50C-11-a
17.1
7.5 M50C-10-b
7.5 M50A-20-a
M55B10-a
5.2 M60A-4-a
12.9
2.6 M50A-16-a
12.9
M50A-9-a
5.0
M60A-15-a
21.3
EAC6B2
(135/191)-a
M56A-9-a
1.5 M61A-2-b
26.7
M50C-5-b
13.9 M58B-17-a
2.5 M50B-7-a
14.0
M61B-15-b
5.0
S1M52-2-a
12.2
M55B-9-a
G
EAC7F10
(124/136)-a
6.2
M60C-15-a
5.2
EAC6E2D(90)-a
1.7
M60A-10-b
8.5
M50A-13-b
4.2 M48B-4-a
12.9
M48A-1-a
8.6
M56B-4-a
M49C-2-a
M56A-21-a
2.8
M56A-20-a
1.8 M60A-13-a
M49C-1-b
13.7
M57A-3-a**
18.9
M49A-10-b*
2.8 M58B-8-b
13.3
M58B-15-a
14.4
M56B-2-a
18.7
M55B-3-b
25.1 M56C-9-a
H
M60A-7-a
15.0 M54A-9-b
17.2 M61C-7-a
18.9
EAC7F8
(218/238)-a
1.4
M61A-9-a
1.4
M58B-12-a
4.4
M60B-9-a
10.4 M58A-7-a**
21.1
M58B-18-a
22.4
M54A-4-a
M54A-2-b
M54A-3-b
18.0
M50A-8-b
5.7
M50B-11-a
4.6
M50C-9-b
2.0 M58B-14-b
19.8
M60B-3B
10.0 S7M52-1-b
3.6 M61C-10-b
5.1 M56B-1-a

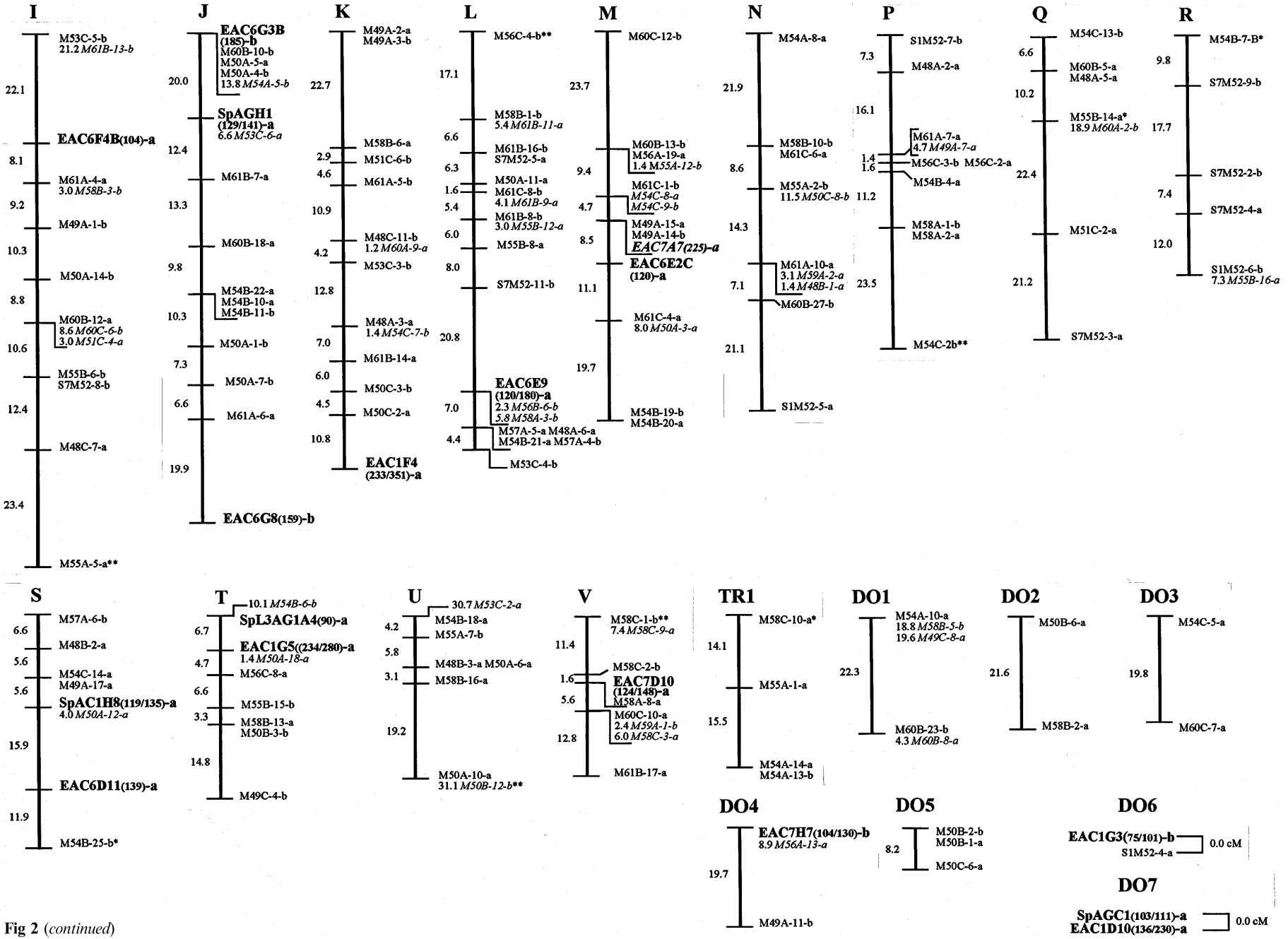


Fig 2 (*continued*)

Table 2 Estimates of *Picea abies* genome length

T	X (cM)[b]	X (cM)[b]	Genome length[a] Estimate 1 (cM)	Genome length[a] Estimate 2 (cM)
3.0	30.7	31.89	2783	2891
4.0	26.2	26.98	2784	2867
5.0	22.6	23.29	2812	2899

[a] Genome length is expressed in map units (cM, Kosambi function) and was calculated (Estimate 1) according to the method described by Hulbert et al. (1987), or (Estimate 2) using method No. 3 of Chakravarti et al. (1991)
[b] X is the map distance between two markers for which the expected value of the LOD score is T, calculated by the methods mentioned above

Table 3 Expected number of markers needed to cover a given fraction of genome length with a maximum distance of 25.54 cM between adjacent markers

Genome coverage[a]	Needed markers
0.80	220
0.85	259
0.90	314
0.95	409

[a] Based on an estimated genomic length of 2839 cM; 25.54 cM is the maximum genetic distance between two markers at a LOD score threshold of 4.3

Table 4 Analysis of clusters of consecutive crossovers and markers

No. of crossovers per interval	Observed		Expected		
	No.	Percentage	No.	Percentage	P (longest run ≥ n) (%)[b]
a Number of crossovers between consecutive random markers[a]					
0	47	18.43	46.66 ± 6.17	18.30	100%
1	15	5.88	38.12 ± 5.69	14.95	100%
2	19	7.45	31.15 ± 5.23	12.21	100%
3	31	12.16	25.45 ± 4.79	9.98	100%
4	22	8.63	20.79 ± 4.37	8.15	100%
5	22	8.63	16.99 ± 3.98	6.66	100%
6	13	5.10	13.88 ± 3.62	5.44	100%
7	15	5.88	11.34 ± 3.29	4.45	100%
8	13	5.10	9.26 ± 2.99	3.63	100%
9	13	5.10	7.57 ± 2.71	2.97	99.9%
10	13	5.10	6.18 ± 2.46	2.42	99.8%
11	9	3.53	5.05 ± 2.23	1.98	99.5%
12	15	5.88	4.13 ± 2.02	1.62	98.6%
13	3	1.18	3.37 ± 1.82	1.32	97.1%
14	3	1.18	2.76 ± 1.65	1.08	94.5%
15	2	0.78	2.25 ± 1.49	0.88	90.7%
Total	255				

b Random markers occurring between consecutive crossovers[c]

No. of markers per block	Observed		Expected		
	No.	Percentage	No.	Percentage	P (longest run ≥ n) (%)[d]
0	930	79.78	953.45 ± 13.21	81.70	100%
1	200	17.05	174.47 ± 12.18	14.95	100%
2	30	2.57	31.93 ± 5.57	2.74	100%
3	3	0.26	5.84 ± 2.41	0.50	99.6%
4	4	0.34	1.07 ± 1.03	0.09	63.2%
Total	1167				

[a] The intervals with ≥1 crossover represent the 208 gaps between consecutive bins of recombinationally inseparable markers
[b] The probability of the longest run is calculated for $p = 0.82$ (see Materials and methods)
[c] The blocks with ≥1 marker represent the 237 bins of recombinationally separable markers
[d] The probability of the longest run is calculated for $p = 0.18$

mapped. Only a small amount of DNA (2.5 ng per locus) is required for SSR genotyping also. Considering that from 1 to 2 μg of DNA are recovered from each megagametophyte in Norway spruce, it will be possible to add new SSR markers (both of di- and tri-nucleotide repeats) to this map as soon as they become available. Special emphasis has been given to the automation and streamlining of the genotyping procedures. The use of Biomek 2000 and Biomek 1000 Workstations for the assembly of PCR reactions and the loading of agarose gels (stained with a dye that is 10 times more luminescent than ethidium bromide; see Materials and methods) greatly enhanced the efficiency and rapidity of the whole mapping project. Bands were obtained with 21 out of the

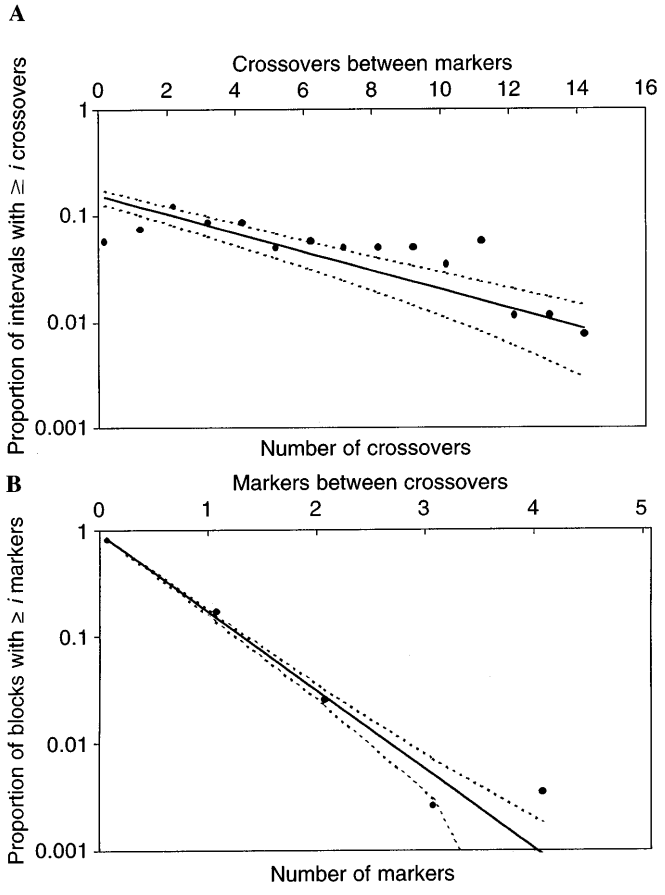


Fig. 3A, B Observed versus expected proportion of markers and crossovers in the Norway spruce linkage map. **A** Observed proportion of genetic intervals in the map having $\geq i$ crossovers, compared to the expected proportion of Pc^i ($Pc = 0.82$). **B** Observed proportion of blocks in the map containing $\geq i$ recombinationally unseparated markers, compared to the expected proportion of Pm^i ($Pm = 0.18$). Data are plotted on a logarithmic scale. The expected data fall on the *solid line*, observed data are plotted as *points*. *Dotted lines* indicate upper and lower confidence limits, corresponding to two standard deviations

total 34 SSR primer pairs used, and mapped using this procedure. This finding is of great importance, as it makes it possible to exploit the whole power of PCR-based technologies for genotyping and constructing second-generation genetic maps (Goodfellow 1992; Hodgson 1994).

Genome characterisation in *P. abies* and other conifers based on the use of genetic mapping data

We tested the distribution of the markers on our map with respect to crossovers in order to see if significant deviations from the assumed random distribution could be observed (Table 4). We observed an excess of large clusters of crossovers that correspond to large genetic intervals; on the other hand the data also suggest that there is a smaller incidence of small clusters of crossovers than expected (Fig. 3a). This is a clear indication that markers may be not distributed randomly through the genome, and that some chromosomal regions could not be appropriately sampled. While SSRs have proven to be randomly distributed in the very dense maps available for some mammals (Dietrich et al. 1996), no detailed studies have been conducted on AFLPs (Barnes et al. 1996). The portions of the genome sampled by this technique strongly depend on the restriction enzymes used. In our case, the choice of a methylation-sensitive endonuclease like *Pst*I as the rare-cutter restriction enzyme was made in order to reduce the complexity of the amplification products. *Pst*I-generated markers are generally derived from hypomethylated sequences that are considered to correspond to the single- or low-copy-number class, and thus may show a distribution corresponding to that of expressed sequences (i.e. genes). Most of the SSR markers employed in this study were also selected for their content of low-copy number DNA in the sequences flanking the repetitive region (Pfeiffer et al. 1997). We therefore biased our marker set towards single-copy or low-copy-number sequences. If large regions are present in the spruce genome that are mainly or wholly comprised of repetitive hypermethylated DNA, and such regions are recombinationally active, these regions will not be adequately sampled using *Pst*I-derived markers, and thus the distribution of markers will not be random. Additional markers of the same type would not help to merge linkage groups and alternative approaches would be needed in order to do this, such as using markers that specifically target repetitive DNA [i.e. Sequence Specific Amplification Polymorphisms (S-SAP; Waugh et al. 1997] or that do not select for non repetitive regions (i.e. RAPDs). An uneven distribution of markers and the presence of large gaps may also be explained by the occurrence of hot spots for recombination.

In the only other map of *P. abies* available, which is entirely composed of RAPDs (Binelli and Bucci 1994), 26 linkage groups were found using 185 marker loci on a panel of 72 megagametophytes, with an average inter-marker distance of 22 cM. The total genetic distance covered by this map was 3584 cM, which is 163% of the 2198 cM genetic distance covered by our map. This dramatic difference can be explained by the difference in the methods used for map construction. In our map those markers which could not be ordered at LOD $\geq$ 2.0 were excluded from the framework and considered as ancillary markers (Fig. 2) because their position could not be confirmed with the same level of statistical significance (i.e. for framework markers the closest alternative order was at least 100 times less probable, rising to up to 1000 times in some portions of the map). The largest genetic interval between two markers was 28.1 cM (close to the theoretical maximum genetic distance of 25.54 cM for two loci linked with a LOD $\geq$ 4.3 given our average population size). In the RAPD map (Binelli and Bucci 1994) some markers were assigned to linkage groups with a LOD $\geq$ 2.0 and the interval-supporting LOD score for the multipoint analysis was in some cases less than 1.0. This may have contributed to the observed inflation of

genetic distances (in several cases two flanking markers were ordered at a genetic distance of 50 cM or more). No estimates of expected genetic distances were made. In our case, following the method-of-moment estimator described by Hulbert et al. (1988) and Chakravarti et al. (1991) we obtained estimates of approximately 2793 and 2886 cM respectively (Table 1). These values are the only ones reported so far for Norway spruce, but other estimates have been made for *Pinus* species, ranging from 1290 cM (K) (*P. pinaster*, Plomion et al. 1995b), to 3000 cM (H) (*P. elliottii*, Nelson et al. 1993) (H = Haldane function, K = Kosambi function). This extreme heterogeneity probably reflects divergence in the criteria used to calculate the estimates rather than having a biological basis due to differences among *Pinus* species. Echt and Nelson (1997) recalculated this parameter, using the linkage data sets derived from maps of different *Pinus* species but the same calculation criterion, and concluded that the expected genetic distance should in fact be of the order of 2000 cM (K) for all of them. The difference in estimated genetic distance between *P. abies* and *Pinus* thus should be around 800 cM (K). It is not clear whether this reflects actual biological differences between these distantly related species or is the result of biases deriving from the use of different classes of markers (and therefore from differences in their genomic distributions).

SSR markers: identification and use in conifers

Efficient strategies for the rapid isolation of hundreds of microsatellite sequences in *P. abies* have been devised (Pfeiffer et al. 1997); nevertheless, the assembly of a set of SSRs suitable for all types of genetic analysis still remains quite a time-consuming and expensive task. Once the set has been established, however, the use and distribution of markers among different labs is easy and inexpensive. In conifers, as well as other species with large genomes, such as wheat (Röder et al. 1995), SSRs show complex patterns of amplification, i.e. the presence of amplified DNA fragments in addition to the ones expected from the specific microsatellite region. This is probably due to the high proportion of repetitive DNA in which the microsatellite sequences can be embedded (Pfeiffer et al. 1997, Paglia, Magni and Morgante, in preparation). Although some countermeasures to minimise this problem can be taken, Pfeiffer et al. (1997) reported that in Norway spruce only 20% of the primer pairs designed for SSRs amplified a single variable locus, which is a prerequisite for the use of these markers for diversity analysis. However, the remaining primers can be successfully employed for mapping. In our genetic linkage map of *P. abies*, we were able to include both single-locus SSRs and those which gave more complex patterns, provided that a clear polymorphism and a Mendelian segregation ratio of 1:1 was observed, which was the case for all the 61 bands scored. Thus, 38 SSR marker loci were mapped to 19 linkage groups, representing the first significant set of STSs included so far in a genetic map of conifers.

SSR markers: utility in population genetics and breeding of forest trees

Mapped SSRs at different levels of linkage could be used for studying linkage disequilibrium in *P. abies* (Epperson and Allard 1987). This would be of importance not only for monitoring the effects of population subdivision, genetic drift and epistatic selection in natural populations of Norway spruce, but also in assessing the utility of such markers for marker-assisted selection. Indeed, since forest trees generally exhibit high levels of genetic diversity and are highly outcrossed, it has been postulated that linkage disequilibrium should be low and thus that alleles at marker loci and alleles at loci of interest (i.e. QTLs, Quantitative Trait Loci) should be randomly associated in different genotypes (Plomion et al. 1995a). This would require that a genetic map be created for individual trees (Grattapaglia and Sederoff 1994), preferably using high- throughput PCR fingerprinting techniques, such as RAPDs or AFLPs. These marker types, although suitable for the quick genotyping of many individuals and for rapid map construction, have the severe drawback of being difficult to transfer between different maps because they are not STS markers (Powell et al. 1996). The availability of a set of evenly distributed multiallelic and fully informative mapped microsatellites will help to create index maps with a subset of common markers bracketing regions in which QTLs have been identified, and will thus facilitate their manipulation in breeding. Mapped genetic markers can also be employed to test for homogeneity of recombination fraction among different individuals and different sexes (Plomion et al. 1995b).

The analysis of the mating system in the outcrossed natural populations of conifers is not only important for understanding population genetic structure but also has practical significance for reforestation with wind-pollinated seed, due to the fact that most coniferous species, which are known to be predominantly outbreeders, usually suffer from strong inbreeding depression (Franklin 1970, Ledig et al. 1983). In this case the availability of sets of several unlinked multiallelic codominant microsatellite markers will be a valuable tool that allows one to perform paternity analysis and to estimate parameters of gene flow, mating system, genetic relatedness of progeny, male reproductive value, mate choice, seed shadow. Such studies have been hampered so far by the paucity of polymorphic loci and by the low number of alleles per locus (Ritland 1983). The 38 SSRs which have been mapped in 19 linkage groups on the present map can provide a tool for all of these applications and represent an important step towards the utilisation of this class of markers in gymnosperms.

Acknowledgements The authors are grateful to Craig Echt for providing information prior to publication and to Christophe Plomion for helpful comments on the manuscript. The excellent technical assistance of Monica Vizzi and Michela Pin is gratefully acknowledged. This study was supported by grants from the European Commission Biotechnology Program (BIO3-CT93-0373, BIO4-CT96-1976)

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphism of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Barnes S, Massaro G, Lefebvre M, Kuiper M, Verstege E (1996) A combined RFLP and AFLP genetic map for sugar beet. Proceedings of the 59th IIRB Congress, pp 555–560

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

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. Curr Opin Genet Dev 3:939–943

Chakravarti A, Lasher LA, Reefer JE (1991) A maximum likelihood method for estimating genome length using genetic linkage data. Genetics 128:175–182

Conkle MT (1981) Isozyme variation and linkage in six conifer species. Proceedings of the Symposium on Isozymes in North American Forest Trees and Forest Insects. Pacific Southwest Experiment Station, USDA Genetics Technical Report PSW-48, pp 11–17

Devey ME, Fiddler TA, Liu BH, Knapp SJ, Neale DB (1994) An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. Theor Appl Genet 88:273–278

Devey ME, Bell JC, Smith DN, Neale DB, Moran GF (1996) A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers. Theor Appl Genet 92:673–679

Dib C, Fauré S, Fizames C, Samson D, Drout N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5264 microsatellites. Nature 380:152–154

Dietrich WF, Miller JC, Steen RG, Merchant M, Damron D, Nahf R, Gross A, Joyce DC, Wessel M, Dredge RD, Marquis A, Stein LD, Goodman N, Page DC, Lander ES (1994) A genetic map of the mouse with 4006 simple sequence length polymorphisms. Nature Genet 7:220–225

Dietrich WF, Miller J, Steen R, Merchant MA, Damron-Boles D, Husain Z, Dredge R, Daly MJ, Ingalls KA, O'Connor TJ, Evans CA, DeAngelis MM, Levinson DM, Kruglyak L, Goodman N, Copeland NG, Jenkins NA, Hawkins TL, Stein L, Page DC, Lander ES (1996) A comprehensive genetic map of the mouse genome. Nature 380:149–152

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Echt CS, Nelson CD (1997) Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). Theor Appl Genet 94:1031–1037

Epperson BK, Allard RW (1987) Linkage disequilibrium between allozymes in natural populations of lodgepole pine. Genetics 115:341–352

Franklin EC (1970) Survey of mutant forms and inbreeding depression in species of the family Pinaceae. USDA Forest Service Research Paper SE-61. Southeastern Forest Experiment Station, Asheville, N.C.

Goodfellow PN (1992) Variation is now the theme. Nature 359:777–778

Goodfellow PN (1993) Microsatellites and the new genetic maps. Curr Biol 3:149–150

Govindaraju RD, Cullis CA (1991) Modulation of genome size in plants: the influence of breeding system and neighbourhood size. Trends Evol 5:43–51

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

Hodgson J (1994) Genome mapping the "easy way". Bio/Technology 12:581–584

Hulbert SH, Ilott TW, Legg EJ, Lincoln SE, Lander ES, Michelmore RW (1988) Genetic analysis of the fungus *Bremia lactucae* using restriction fragment length polymorphism. Genetics 120:947–958

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, van de Wiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol Breeding 3:381–390

Karp A, Edwards KJ, Bruford M, Funk S, Vosman B, Morgante M, Seberg O, Kremer A, Boursot P, Arctander P, Tautz D, Hewitt GM (1997) Molecular technologies for biodiversity evaluation: opportunities and challenges. Nature Biotechnol 15: 625–628

Kesseli RV, Paran I, Michelmore RW (1994) Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. Genetics 136:1435–1446

Kostia S, Varvio SL, Vakkari P, Pulkkinen P (1995) Microsatellite sequences in a conifer, *Pinus sylvestris*. Genome 38:1244–1248

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary linkage maps of experimental and natural population. Genomics 1:174–181

Lange K, Boehnke M (1982) How many polymorphic marker genes will it take to span the human genome? Amer J Hum Genet 34: 842–845

Ledig FT, Guries RP, Bonefeld BA (1983) The relation of growth to heterozygosity in forest trees. Evolution 37:1227–1238

Miksche JP, Hotta Y (1973) DNA base composition and repetitious DNA in several conifers. Chromosoma 41:29–36

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Morgante M, Vogel J (1994) Compound microsatellite primers for the detection of genetic polymorphisms. US Patent Application Number 08/326456

Morgante M, Rafalski A, Biddle P, Tingey S, Olivieri AM (1994) Genetic mapping and variability of seven soybean simple sequence repeat loci. Genome 37:763–769

Nelson CD, Nance WL, Doudrick RL (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

Nelson CD, Kubisiak TL, Stine M, Nance WL (1994) A genetic linkage map of longleaf pine (*Pinus palustris* Mill.) based on random amplified polymorphic DNAs. J Hered 85:433–439

Olson M, Hood L, Cantor C, Botstein D (1989) A common language for physical mapping of the human genome. Science 245:1434–1435

Paglia GP, Morgante M (1998) PCR-based multiplex DNA fingerprinting techniques for the analysis of conifer genomes. Mol Breeding, in press

Pfeiffer A, Olivieri AM, Morgante M (1997) Identification and characterisation of microsatellites in Norway spruce (*Picea abies* K). Genome 40:411–419

Plomion C, Bahrman N, Durel CE, O'Malley DM (1995a) Genomic mapping in *Pinus pinaster* (maritime pine) using RAPD and protein markers. Heredity 74:661–668

Plomion C, O'Malley DM, Durel CE (1995b) Genomic analysis in maritime pine (*Pinus pinaster*). Comparison of two RAPD maps using selfed and open pollinated seeds of the same individual. Theor Appl Genet 90:1028–1034

Plomion C, Durel CE, O'Malley DM (1996) Genetic dissection of height in maritime pine seedlings raised under accelerated growth conditions. Theor Appl Genet 93:849–858

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalsky A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breeding 2:225–238

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet 9:275–280

Rake AV, Miksche JP, Hall RB, Hansen KM (1980) DNA reassociation kinetics of four conifers. Can J Genet Cytol 22: 69–79

Ritland K (1983) Estimation of mating system. In: Tanksley SD, Orton TJ (eds). Isozymes in plant genetics and breeding, part A. Elsevier Science Publishers, Amsterdam, pp 289–302

Röder MS, Plaschke J, König SU, Börner A, Sorrels ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246:327–333

Rohrer GA, Alexander LJ, Keel JW, Smith TP, Beattie CW (1994) A microsatellite linkage map of the porcine genome. Genetics 136:231–245

Senior NL, Heun M (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. Genome 36:884–889

Smith D, Devey ME (1994) Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37:977–983

Taramino G, Tingey S (1996) Simple Sequence Repeats for germplasm analysis and mapping in maize. Genome 39:277–287

Tulseriam LK, Glaubitz JC, Kiss G, Carlson J (1992) Single tree genetic linkage mapping using haploid DNA from megagametophytes. Bio/Technology 10:686–690

Van der Ven WTG, McNicol RJ (1996) Microsatellites as DNA markers in Sitka spruce. Theor Appl Genet 93:613–617

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 21:4407–4414

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BB, Powell W (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol Gen Genet 253:687–694

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. Nature 359:794–801

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241:225–235

Zhao X, Kochert G (1993) Phylogenetic distribution and genetic mapping of a (GGC)n microsatellite from rice (*Oryza sativa* L.). Plant Mol Biol 21:607–614