Theor Appl Genet (2002) 104:987–997
DOI 10.1007/s00122-001-0823-y

I. Gosselin · Y. Zhou · J. Bousquet · N. Isabel

Megagametophyte-derived linkage maps of white spruce (*Picea glauca*) based on RAPD, SCAR and ESTP markers

Received: 26 March 2001 / Accepted: 25 September 2001 / Published online: 13 February 2002

Abstract We have constructed linkage maps for two parents of white spruce [*Picea glauca* (Moench) Voss]. Haploid megagametophytes from 92 and 96 seeds of parents M2 and 80132, respectively, were analysed with RAPD, SCAR and ESTP markers. Fragments segregating in a 1:1 Mendelian ratio were classified and mapped using MAPMAKER, GMENDEL and JOINMAP. For M2, the analysis with JOINMAP resulted in 165 loci (152 RAPDs, 3 SCARs and 10 ESTPs) mapping to 23 linkage groups and covering 2,059.4 cM(Kosambi function, K). For 80132, the analysis resulted in 144 loci (137 RAPDs, 1 SCAR and 7 ESTPs) mapping to 19 linkage groups and covering 2,007.7 cM(K). The maps covered 87 and 73% of the entire genome of parents M2 and 80132, respectively. Similar results were obtained with MAPMAKER and GMENDEL. A comparison was made between the two individual maps and 16 loci were shared between the two maps.

Keywords Codominant markers · Conifers · Genetic linkage map · GMENDEL · JOINMAP · MAPMAKER

Introduction

White spruce [*Picea glauca* (Moench) Voss] is widely distributed in Canada, from the Atlantic to the Pacific ocean. It is used for lumber, as well as pulp and paper production, and is the second conifer species of importance for reforestation in Canada. White spruce is highly diverse genetically for quantitative characters and marker loci (Furnier et al. 1991; Li et al. 1993, 1997; Jaramillo-Correa et al. 2001). Besides breeding for improved growth, the high heritability of various traits such as outer wood density (Corriveau et al. 1991) and embryogenic capacity (Y.S. Park, personal communication) indicates that significant gains could be realized from selection and breeding in this species.

In white spruce, individual linkage maps are needed in order to estimate a consensus conspecific map and identify genomic regions that contain genes affecting wood density and embryogenic capacity, for possible deployment of marker-aided selection programs. At a more fundamental level, these maps will contribute to our understanding of genome organization and the building of a congeneric consensus map. Linkage maps have been constructed in numerous conifer species using restriction fragment length polymorphisms (RFLPs) (e.g., Devey et al. 1996; Jermstad et al. 1998; Sewell et al. 1999), random amplified polymorphic DNA markers (RAPDs) (e.g., Tulsieram et al. 1992; Binelli et Bucci 1994; Arcade et al. 2000), and amplified fragment length polymorphisms (AFLPs) (e.g., Paglia et al. 1998; Travis et al. 1998). In conifers, map construction is complicated due to their large genome size, late sexual maturity, long generation time, inbreeding depression and by the lack of suitable pedigrees. To alleviate these shortcomings, an alternative approach to mapping in conifers consists of using haploid megagametophytes of a single tree as a mapping population, with the limitation that only PCR-based markers can be used because of the small amounts of DNA in megagametophytes (Tulsieram et al. 1992). For such a demonstration, a preliminary map was constructed in white spruce, with 47 RAPD loci distributed among 12 linkage groups and covering a distance of 873.8 cM (Tulsieram et al. 1992).

Each type of molecular marker has advantages and limitations, and many factors can influence the choice of markers for a given purpose. RAPD markers appear to be well suited for developing single-tree linkage maps due to their rapidity and simplicity of detection. No prior knowl-

Communicated by H.C. Becker

I. Gosselin · Y. Zhou · J. Bousquet · N. Isabel (✉)
Centre de recherche en biologie forestière
and Chaire de recherche du Canada en génomique forestière
et environnementale, Pavillon Charles-Eugène-Marchand,
Université Laval, Ste-Foy Québec, Canada G1K 7P4
e-mail: nisabel@exchange.cfl.forestry.ca

I. Gosselin · N. Isabel
Service canadien des forêts, Ressources Naturelles Canada,
Centre de foresterie des Laurentides, 1055 du P.E.P.S.,
P.O. Box 3 800, Ste-Foy, Québec, Canada G1V 4C7

edge of DNA sequences is required and it is easy to generate a large number of markers due to the wide availability of synthetic oligonucleotides. The dominance property of RAPDs may be overcome by the use of haploid megagametophytes (Isabel et al. 1995) or appropriate pedigrees and configurations such as the pseudotestcross (Grattapaglia and Sederoff 1994). However, RAPDs are often criticized for their lack of reproductibility over time as well as between laboratories (Jones et al. 1997). The transferability/reproducibility of RAPD markers in different backgrounds (individuals or species) may be limited to only very intensely amplified polymorphisms (Kubisiak et al. 1996). Moreover, the amplification products from different backgrounds may not represent the same locus, so it becomes difficult to compare or combine linkage groups (Torres et al. 1993). On the other hand, RAPD markers appear to be repeatable in terms of their amplification and inheritance pattern using various haploid megagametophyte DNA arrays obtained from the same individual (Plomion et al. 1995; Kubisiak et al. 1996). RAPDs are increasingly replaced by AFLPs since the number of putative gene loci revealed per PCR reaction is greater (Vos et al. 1995). Otherwise, the same limitations reported above for RAPDs generally apply to AFLPs.

Sequence-characterized amplified regions (SCARs), which are usually derived from sequenced RAPD loci, are most often dominant but are derived from specific primers, which should make them transferable and reproducible between conifer species (Scotti et al. 1998). The development of sequence-tagged-site (STS) markers and microsatellites is long and expensive, but their use is easy and informative because they are codominant and derived from specific primers (Perry and Bousquet 1998a; Lefort et al. 1999). Also, they can be indicative of coding regions (Perry and Bousquet 1998a), and thus represent expressed sequence tag polymorphisms (ESTPs). These genetic markers are easily transferable from one species to another in the genus *Picea* (Perry and Bousquet 1998b; Perry et al. 1999) and, accordingly, they can be used for the construction of intraspecific and interspecific consensus maps.

Various softwares are available to construct linkage maps and each has its own criteria and merit. A survey of the literature reveals that MAPMAKER appears to be the most frequently used software. A multi-point linkage analysis, which considers all of the raw (genotypic) data available simultaneously in each computation, is used by MAPMAKER to find map orders and map distances (Lincoln et al. 1993). GMENDEL uses simulated annealing and multiple pairwise methods for locus ordering (Holloway and Knapp 1993), which is different from the locus ordering method used by MAPMAKER. But most often, GMENDEL results in maps equivalent to those found by MAPMAKER (Holloway and Knapp 1993). JOINMAP is useful for combining linkage data that have been collected in differents experiments. Unlike the preceeding softwares, JOINMAP uses the raw segregation data of various types, especially haploid, as well as "independent" estimates of recombination to construct integrated linkage maps (Stam 1993). The map orders are found by calculating pairwise recombination frequencies, and map distances are estimated by a least-squares procedure.

The goals of this study were to: (1) construct two individual linkage maps of white spruce with RAPD, SCAR and ESTP loci, and estimate the genome length of white spruce and the number of loci necessary to achieve map saturation; (2) find common loci between the two maps; and (3) compare the results obtained with MAPMAKER V3.0, JOINMAP V2.0 and GMENDEL V2.0.

Materials and methods

Plant materials and DNA extraction

The two mapping populations consisted of 92 megagametophytes and 96 megagametophytes derived from the two white spruce parents M2 and 80132, respectively. These trees are native to the Ottawa Valley in eastern Canada, and they are part of partial diallels on wood density and embryogenic capacity, respectively.

Seeds were surface-sterilized with sodium hypochlorite. The seed coat and embryo were removed and each megagametophyte was ground in a 1.5-ml Eppendorf tube, using a grinding (Caframo), with 50 µl of extraction buffer (2% CTAB, 100 mM Tris, 1.4 M NaCl, and 20 mM EDTA, pH 9.5). Then, DNA was extracted using the procedure of Isabel et al. (1995).

RAPD assay

All RAPD primers used in this study were random sequence, 10-base, oligonucleotide primers with C+G contents ranging from 50 to 80%. One set of primers was obtained from Operon Technologies Inc. (Alameda, Calif.) and an additional set from the University of British Columbia (John Carlson, Biotechnology Laboratory; now at the School of Forest Resources, Pennsylvania State University). The RAPD locus names denote the supplier (letters = Operon, only number = UBC), the primer number and the approximate marker size in base pairs.

RAPD reactions were based on the protocol of Isabel et al. (1993) with some modifications. Amplification reactions were performed in volumes of 13 µl containing 1 × Boehringer Mannheim reaction buffer, 200 µM of each dNTP (Pharmacia), 25 (or 50) µM of primer, 1 ng of DNA, and 0.5 units of *Taq* DNA polymerase (Boehringer Mannheim). For DNA amplification, a Perkin Elmer Gene Amp PCR System 9600 was programmed for 45 cycles, each consisting of a denaturating step of 15 s at 94 °C, followed by an annealing step of 15 s at 35 °C, and an extension step of 1 min 30 s at 72 °C. The last 25 extension steps were progressively extended by 5 s per cycle, and the last cycle was followed by 10 min at 72 °C. After amplification was completed, 10 µl of each sample were loaded and electrophoresed on 0.5% Synergel™ and a 1% agarose gel during 4 h at 190 V, and visualized by ethidium bromide staining. Gels were photographed under UV light with an Imager System (Canberra Packard, Ont.). The molecular-weight marker used was a 100-bp ladder (Pharmacia).

SCAR assay

SCAR markers, derived from RAPD markers, were previously developed for genetic mapping in *Picea abies* (Scotti et al. 1998). In this study, nine primers pairs were used, which resulted in positive amplification in white spruce (Scotti et al. 1998). SCAR reactions were based on the protocol of Paran and Michelmore (1993) with some modifications. Amplifications were performed in a volume of 15 µl containing 1 × reaction buffer (Boehringer Mannheim), 150 µM of each dNTP (Pharmacia), 0.2 µM of primer, 2 ng of

DNA, and 1 unit of *Taq* DNA polymerase (Boehringer Mannheim). The same PCR cycler as above was programmed for 30 cycles, each consisting of a denaturating step of 1 min at 94 °C, followed by an annealing step of 1 min at 55, 58, 62 or 67 °C [depending on the marker tested, Scotti et al. (1998)], and an extension step of 2 min at 72 °C. Electrophoresis and detection conditions were the same as in the RAPD assay. In all cases, PCR reactions were performed at least twice in order to ensure that absence was a real absence and not a failed reaction.

EST and EST-SSCP assays

Thirty six EST primers developed for *Picea mariana* (Perry and Bousquet 1998a) and further assessed in white spruce (Perry and Bousquet 1998b) were tested. These cDNA-based markers may be useful for establishing anchor loci in genomic mapping. The amplification procedure was followed as described by Perry and Bousquet (1998a). Electrophoresis conditions were as described by Fournier et al. (2001). When PCR products remained monomorphic on an agarose gel, a more-sensitive method of detection, SSCP (single-strand conformation polymorphism), was used according to Fournier et al. (2001). Electrophoresis was run in 0.6 × TBE buffer at 4, 10 or 14 °C, depending on the locus tested during 18 h at 5 mA. The gel was then silver-stained according to the Bio-Rad manufacturer's protocol (derived from the protocol of Merril et al. 1981). The gels were fixed in 40% methanol for 30 min, and two times with 10% ethanol for 15 min, then with oxidizing for 5 min. The gels were washed three times with deionized water for 5 min. Color impregnation lasted for 20 min with silver stain (Bio-Rad). The gels were then rinsed for 1 min with deionized water and developed three times with the silver stain developer (Bio-Rad) for 30 s, 5 min and 5 min. The staining reaction was stopped with 5% acetic acid. The gels were dried during 2 h at 60 °C.

Primer screening, segregation analysis and marker scoring

All primers were initially screened against a panel of eight megagametophytes for each parent. For parent 80132, 809 RAPD primers were tested as well as 9 SCAR and 36 EST primer pairs. Of these, 456 RAPD primers resulted in poor or null amplification and were simply discarded, resulting in 353 operational RAPD primers. For parent M2, the 353 operational RAPD primers as well as the 9 SCAR and 36 EST primer pairs were also tested. All polymorphisms with an χ^2 -value smaller than 3.84 ($P > 0.05$) were classified as Mendelian genetic markers and used for further linkage analysis, which involved 92 and 96 megagametophytes for parents M2 and 80132, respectively. The RAPD and SCAR markers were scored as A for absent, B for present, and – for missing (i.e., failed or unscorable). Only clear and reproducible fragments were kept and markers with more than 5% missing data were left out of further analyses. The ESTPs and ESTPs-SSCPs were scored, respectively, according to Perry and Bousquet (1998a,b) and to Fournier et al. (2001), from the bottom to the top of the gels.

Linkage analysis

Linkage analysis was performed using: (1) MAPMAKER/EXP V3.0 (Lander et al. 1987), (2) JOINMAP V2.0 (Stam and van Ooijen 1995), and (3) GMENDEL V3.0 (Holloway and Knapp 1993). The Kosambi (1944) function was used to calculate genetic distances between linked markers.

MAPMAKER V3.0

MAPMAKER does not recognize linkages for markers in repulsion, so the repulsion-phase linkages can be detected by analysing re-coded data (i.e., presence re-coded to absence, and vice versa) together with the original data set using the F_2 backcross data type. Preliminary grouping of marker loci was conducted using the "GROUP" command (two-point analysis), with a LOD score ≥ 4.0 and a recombination fraction $\theta \leq 0.30$ (equivalent to 37 cM) (Kosambi function, K). Because of the double-up data set, all identified linkage groups were also doubled up in identical pairs in complementary linkage phases. Arbitrarily, the first of these groups is reported. The same criteria (LOD and θ), a minimal number of missing data (5%), and error detection off, were used to find the orders of each group. The "COMPARE" command was used for groups with less than eight loci. For larger linkage groups, the three-point command was used to pre-compute the likelihood of all three-point crosses of each groups. Then, the "ORDER" command was used to determine a linear order using multi-point analysis. The remaining loci in each group were then placed with the "TRY" command and were then considered as accessory loci.

JOINMAP V2.0

The original data set (haploid population type) was analysed by the JOINMAP software. Preliminary grouping of loci was performed using the "JMGRP" command, with a LOD score ≥ 4.0. Then, loci within the groups were analysed for pairwise linkages using the "JMREC" command, employing REC and LOD thresholds of 0.499 and 0.01, respectively (Jermstad et al. 1998). The linkage groups were then ordered with the "JMMAP" command using the following parameters: 0.1 LOD threshold, 0.49 REC threshold, 3 jump threshold, 7 triplet threshold, and 3 ripple value.

GMENDEL V3.0

The analysis with GMENDEL used the same double-up data set as MAPMAKER. Preliminary grouping was performed using the "COMPUTE GROUPS" command with a minlod of 4.0 and a maximum recombination value (rmax) of 0.30. Simulating annealing and multiple pairwise methods were used for locus ordering. Monte Carlo simulations and bootstrap re-samplings with 500 and 250 iterations were applied to test the ordering and obtain confidence intervals for the mapped positions of marker loci.

Genome length and map coverage estimations

Genome lengths, G, were estimated using the method E(G) = 2MX/K (Hulbert et al. 1988), where M is the number of informative meioses, X is the maximum observed map distance among the locus pairs above a threshold lod, value K is the number of locus pairs having LOD scores at or above Z. The values used for Z were 4.0. Since all meioses were informative in the haploid mapping population, M = n(n-1)/2, where n is the number of loci analysed. The values of X and K were obtained from an inspection of the list of values generated by the "LOD TABLE" function of MAPMAKER and by the "JMREC" function of JOINMAP. The confidence interval for G, I_α (G), was calculated from the equation, $I_\alpha (G) = E(G) (1 \pm n_\alpha K^{-\frac{1}{2}})^{-1}$, where $n_\alpha = 1.96$ for an α of 5% (Gerber and Rodolphe 1994). Observed genome map coverage, C_O, was calculated from the equation $C_O = G_F + X(L–R)$ (Bishop et al. 1983), where G_F is the total cM length of the framework map, L is the total number of linkage groups, pairs and unliked loci, and R is the number of chromosomes. This method for calculating map coverage takes into account the maximum distance, X, of detectable linkages to unliked loci and the ends of linkage groups, as well as the reduction in coverage, $\frac{1}{2}$ X, expected at the ends of chromosomes (Nelson et al. 1994). The minimal number, *n*, of randomly distributed marker loci required to cover a proportion $P = 95\%$ of a genome size of K = E(G), at a maximum distance between marker loci of c = 20 cM, was calculated from the equation $n = [\log (1–P)/\log (1–2c/K)]\ 1.25$ (Lange and Boehnke 1982).

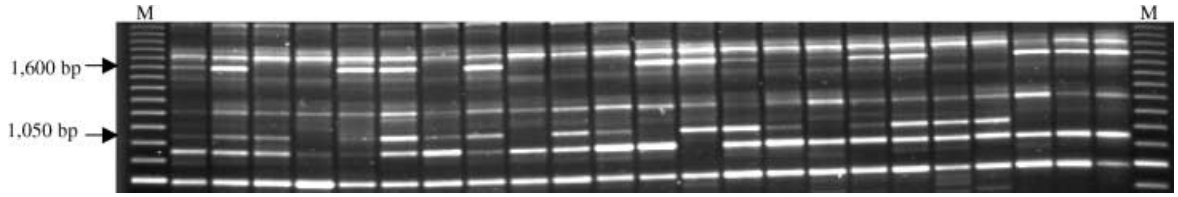

Fig. 1 Segregation of RAPD markers revealed by primer UBC-408 among an array of 23 megagametophytes of parent M2. M is the 100-pb size ladder standard (Pharmacia)

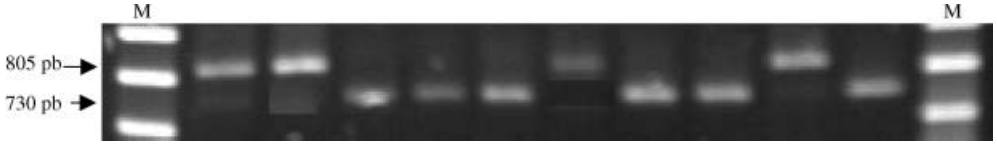

Fig. 2 Segregation of ESTP markers revealed by primer SB24 among an array of 10 megagametophytes of parent 80132. M is the 100-pb size ladder standard (Gibco BRL)

Table 1 Allelic variation observed at ESTP loci in two *P. glauca* individuals

Loci	Segregation classes (A/B)	Observed ratio (A/B)	Expected ratio (A/B)	χ^2 test[a]
Parent M2[b]				
SB01	1,880/1,920	42/36	39/39	0.462 ns[c]
SB16	1,050/1,100	45/45	45/45	0.000 ns
SB24	730/730–805	38/46	42/42	0.76 ns
SB29	575/590	48/40	44/44	0.727 ns
SB35	450/450–475	46/43	44.5/44.5	0.100 ns
SB08 (14 °C)[d]	13/23	45/46	45.5/45.5	0.01 ns
SB14 (4 °C)[d]	12/23	38/54	46/46	2.783 ns
SB21 (14 °C)[d]	13/24	49/42	45.5/45.5	0.538 ns
SB49 (4 °C)[d]	12/13	48/44	46/46	0.17 ns
SB60 (14 °C)[d]	12/13	46/45	45.5/45.5	0.011 ns
Parent 80132[e]				
SB16	1,050/1,100	47/49	48/48	0.04 ns
SB17	640/650	36/58	47/47	5.15 s[f]
SB24	730/805	44/48	46/46	0.17 ns
SB29	575/590	41/52	46.5/46.5	1.3 ns
SB08 (14 °C)[d]	13/23	55/41	48/48	2.04 ns
SB11 (10 °C)[d]	13/24	42/52	47/47	1.06 ns
SB21 (14 °C)[d]	13/24	45/50	47.5/47.5	0.26 ns
SB32 (10 °C)[d]	13/23	53/42	47.5/47.5	1.27 ns
SB49 (4 °C)[d]	12/13	44/51	47.5/47.5	0.52 ns

[a] Critical value of χ^2 at $P = 0.05$ is 3.84 (1 *df*)
[b] Mendelian segregation was tested with an array of 92 megagametophytes
[c] Non-significant
[d] ESTP-SSCP locus and migration temperature
[e] Mendelian segregation was tested with an array of 96 megagametophytes
[f] Significant at $P < 0.05$

Results

Marker screening

For the parent M2, among 353 RAPD primers resulting in positive amplification as well as 9 SCAR and 36 EST primer pairs, 93 RAPD primers (211 markers) as well as 3 SCAR and 10 EST (including 5 EST-SSCP) primer pairs resulted in polymorphic markers. The rate of polymorphic markers per primer or primer pair was 26%, 33% and 28% for RAPDs, SCARs, and ESTs, respectively. From these, 182 showed a 1:1 segregation ($P < 0.05$). Forty one (19%) of the RAPD markers did not show a Mendelian segregation ($P < 0.05$). Figures 1 and 2 show examples of segregation of a RAPD and an ESTP. For the parent 80132 and for the same primers or primer pairs, 80 RAPD primers (180 markers) as well as 1 SCAR and 9 EST (including 5 EST-SSCP) primer pairs resulted in polymorphic markers. The rate of polymorphic markers per primer or primer pair was 23%, 11% and 25% for RAPDs, SCARs, and ESTs, respectively. From these, 166 showed a 1:1 segregation ($P < 0.05$). Twenty three (13%) of the RAPDs and one ESTP (*SB17*) did not show a Mendelian segregation ($P < 0.05$) (Table 1). However, the ESTP *SB17* did segregate in a 1:1 ratio at $P < 0.01$. Distorted markers ($P < 0.05$) were not mapped in this study, except ESTP marker *SB17* for parent 80132 (see above). In total, 182 and 167 polymorphic markers were used for linkage analysis with the parents M2 and 80132, respectively. Among these, there was 13 RAPD, 1 SCAR and 6 ESTP (including 3 ESTP-SSCP) markers being simultaneously polymorphic for both parents, but only 16 were shared between the two maps (see below).

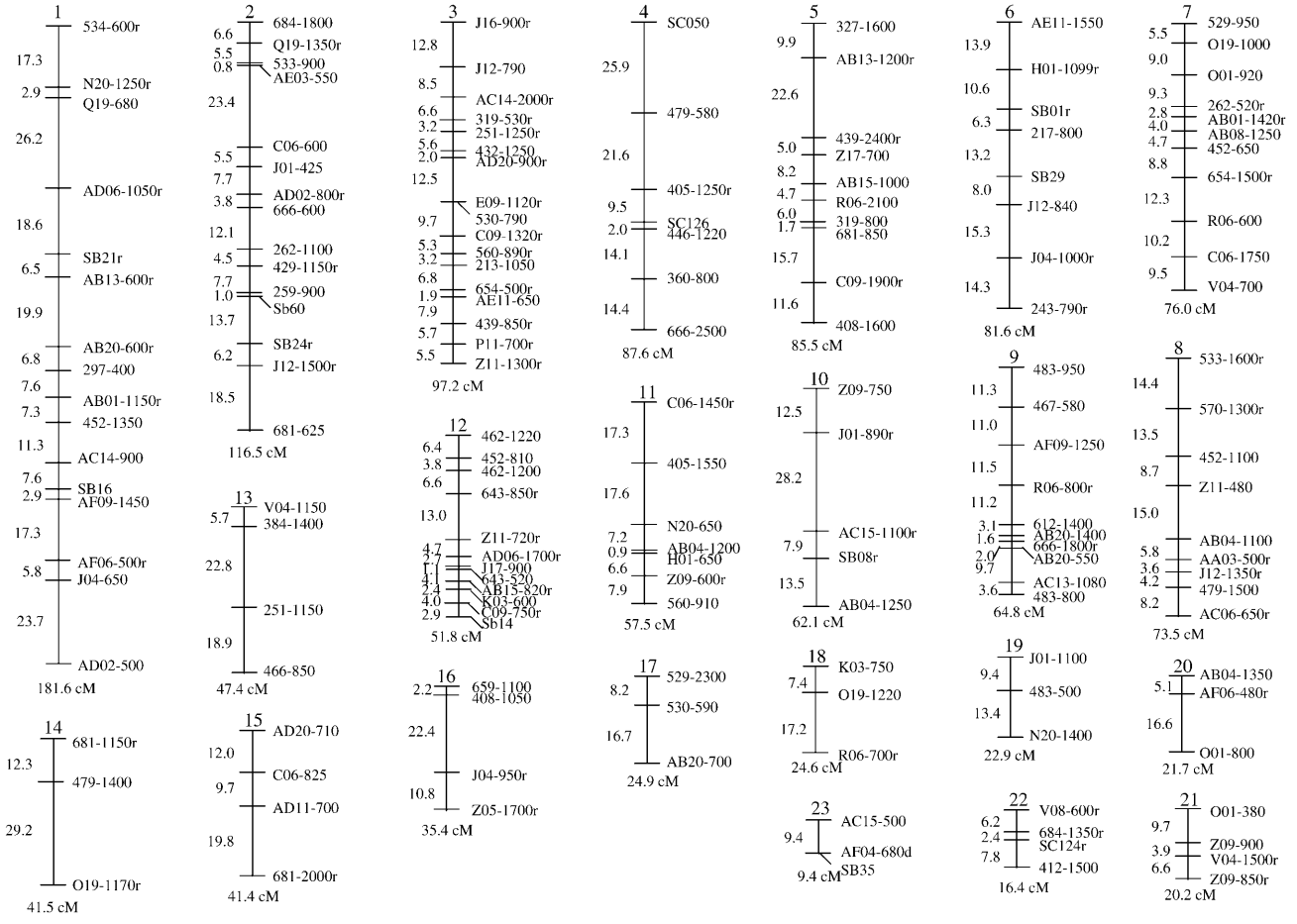

Fig. 3 Linkage map of parent M2 obtained with JOINMAP. Names of framework markers are indicated on the right of the linkage groups (RAPD named according to the text, SB=ESTP markers, and SC=SCAR markers). Those annotated with the letter r indicate markers in repulsion phase. Genetic distances in cM using the Kosambi's mapping function are indicated on the left of the linkage groups. The length of each linkage group (in cM) is shown below the groups

Table 2 Description of the map data

MAP feature	Parent M2			Parent 80132		
	MM[a]	JM[b]	GM[c]	MM	JM	GM
Number of mapped loci	162	165	163	142	145	145
Number of unlinked loci	12	11	11	16	14	14
Number of linkage groups[d]	23	23	23	20	19	19
Number of pairs	3	3	3	4	4	4
Number of accessory loci	32	–	–	14	–	–

[a] MAPMAKER; [b] JOINMAP; [c] GMENDEL
[d] Number of linkage groups of three and more loci, assuming a minimum LOD score ≥ 4.0 and a frequency of recombination θ ≤ 0.30

Construction of the M2 and 80132 maps

M2 maps

The map generated by MAPMAKER, using a LOD value ≥ 4 and θ ≤ 0.30, had a total of 162 loci (with 32 accessory loci) distributed in 23 linkage groups including six groups of three loci, while 12 loci remained unlinked (including one ESTP and one SCAR) (Table 2). The size of the linkage groups ranged from 9.1 cM to 176.8 cM. The average distance between two framework loci was 15.5 cM, with 17 gaps exceeding 25 cM. Nine intervals between two loci were larger than 30 cM. Using the same criteria, the maps constructed by JOINMAP (Fig. 3) and GMENDEL had a total of 165 loci distributed in 23 linkage groups, including six groups of three loci, while 11 loci remained unlinked. With JOINMAP, the size of linkage groups ranged from 9.4 cM to 181.6 cM. The average framework locus spacing was 9.5 cM, with ten gaps exceeding 20 cM. Only four intervals between two loci were larger than 25 cM. The unlinked loci were the same between the three softwares, except for one RAPD locus (*AB04-1200*). The ESTP loci were randomly distributed over the genome, with the nine ESTP loci being distributed on six different linkage groups. The two SCAR loci were on two different linkage groups.

80312 maps

The map generated by MAPMAKER (LOD ≥ 4, θ≤ 0.30) had a total of 142 loci (with 14 accessory loci) distributed

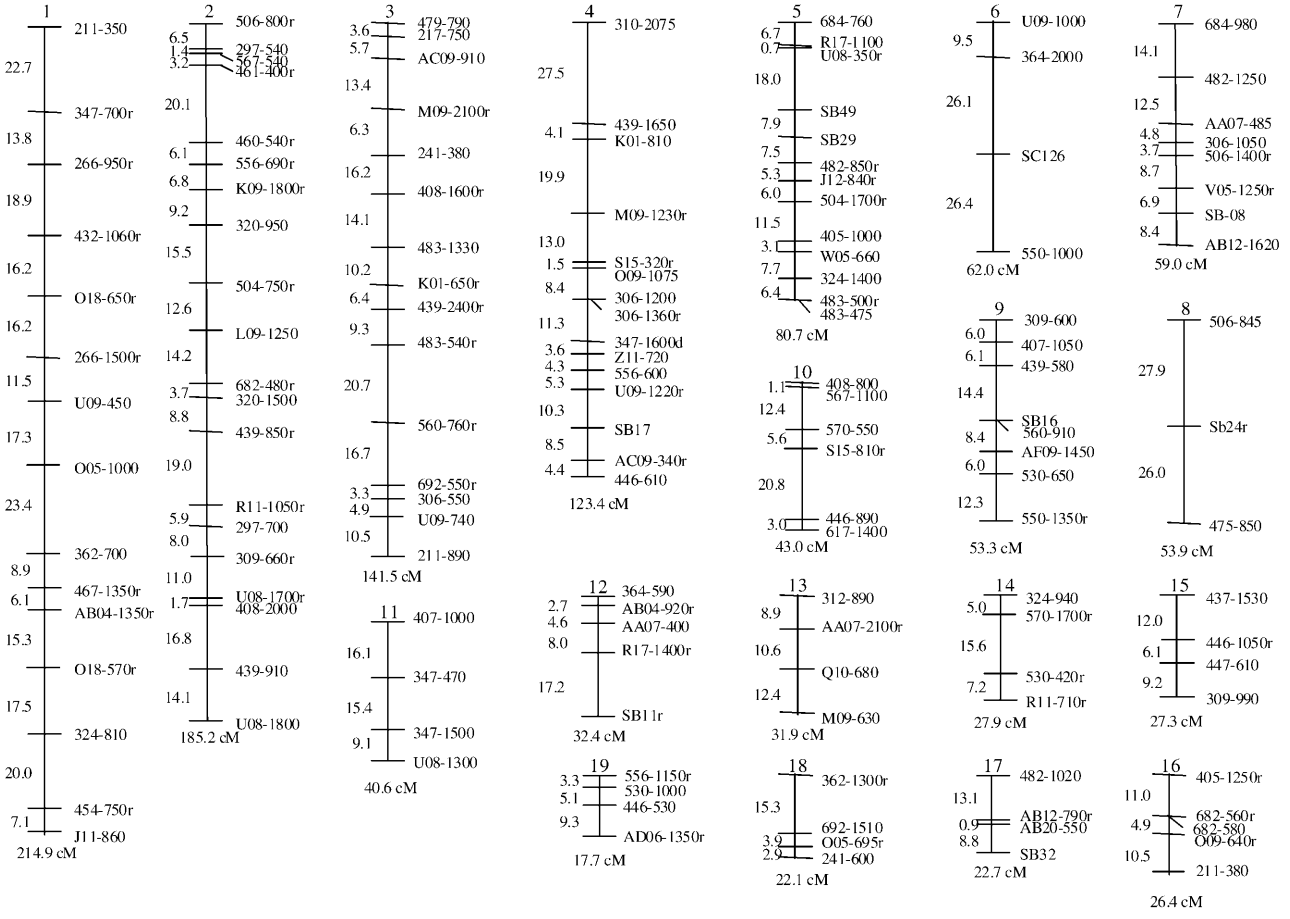

Fig. 4 Linkage map of parent 80132 obtained with JOINMAP. Names of framework markers are indicated on the right of the linkage groups (RAPD named according to the text, SB=ESTP markers, and SC=SCAR markers). Those annotated with the letter r indicate markers in repulsion phase. Genetic distances in cM using the Kosambi's mapping function are indicated on the left of the linkage groups. The length of each linkage group (in cM) is shown below the groups

Table 3 Shared loci between the two individual maps

RAPD	SCAR	ESTP
405–1250	*SC126*	*SB08*[a]
408–1600		*SB16*
439–850		*SB24*
439–2400		*SB29*
483–500		
560–910		
AB04-1350		
AB20-550		
AF09-1450		
J12-840		
Z11-720		

[a] ESTP-SSCP locus (14 °C)

in 20 linkage groups including four groups of three loci, while 16 loci remained unlinked (including two ESTP loci) (Table 2). The size of the linkage groups ranged from 13.8 cM to 247.5 cM. The average distance between two framework loci was 13.9 cM, with ten gaps exceeding 25 cM. Three intervals between two loci were larger than 30 cM. Using the same criteria, the maps constructed by JOINMAP (Fig. 4) and GMENDEL had a total of 145 loci distributed in 19 linkage groups, including one group of three loci, and 14 loci remained unlinked (including one ESTP). The unlinked loci were the same between the three softwares, but there were two more unlinked loci with MAPMAKER. The eight ESTP loci were distributed on seven different linkage groups. With JOINMAP, the size of the linkage groups ranged from 17.7 cM to 214.9 cM. The average framework locus spacing was 10.0 cM, with 12 gaps exceeding 20 cM. Only six intervals between two loci were larger than 25 cM.

Out of 20 markers being polymorphic simultaneously for both parents, there were 16 loci (11 $RAPD_s$, 1 SCAR and 4 $ESTP_s$, including 1 ESTP-SSCP) shared between the two maps (Table 3). The difference between the number of shared polymorphic markers and the number of shared loci is explained by the number of unlinked loci (four). These results implicate that 44% of the mapped ESTP loci and 9% of the mapped RAPD loci were shared between the two individual maps. Figure 5 shows six of these shared loci.

Genome length and map coverage

Using MAPMAKER, the estimated genome length for the M2 map was E(G) = 2,744 cM and for the 80132 map, E(G) = 3,072 cM (Table 4). Using JOINMAP, the

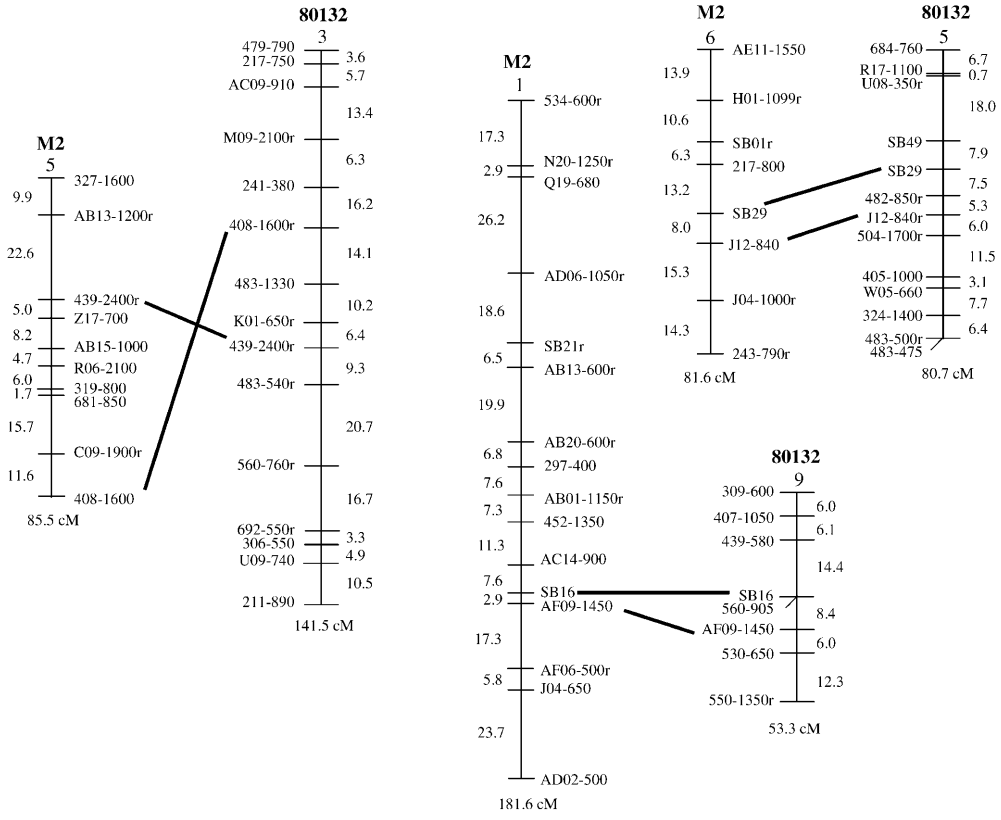

Fig. 5 Three linkage groups of the M2 and 80132 maps obtained with JOINMAP which have common markers. Markers in common between maps are indicated with bold lines

Table 4 Estimations of genome length and map coverage[a]

Parent	Software	E(G), (K; X) (cM)	C_O (cM)	% Coverage	I_α (G)	Expected number of loci required for near saturation
M2	MAPMAKER	2,743.7 (377; 31.4)	2,476.5	90.3	(2,492–3,052)	254
	JOINMAP	2,363.3 (396; 28.4)	2,059.4	87.1	(2,151–2,622)	220
80132	MAPMAKER	3,071.6 (287; 31.8)	2,330.2	75.9	(2,753–3,473)	286
	JOINMAP	2,705.6(304; 29.7)	2,006.6	74.2	(2,432–3,048)	251

[a] Abbreviations used: E(G), genome length; K, the number of locus pairs having LOD values at or above 4; X, the maximum observed map distance among the locus pairs above a threshold lod; C_o, the observed genome map coverage; I_α (G), the confidence interval for G; cM, centimorgan

estimated genome length was smaller: 2,363 cM and 2,706 cM for the M2 and 80132 maps, respectively. For the maps constructed with MAPMAKER, the observed genome map coverage, C_O, was 2,477 cM for the M2 map and 2,330 cM for the 80132 map. For the two maps constructed with JOINMAP, C_O was 2,059 cM and 2,007 cM for the M2 and 80132 maps, respectively. The M2 map covered between 87 and 90% of the total genetic distance, according to the different softwares (Table 4). The map of 80132 covered between 74 and 76% of the total genetic distance, depending on the software used (Table 4). This indicates that the mapped loci did not provide a full coverage of the white spruce genome, and the number of linkage groups did not correspond to the number of haploid chromosomes (n = 12). The expected number of loci required to achieve near saturation was between 220 and 254 for M2 and between 251 and 286 for 80132, depending on the software used (Table 4).

Discussion

From the analysis of haploid megagametophyte DNA arrays, the proportion of RAPD primers revealing polymorphisms was 26% and 23% for the parents M2 and 80132, respectively. The average proportion of 24% RAPD primers revealing polymorphisms is slightly higher than results obtained in other conifer genome-mapping studies using RAPDs, such as 19% in *Pseudotsuga menziesii* (Krutovskii et al. 1998), 16% for *Pinus elliottii* (Nelson et al. 1993) and 19.6% for *Pinus pinastes* (Plomion et al. 1995). However, if primers rejected at the preliminary stage for poor or null amplification are taken

into account (456 primers rejected, see Materials and Methods), the rate of primers revealing polymorphisms drops to 12% and 10% for the parents M2 and 80132, respectively. The proportion of EST primer pairs revealing ESTPs was 28% and 25% for each of the parents M2 and 80132. In *Pinus taeda*, Temesgen et al. (2001) have reported a rate of 50 and 70% of ESTs revealing ESTPs for two cDNA banks. They used three methods to reveal polymorphisms: agarose gels, SSCP and DGGE (denaturing gradient gel-electrophoresis). The third method revealed more DNA polymorphisms than the other two methods, which would explain their higher success rate, on average, than that observed in the current study where only agarose gels and SSCP were used to reveal ESTPs. For SCAR primer pairs, such comparisons of efficiency are difficult to conduct due to the paucity of data in other species. If we consider the average rate of polymorphism per map revealed for each type of marker, RAPD 24%, SCAR 22% and EST 26%, these values are just about in the same range with the detection methods used in this study. Similar rates of polymorphism per primer or primer pair were reported in *P. mariana* for RAPDs (Isabel et al. 1995) and ESTs (Perry and Bousquet 1998a).

Of the 223 polymorphic markers analysed for the parent M2, 41 (19.4%, all of which were RAPDs) did not exhibit the expected 1:1 segregation ratio. For the parent 80132, 12.8% (all of which were RAPDs but with one ESTP) did not exhibit the expected 1:1 ratio. These rates of rejection are higher than those reported for RAPD markers in *P. elliottii*. (10%) (Nelson et al. 1993), *P. abies* (2.8%) (Skov and Wellendorf 1998), *Pinus pinaster* (1.3%) (Plomion et al. 1995), and *Pinus strobus* (3%) (Echt and Nelson 1997). However, our rejection rates are similar to those for RAPD markers in *P. menziesii* (18%) (Krutovskii et al. 1998) and lower than those obtained for white spruce by Tulsieram et al. (1992). Segregation distortion of RAPD fragments could result from comigration of fragments coded by multiple loci, competition among annealing sites for reaction components, or other non-biological causes (Krutovskii et al. 1998). Distorted markers were not mapped in this study, but there is no consensus among researchers as to whether markers showing segregation distortion should be used in linkage studies.

In general, the LOD values used in conifer mapping studies vary from 3.0 to 5.0. In this study, the number of linkage groups with three or more loci was the same with a LOD value of 3.5 or 4.0 for the parent M2, using three different softwares. For the parent 80132, the number of groups was slightly different with a LOD value of 3.5 or 4.0. Since the number of groups was basically the same with a LOD value of 3.5 or 4.0, we chose 4.0 as a more severe threshold value for the rest of the analyses, and kept the same criterion for the ordering step. About 7% and 10% of the loci remained unlinked on maps M2 and 80132, respectively, after the grouping step. After the ordering step, there was 20% and 10% of the loci not ordered for maps M2 and 80132, respectively. Our results are similar to those found in larch by Arcade et al. (2000). In *P. elliottii*, Nelson et al. (1993) found about 42% of unlinked loci and 21.6% of accessory loci. If we were to decrease the LOD value and increase the recombination fraction, we would have obtained less unlinked and unordered loci.

In *Picea*, the number of chromosomes per haploid genome is 12. In this study, we found 17 major linkage groups, plus six triplets of loci for the map M2 and 17 major linkage groups, plus three triplets of loci, for the map 80132. The failure to obtain the basic number of chromosomes has also been reported for other tree species (e.g., Binelli and Bucci 1994 in *P. abies*; De Simone et al. 1997 in *Cichorium intybus*; Travis et al. 1998 in *Pinus edulis*; Krutovskii et al. 1998 in *P. menziesii*; Arcade et al. 2000 in larch; Wu et al. 2000 in *Populus deltoides*). In each of these studies, and in our study, a large number of RAPD markers were used but only a few markers of other types, such that the lack of polymorphic markers in particular chromosomal regions may be attributable either to the marker system used and/or to the presence of large homozygous regions in the spruce genome. Further analyses involving an increase in the type and the number of markers used will be needed in order to fill-in gaps.

In the only other map estimated for *P. glauca*, which was estimated with MAPMAKER and entirely composed of RAPD markers (Tulsieram et al. 1992), 12 linkage groups were found using 61 loci on a panel of 47 megagametophytes, with an average inter-locus distance of 23.5 cM. The total genetic distance covered was 873.8 cM, approximately 1/3 of the genome. Our maps showed an improved coverage of the genome, with 90 and 76% (using MAPMAKER) of the genome covered with an average inter-locus distance of 15.5 and 13.9 cM, respectively, for each of the M2 and 80132 maps. However, the main objective of the Tulsieram's (1992) study was to assess the utility of RAPD markers for genetic mapping using haploid DNA while our main objective was to estimate two individuals maps.

Based on the haploid DNA content, it is estimated that the average conifer genome size is about 2,500 cM (Neale and Williams 1991). Our maps (using JOINMAP) had an observed genome map coverage (C_O) between 2,006.6 cM and 2,476.5 cM, and they covered between 74.2 and 90.3% of the genome, respectively, for 80132 and M2. The length of the genome of *P. glauca* appears to fall in the range of the total genetic distances estimated in conifer mapping: 3,584 cM in *P. abies* (Binelli and Bucci 1994), 2,198 cM in *P. abies* (Paglia et al. 1998), 1380 cM in *P. pinaster* (Plomion et al. 1995), 1,204 cM in *P. strobus* (Echt and Nelson 1997), 2,012 cM in *P. edulis* (Travis et al. 1998), 2,020 and 2,421 cM in *Larix decidua* and *Larix kaempferi*, respectively (Arcade et al. 2000).

Our estimates of genome length [E(G)], 2,363 to 3,071 cM, are also in agreement with previous conifer mapping studies: 2,839 cM in *P. abies* (Paglia et al. 1998), 2,071 cM in *P. strobus* (Echt and Nelson 1997),

2,390 cM in *P. edulis* (Travis et al. 1998), 2,537 and 2,997 cM in *L. decidua* and *L. kaempferi*, respectively (Arcade et al. 2000). Some of the differences in genome length between studies can be accounted for by the different methods used for map construction, types of markers, number of individuals, number of markers, frequency of recombination and LOD values, and the software employed. Another significant issue that hampers all aspects of map development is the experimental error derived from misscored individuals or contamination. Such errors tend to inflate the number of apparent recombinants and expand map distance, which is especially severe when markers are tightly linked (Wu et al. 2000).

Despite a low average distance between adjacent loci (between 13.9 and 15.5 cM) and the high percentage of genome coverage, based on a comparison of observed and predicted total map sizes (between 74 and 90%), there were many inter-locus intervals larger than 20 cM. Assuming a random distribution of loci, we would expect to achieve near saturation with 220 loci for map M2 and 251 loci for map 80132, using JOINMAP. Thus, about 55 loci should be added on map M2 and 106 loci on map 80132 to obtain saturated maps.

The few mapped ESTP loci were widely distributed along the estimated maps. Also, the percentage of shared loci between the maps M2 and 80132 was higher for ESTPs than for RAPDs (44% versus 9%). The difference is even greater if one considers that RAPD amplification products of the same size from different parental backgrounds may not represent the same locus in all cases since no sequencing was performed to ascertain marker-locus homology. In addition, the ESTP markers resolved were codominant, locus specific, easy to use, with a proportion of polymorphism slightly higher per EST primer pair than per RAPD primer. Hence, the development of additional ESTP markers appear well supported, especially if one wants to construct concensus maps. ESTPs are also markers for coding regions, thus they can assist in locating candidate genes on current maps. The development of additional ESTPs and candidate ESTs is in progress (Temesgen et al. 2001 for pines; Neale, Devey, Isabel, Plomion, Richardson, Sederoff and Wheeler for the Conifer Comparative Genomics Project, CCGP). Also, because of their informativeness across pedigrees or populations, ESTP markers should be most useful for marker-assisted breeding and selection programs. Unlike many other markers commonly used in forestry, ESTPs can be used as orthologous markers for comparative mapping, to map genes of known function, or to identify candidate genes affecting important traits (Temesgen et al. 2001). Until sufficient numbers of ESTPs, or similarly informative markers, are available for constructing full-coverage genome maps, framework maps built from anonymous markers, such as RAPDs, can be used to construct maps rapidly. Once a full coverage *P. glauca* ESTP framework map is constructed, it should be useful for comparative mapping and breeding applications within and among *Picea* species.

The results obtained with the three different softwares were almost similar: the number of linkage groups and the ordering was almost the same. The main difference was at the level of the estimated genome length. The genome length and the observed genome map coverage were larger with MAPMAKER than with JOINMAP. Such a difference was also reported in barley and was attributed to how each program calculates map distance when the actual interference differs from that assumed (Qi et al. 1996). Jermstad et al. (1998) in *P. menziesii* and Sewell et al. (1999) in *P. taeda* also observed such a difference. JOINMAP accepts haploid data unlike MAPMAKER and GMENDEL where data have to be re-coded. Double cross-overs and possibly misscored individuals or loci can be identified by the command“ERROR DETECTION” in the MAPMAKER software, but it appears that this command has to be used carefully. When the error detection option is on, the software rectifies some data to facilitate the analyses, but it is preferable to check the supposed errors with the command “GENOTYPE” and to score again the data entry. If the software corrects some false errors, this will result in differences in group ordering and in genome length. GMENDEL does not estimate the genome length and the genome map coverage in cM, which is an obvious disadvantage. On the other hand, GMENDEL appears to be a reliable software to validate or compare the results obtained with other methods because it estimates error statistics that other softwares do not estimate, using Monte Carlo simulations or bootstrap resampling.

Our results show that white spruce has a rather large and complex genome, and that a large number of marker loci will be needed before comprehensive QTL and marker-aided selection projects take place successfully in this species. The construction of the current maps showed that dominant RAPD markers could be used to rapidly construct individual linkage maps, but some other types of markers, such as co-dominant ESTPs, will become increasingly important for consensus mapping and to further detect associations between candidate genes and quantitative traits or known metabolic functions. Collaborative efforts are currently underway to integrate genome maps and genetic markers in conifers (Conifer Comparative Genomics Project).

Acknowledgements We thank D. Fournier, S. Gagnon, and M. Lamothe (Canadian Forest Service, Laurentian Forestry Centre) for help in the laboratory, and Dr. Jean Beaulieu for providing some of the seed materials. This project was made possible by grants from the Canadian Forest Service, Fonds FCAR du Québec and the National Sciences and Engineering Research Council of Canada.

References

Arcade A, Anselin F, Faivre Rampant P, Lesage MC, Pâques LE, Prat D (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. Theor Appl Genet 100:299–307

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

Bishop DT, Cannings C, Skolnick M, Williamson JA (1983) The number of polymorphic DNA clones required to map the human genome. In: Weir BS (ed) Statistical analysis of DNA sequence data. Marcel-Dekker, New York, pp 181–200

Corriveau A, Beaulieu J, Daoust G (1991) Heritability and genetic correlations of wood characters of Upper Ottawa Valley white spruce populations grown in Quebec. For Chron 67:698–705

De Simone M, Morgante M, Lucchin M, Parrini P, Marocco A (1997) A first linkage map of *Cichorium intybus* L. using one-way pseudo-testcross and PCR-derived markers. Mol Breed 3:415–425

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

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

Fournier D, Perry DJ, Beaulieu J, Bousquet J, Isabel N (2001) Optimizing expressed sequence tag polymorphisms by single strand conformation polymorphism in spruces. For Genet (in press)

Furnier GR, Stine M, MohnCA, Clyde MA (1991) Geographic patterns of variance in allozymes and height growth in white spruce. Can J For Res 21:707–712

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

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

Holloway JL, Knapp SJ (1993) GMENDEL 3.0 Users Guide. Oregon State University, Corvallis

Hulbert S, Ilott T, Legg EJ, Lincoln S, Lander E, Michelmore R (1988) Genetic analysis of the fungus *Bremia lactucae*, using restriction length polymorphism. Genetics 120:947–958

Isabel N, Tremblay L, Michaud M, Tremblay FM, Bousquet J (1993) RAPDs as an aid to evaluate the genetic integrity of somatic embryogenesis-derived populations of *Picea mariana* (Mill.) B.S.P. Theor Appl Genet 86:81–87

Isabel N, Beaulieu J, Bousquet J (1995) Complete congruence between gene diversity estimates derived from genotypic data at enzyme and random amplified polymorphic DNA loci in black spruce. Proc Natl Acad Sci USA 92:6369–6373

Jaramillo-Correa JP, Beaulieu J, Bousquet J (2001) Constrasting evolutionary forces driving population structure at ESTPs, allozymes and quantitative traits in white spruce. Mol Ecol 10:2729–2740

Jermstad KD, Bassoni DL, Wheeler NC, Neale DB (1998) A sex-averaged genetic linkage map in coastal Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco var "*menziesii*"] based on RFLP and RAPD markers. Theor Appl Genet 97:762–770

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) Reproductibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol Breed 3:381–390

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Krutovskii KV, Vollmer SS, Sorensen FC, Adams WT, Knapp SJ, Strauss SH (1998) RAPD genome maps of Douglas-fir. J Hered 89:197–205

Kubisiak TL, Nelson CD, Nance WL, Stine M (1996) Comparison of RAPD linkage maps constructed for a single longleaf pine from both haploid and diploid mapping populations. For Genet 3:203–211

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

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

Lefort F, Echt C, Streiff R, Vendramin GG (1999) Microsatellite sequences: a new generation of molecular markers for forest genetics. For Genet 6:15–20

Li P, Beaulieu J, Corriveau A, Bousquet J (1993) Genetic variation in juvenile growth and phenology in a white spruce provenance-progeny test. Silvae Genet 42:52–60

Li P, Beaulieu J, Bousquet J (1997) Genetic structure and patterns of genetic variation among populations in eastern white spruce (*Picea glauca*). Can J For Res 27:189–198

Lincoln SE, Daly MJ, Lander ES (1993) Constructing genetic linkage maps with MAPMAKER/EXP version 3.0: a tutorial and reference manual. Whitehead Institute, M.I.T. Cambridge, Massachusetts

Merril CR, Goldman D, Sedman SA, Ebert MH (1981) Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. Science 211:1437–1438

Neale DB, Williams CG (1991) Restriction fragment length polymorphism mapping in conifers and applications to forest gentics and tree improvement. Can J For Res 21:545–554

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

Paglia GP, Olivieri AM, Morgante M (1998) Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). Mol Gen Genet 258:466–478

Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985–993

Perry DJ, Bousquet J (1998a) Sequence-tagged-site (STS) markers of arbitrary genes: development, characterization and analysis of linkage in black spruce. Genetics 149:1089–1098

Perry DJ, Bousquet J (1998b) Sequence-tagged-site (STS) markers of arbitrary genes: the utility of black spruce-derived STS primers in other conifers. Theor Appl Genet 97:735–743

Perry DJ, Isabel N, Bousquet J (1999) Sequence-tagged-site (STS) markers of arbitrary genes: the amount and nature of variation revealed in Norway spruce. Heredity 83:239–248

Plomion C, O'Malley DM, Durel CE (1995) 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

Qi X, Stam P, Lindhout P (1996) Comparison and integration of four barley genetic maps. Genome 39:379–394

Scotti I, Troggio M, Soranzo N, Vendramin GG, Bucci G (1998) A new set of PCR-based, locus-specific markers for *Picea abies* (L.) Karst. Mol Ecol 7:789–792

Sewell MM, Sherman BK, Neale DB (1999) A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. Genetics 151:321–330

Skov E, Wellendorf H (1998) A partial linkage map of *Picea abies* clone V6470 based on recombination of RAPD-markers in haploid megagametophytes. Silvae Genet 47:273–282

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. Plant J 3:739–744

Stam P, van Ooijen JW (1995) JOINMAP version 2.0: software for the calculation of genetic linkage maps. DLO-Center for Plant Breeding and Reproduction Research, Wageningen, The Netherlands

Temesgen B, Brown GR, Harry DE, Kinlaw CS, Sewell MM, Neale DB (2001) Genetic mapping of expressed sequence tag polymorphism (ESTP) markers in loblolly Pine (*Pinus taeda* L.). Theor Appl Genet 102:664–675

Torres AM, Weeden NF, Martin A (1993) Linkage among isozyme, RFLP and RAPD markers in *Vicia faba*. Theor Appl Genet 85:937–945

Travis SE, Ritland K, Whitham TG, Keim P (1998) A genetic linkage map of Pinyon pine (*Pinus edulis*) based on amplified fragment length polymorphisms. Theor Appl Genet 97:871–880

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

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 23:4407–4414

Wu RL, Han YF, Hu JJ, Fang JJ, Li L, Li ML, Zeng Z-B (2000) An integrated genetic map of *Populus deltoides* based on amplified fragment length polymorphisms. Theor Appl Genet 100:1249–1256