Theor Appl Genet (2005) 111: 1466–1488
DOI 10.1007/s00122-005-0068-2

ORIGINAL PAPER

**Betty Pelgas · Jean Bousquet · Stéphanie Beauseigle
Nathalie Isabel**

A composite linkage map from two crosses for the species complex *Picea mariana* × *Picea rubens* and analysis of synteny with other Pinaceae

Received: 22 February 2005 / Accepted: 4 July 2005 / Published online: 7 October 2005
© Springer-Verlag 2005

Abstract Four individual linkage maps were constructed from two crosses for the species complex *Picea mariana* (Mill.) B.S.P. × *Picea rubens* Sarg in order to integrate their information into a composite map and to compare with other Pinaceae. For all individual linkage maps, 12 major linkage groups were recovered with 306 markers per map on average. Before building the composite linkage map, the common male parent between the two crosses made it possible to construct a reference linkage map to validate the relative position of homologous markers. The final composite map had a length of 2,319 cM (Haldane) and contained a total of 1,124 positioned markers, including 1,014 AFLPs, 3 RAPDs, 53 SSRs, and 54 ESTPs, assembled into 12 major linkage groups. Marker density of the composite map was statistically homogenous and was much higher (one marker every 2.1 cM) than that of the individual linkage maps (one marker every 5.7 to 7.1 cM). Synteny was well conserved between individual, reference, and composite linkage maps and 94% of homologous markers were colinear between the reference and composite maps. The combined information from the two crosses increased by about 24% the number of anchor markers compared to the information from any single cross. With a total number of 107 anchor markers (SSRs and ESTPs), the composite linkage map is a useful starting point for large-scale genome comparisons at the intergeneric level in the Pinaceae. Comparisons of this map with those in *Pinus* and *Pseudotsuga* allowed the identification of one breakdown in synteny where one linkage group homoeologous to both *Picea* and *Pinus* corresponded to two linkage groups in *Pseudotsuga*. Implications for the evolution of the Pinaceae genome are discussed.

Keywords Codominant markers · Colinearity · Comparative mapping · Consensus map · Pinaceae · Synteny

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00122-005-0068-2

Communicated by O. Savolainen

B. Pelgas · J. Bousquet · S. Beauseigle · N. Isabel
Chaire de recherche du Canada en génomique forestière et environnementale, Centre de recherche en biologie forestière, Pavillon Charles-Eugène-Marchand, Université Laval, Sainte-Foy, QC, G1K 7P4, Canada

N. Isabel (✉)
Natural Resources Canada, Laurentian Forestry Centre, Canadian Forest Service, 1055 Rue du PEPS, C.P. 3800, Sainte-Foy, QC, G1V 4C7, Canada
E-mail: nisabel@cfl.forestry.ca
Tel.: +1-418-648-7137
Fax: +1-418-648-5849

Introduction

For species with large and unsequenced genomes such as in conifers, comparative mapping is an alternative strategy to understand genome organization and to highlight homoeologous chromosomal segments involved in economical and adaptive traits (Kumar et al. 2000; Chagné et al. 2003; Devey et al. 2004). Within the conifers, genetic linkage maps have been constructed for numerous species including *Cryptomeria japonica* (Mukai et al. 1995; Nikaido et al. 2000), *Pinus elliottii* and *Pinus caribaea* (Brown et al. 2001; Shepherd et al. 2003), *Pinus pinaster* (e.g. Plomion et al. 1995; Costa et al. 2000; Chagné et al. 2002), *Pinus radiata* (Cato et al. 1999), *Pinus sylvestris* (e.g. Yin et al. 2003), *Pinus taeda* (e.g. Devey et al. 1994; Remington et al. 1999; Brown et al. 2001), *Pseudotsuga menziesii* (Krutovskii et al. 1998), *Picea abies* (e.g. Bucci et al. 1997; Paglia et al. 1998; Acheré et al. 2004), and *Picea glauca* (Tulsieram et al. 1992; Gosselin et al. 2002). Recently, composite maps have been published for *P. abies* (Acheré et al. 2004), *P. taeda* (Sewell et al. 1999), and *C. japonica* (Tani et al. 2003), as well as preliminary comparative studies of linkage maps of the hard pines between *P. taeda* and *P. radiata* (Devey et al. 1999), *P. elliottii*

(Brown et al. 2001), *P. pinaster* (Chagné et al. 2003), and *P. sylvestris* (Komulainen et al. 2003). Such comparative studies are also being conducted at the intergeneric level (Krutovsky et al. 2004).

Comparative mapping and the identification of homoeologous regions between crosses within species and between species require a large number of orthologous and anchor markers. Although amplified fragment length polymorphisms (AFLPs, Vos et al. 1995) are useful to help covering genomes conveniently, their value as anchor markers is limited because their orthology remains difficult to assess, especially at the interspecific level (Waugh et al. 1997; Marques et al. 1998). Sequence-based markers such as expressed sequence tag polymorphisms (ESTPs) and single-locus simple sequence repeats (SSRs) are more appropriate to anchor maps because they are usually orthologous across congeneric species (Pelgas et al. 2004). However, the development of these markers is labor intensive, which seriously hampers comparative genome mapping. Indeed, only a few hundreds of ESTP and SSR markers are available for a limited number of conifer species, e.g. *P. taeda*, *P. pinaster*, *P. abies*, *P. glauca*, *Picea mariana*, *P. menziesii*, and *C. japonica* (Tsumura et al. 1997; Harry et al. 1998; Perry and Bousquet 1998a; Hodgetts et al. 2001; Temesgen et al. 2001; Krutovsky et al. 2004; Pelgas et al. 2004; http://www.pierroton.inra.fr/genetics/pinus/primers.html).

An additional challenge with comparative genome mapping is that only polymorphic markers showing Mendelian segregation within a given progeny can be mapped. One way to alleviate this problem is to use multiallelic, codominant markers such as SSRs or ESTPs to maximize the number of informative mapped markers. Another way is to use more than one cross for a given species, with at least one parent in common. Such strategies should increase the probability that a marker segregates (Beavis and Grant 1991; Kowalski et al. 1994; Lan et al. 2000; Tani et al. 2003), thus increasing marker informativeness for comparative mapping. The use of a common parent between mapping populations also allows verification of the consistency of marker ordering between the various mapping populations (Gentzbittel et al. 1995). These strategies have however rarely been applied in plant genome mapping.

In this study, we constructed the first linkage maps for the North American species complex *P. mariana* × *P. rubens*. *P. mariana* is an ecologically and economically important component of the North American boreal forests with a distribution area extending from the Pacific to the Atlantic coast. *P. rubens* has a more restricted distribution and is limited to northeastern North America. These two species hybridize naturally (e.g. Perron and Bousquet 1997) and represent a recent progenitor-derivative species pair (Perron et al. 2000). A total of four individual linkage maps as well as a reference and a composite linkage map were estimated, an unprecedented effort in the conifers. In particular, our intent was to assess the usefulness of codominant SSR and ESTP markers, in combination with the use of two crosses with one parent in common, for the assembly of a composite map reflecting the marker order of parental maps. In addition, we determined possible homoeologous linkage groups with other Pinaceae taxa in an effort to examine synteny between genera in this family.

Materials and methods

Plant material and DNA extraction

Plant material consisted of 118 and 85 progeny derived from two crosses between presumed *P. mariana* individuals: a backcross-like cross, hereafter called BC1 (#9920002: ♀11307-03 [♀83×♂425] ×♂425), and an outbred *F1* cross, hereafter called F1 (#S11991V: ♀422×♂425). However, after verification with species-specific molecular markers (Perron et al. 1995, 2000), both crosses harbored some genetic background from *P. rubens* (see "Results" section). Two crosses were used instead of a single cross to increase the number of segregating codominant markers. In addition, the two crosses had one male parent (425) in common to allow a better comparison of colinearity between crosses and to insure a better anchoring of markers. For both crosses, needle tissue was collected from the two parents and the progeny. Genomic DNA was extracted from the collected needles with the DNeasy Plant Mini Kit (Qiagen, Mississauga, ON, CA).

Markers analyzed

AFLP markers

The AFLP protocol used was essentially based on Vos et al. (1995) with some modifications. The enzyme combinations *Eco*R1/*Mse*1 and *Pst*1/*Mse*1 were used (Paglia and Morgante 1998; Remington et al. 1999; Wu et al. 2000; Hayashi et al. 2001). The digestion of the genomic DNA (2.5 h at 37°C) was conducted as described by Vos et al. (1995), except that the reaction volume was 25 µl. The ligation solution (10 µl) containing 1.75 pmol of *Eco*R1 adapter (or *Pst*1), 17.5 pmol of *Mse*1 adapter, 1.2 mM ATP, 1 U of T4 DNA ligase (Invitrogen, UK) and 5× reaction buffer was added to the digested DNA template, which was then incubated for 2.5 h at 37°C.

A total of 99 *Eco*R1/*Mse*1 and 14 *Pst*1/*Mse*1 primer combinations were used, with one or two selective nucleotides for the pre-amplification and three or four selective nucleotides for the selective amplification (Table S1, electronic supplementary material). *Eco*R1 and *Pst*1 primers were labeled with infrared dye IRDye™ 700 or IRDye 800 (LI-COR Biosciences, Lincoln, NE, USA). Pre-amplification reactions were done in 50 µl volume containing 0.4× PCR buffer (Invitrogen), 2 U of *Taq* DNA polymerase (Invitrogen), 0.2 mM of dNTPs (Amersham Biosciences), 75 ng of each primer, 0.6 mM

of $MgCl_2$, and 5 µl of digested-ligated DNA fragments. PCR was carried out using a PTC-225 thermal cycler (MJ Research, Reno, NV, USA) with the following conditions: 5 min at 94°C for initial denaturation, 35 cycles of 1 min at 94°C, 1 min 30 s at 56°C and 1 min 30 s at 72°C. Selective amplification reactions consisted of 1.2× reaction buffer (Invitrogen), 1 U of *Taq* DNA polymerase (Invitrogen), 0.2 mM of dNTPs, 5.2 ng of IRDye 700 labeled *Eco*R1 (or *Pst*1) primer, or 6.6 ng of IRDye 800 labeled *Eco*R1 primer, 15 ng of *Mse*1 primer, 1.75 mM of $MgCl_2$ and 2.5 µl of diluted pre-amplification products (1:190) in a volume of 10 µl. Cycle parameters were similar to those described by Vos et al. (1995), except that a denaturation step of 5 min at 94°C was added at the beginning.

AFLP reaction products were visualized on a two-dye IR2 DNA Analyzer System (model 4200; LI-COR Biosciences) according to Myburg et al. (2001), except that 7.5% of Long Ranger polyacrylamide (BioWhittaker Molecular Applications, Rockland, Maine, USA) was added to denaturing gels. Run parameters were as follows: 1500 V, 23 mA, 33 W, 45°C, signal channel 3, and motor speed 3. The digital AFLP gel images were scored using the AFLP-Quantar software version 1.09 (Key Gene Products B.V., Wagenningen, The Netherlands).

SSR markers

A total of 43 SSR markers previously developed by several authors (Pfeiffer et al. 1997; Hodgetts et al. 2001; Rajora et al. 2001; Scotti et al. 2000, 2002a, 2002b; Besnard et al. 2003) were screened. PCR reactions were based on the protocol of Acheré et al. (2004) with some minor modifications. DNA amplifications were performed in volumes of 10 µl containing 20 ng of DNA, 1× reaction buffer, between 2.5 and 5 mM of $MgCl_2$ depending on the marker tested (Table S2, electronic supplementary material), 2 µM of dNTPs, 2 µM of each primer and 0.3 U of Platinum® *Taq* DNA polymerase (Invitrogen). To reduce the occurrence of multiple-banding patterns, four different PCR programs were used according to the primer pairs tested (Table S2, electronic supplementary material). DNA samples were amplified on a PTC-225 thermal cycler (MJ Research).

Fragment analyses were carried out on a two-dye IR2 DNA Analyzer System (Model 4200; LI-COR Biosciences). For each sample, 7 µl of formamide tracking-dye (EMD Chemicals Inc., Gibbstown, NJ, USA) were added to 3 µl of PCR reaction and then run on a 6.5% Long Ranger polyacrylamide gel (BioWhittaker Molecular Applications). Digital SSR gel images were scored using the SAGA^{GT} software Version 3.0 (LI-COR Biosciences).

ESTP markers

A total of 107 ESTP markers developed from various conifer species were screened, as detailed in Pelgas et al. (2004). PCR conditions and the methods used to detect polymorphisms were as described elsewhere (Pelgas et al. 2004), except that the primer concentration was 0.12 µM in PCR. Additionally, six transcription factor genes (*KN1, KN2, KN3, KN4, HAP3a*, and *HAP3b*) were mapped. PCR conditions and the method used to detect polymorphism for the *Knox-I* genes are described elsewhere (Guillet-Claude et al. 2004), except that the primer concentration was 0.2 µM in PCR. For the last two genes (*HAP3a* and *HAP3b*), the detailed PCR conditions and the method used to detect polymorphism are available from the journal web site as electronic supplementary material (Table S3, electronic supplementary material).

Species-specific markers

Because both crosses used in this study were derived from trees located in the sympatric zone between *P. mariana* and *P. rubens*, species-specific markers were used to infer the genetic background of the parents. These markers were also used for mapping when at least one parent was heterozygote (Aa) and the other one heterozygote (Aa) or recessive homozygote (aa). The species-specific markers consisted of six RAPDs ($OPA06_{900}$, $OPA19_{1,250}$, $OPD14_{590}$, $OPE09_{603}$, $OPF17_{800}$, $OPG12_{1,000}$; Perron et al. 1995) and two ESTPs (*Sb62* and *Sb70*; Perron et al. 2000). RAPD primers were obtained from Operon Technologies (Alameda, CA, USA), and ESTP primer pairs were previously developed by Perry and Bousquet (1998a). For RAPDs and ESTPs, PCR reactions and amplification conditions were as described by Gosselin et al. (2002) and Pelgas et al. (2004; see above for primer concentration), respectively.

Screening procedure for data quality

All types of markers were first screened on the parents and a subset of the progeny of each cross, in order to identify the most informative AFLP primer combinations and the polymorphic SSR and ESTP markers. All selected loci (AFLPs, SSRs, and ESTPs) were scored independently from the gels by two observers. In case of conflicting observations, the locus was scored as "missing data" for the particular genotype. Segregation analyses were conducted for all AFLP, SSR, and ESTP markers for each cross in order to test departures from expected Mendelian ratios. To reduce the number of false-positives, a Bonferroni-corrected chi-square ($P < \alpha_{Bonf}$, $\alpha_{Bonf} = 0.01/n$) was used, where n is the number of tests performed (Sokal and Rohlf 1998). Distorted loci were excluded from further analyses in the particular cross where distortion was detected. The software Mapmaker v.3.0 (Lander et al. 1987) was used to check for double recombinants, but none were found. Moreover, no contaminant progeny was observed. For each cross, individuals with more than 10% of missing

data were discarded from the input data before map construction in order to obtain a homogeneous data set (Myburg et al. 2003).

Estimation of linkage maps

The experimental design allowed the construction of four individual linkage maps from both crosses, the estimation of a reference map for the common male parent between the two crosses, and the assembly of a composite map of the species complex *P. mariana* × *P. rubens*. The reference linkage map was useful to obtain the best representation of the genome before assembling the composite linkage map. Each cross was analyzed using the "two-way pseudo-testcross" mapping approach (Grattapaglia and Sederoff 1994). All linkage analyses and map estimations were performed with JoinMap 3.0 (Stam 1993; Van Ooijen and Voorrips 2001) using the parameter CP (cross-pollination), because the parental genotype 83 (backcross-like) was unavailable, with a maximal threshold value of 5 for the jump, a ripple value of 1, and Haldane's mapping function (Haldane 1919). Kosambi's mapping function was also used to enable comparisons with other studies (Kosambi 1944). The grouping of markers was also conducted with CarthaGene (Schiex and Gaspin 1997) in order to compare results with those obtained with JoinMap 3.0.

Individual linkage maps

Only markers segregating 1:1 or 1:1:1:1 were used in the construction of individual linkage maps. Exceptionally, markers in configuration 3:1 or 1:2:1 for one cross and 1:1 for the other cross were also kept because they were informative in the other cross. When a pair of markers was considered identical, such as having a similarity rate of one (JoinMap command "similarity of loci"), only one of the markers was kept for the remaining analyses. For each of the four parental maps, marker grouping and ordering of linked loci were carried out using a LOD of 4.0 and a minimum recombination fraction (θ) of 0.35. For each marker, a minimum number of valid genotypes (at least 90% of the progeny sample) were required to include it into the data set for initial analyses (Hackett and Broadfoot 2003). In order to get the best possible order, the average contribution to the goodness-of-fit for each locus (JoinMap tabsheet "mean chi-square contribution") and the genotype probabilities with "the locus average $-\log_{10}(P)$" were checked. When the −10 base logarithm of the probability was higher than one, markers were excluded from the analyses to avoid erroneous locations. The order obtained for "map 2 of the second round" was retained as well as markers having a "jump" value ≤ 5 for "map 3 of the third round" (with the JoinMap command "Calculate Map"). This order was fixed to allow the positioning of additional markers that were excluded from the data sets for initial analyses because they did not fulfill one of the criteria for a priori inclusion (i.e. minimum number of genotypes and/or maximal jump value). Markers having a "jump" value >5 were considered as accessory markers because they could not be ordered with confidence.

Reference linkage map

The reference linkage map was constructed from the integration of both individual linkage maps of the male parent 425 common to the two crosses. At this step, AFLP markers segregating 3:1 were included in the data sets. All linkage analyses were conducted following the same procedure as described above. The "mean chi-square contribution" and the "$-\log_{10}(P)$" of each additional marker value were checked again. Then, to obtain the reference linkage map, the two data sets of homologous linkage groups, one from each cross, were merged using the JoinMap function "combine groups for map integration." Before marker ordering within each integrated linkage group, a "heterogeneity test" (JoinMap command) was conducted with the help of homologous markers to compare recombination frequency estimates between both data sets for each homologous pair of loci. These pairs of loci (file "Linkages" in JoinMap) showing significant differences ($P<0.01$) were eliminated to avoid erroneous marker ordering. Markers having a "jump" value >5 were considered as accessory markers.

Composite linkage map

The composite map was obtained from the combined analyses of the two crosses BC1 and F1. The assembly was conducted in three steps under the same conditions as for the determination of previous linkage maps (see above) with the inclusion of AFLP markers segregating 3:1 in each data set for each cross. The first step was to assemble female and male data sets of each cross in order to obtain one "sub-composite" linkage map for each of both crosses (Maliepaard et al. 1997), with the help of markers segregating 1:1:1:1, 1:2:1, and 3:1 between both individuals of each cross. This strategy allowed to integrate as much information as possible onto the composite map by the incorporation of the minor linkage groups, obtained previously during the construction of individual linkage maps, into the major linkage groups. The second step was to verify, for the two "sub-composite" maps, the average contribution to the goodness-of-fit for each locus (JoinMap tabsheet "mean chi-square contribution"). In addition, to get the best possible marker order, markers having a "locus average $-\log_{10}(P)$" with a value higher than 1, were excluded from analyses in order to avoid erroneous locations. The last step to obtain the composite map consisted in merging the data sets of homologous linkage groups from the two "sub-composite" maps using

the JoinMap function "combine groups for map integration". For the homologous pairs of loci, a "heterogeneity test" (JoinMap command) was conducted, as previously described for the reference linkage map. Markers having a "jump" value >5 were considered as accessory markers.

Linkage group nomenclature

For all maps, the linkage groups were numbered following the decreasing order of their length obtained on the composite map.

Genome lengths and map coverages

Observed and estimated genome lengths

The observed genome length (G_O) was estimated according to the method of Nelson et al. (1994), which takes into account all markers, linked and unlinked, such as

$$G_O = G_F + X_O(L - R) \quad (1)$$

where G_F is the total length of the map in cM, X_O is the observed maximum distance in cM between two linked markers at or above a minimum LOD threshold value of 4.0 ($Z=4$), L is the total number of linkage groups, pairs of loci and unlinked loci, and R is the haploid number of chromosomes. Observed genome lengths were adjusted in two ways: (1) by taking into account all linked and unlinked markers or (2) by taking into account only linked markers. The expected genome length, G_e, of each parental map was estimated under the assumption of random marker distribution according to Hulbert et al. (1988) and Chakravarti et al. (1991) such as

$$G_e = 2MX/K \quad (2)$$

with a confidence interval of $G_e/\left(1 \pm 1.96/\sqrt{K}\right)$, where $M = N(N-1)/2$ is the number of informative meioses (N is the number of linked loci), X is the maximum observed map distance among the locus pairs at/or above a minimum threshold LOD, and K is the number of locus pairs having LOD values at/or above Z. These estimations were obtained by taking into consideration only placed markers and by eliminating multiple markers at same loci to avoid an overestimation of the genome size.

Observed and expected genome map coverages

The observed map coverage C_O was evaluated in two ways: (i) the ratio of the observed genome length G_O on the estimated genome length G_e (Tani et al. 2003); or (ii) the ratio of the total length of the map G_F on the estimated genome length G_e (Cervera et al. 2001). The expected genome map coverage C_e was calculated according to Lange and Boehnke (1982) from the equation:

$$C_e = 1 - e^{-X_e N/1.25 G_e} \quad (3)$$

with an adjustment for chromosome ends, and where N is the number of linked loci, X_e is the maximum distance between two adjacent linked markers in cM at a minimum threshold LOD value (4.0 in this case), and G_e is the estimated genome length.

Distribution of markers

Two analyses were conducted to evaluate the distribution of markers with different types of inheritance (1:1, 1:1:1:1, 1:2:1, and 3:1) on linkage groups. These analyses were conducted for the composite linkage map and also for the four individual linkage maps, because the composite map is a representation of the assembly of both parents for each cross.

Heterogeneity of marker distribution among linkage groups

For the four individual maps and the composite linkage map, G-tests for goodness-of-fit (Sokal and Rohlf 1998) were used to test for heterogeneity of marker distribution among linkage groups (heterogeneity if $P<0.01$). Marker distribution was analyzed by taking into account the number of markers per linkage group and the size of the linkage groups. An arbitrary index of genome size, one genome unit = 10 cM, was defined to estimate the expected number of markers for each linkage group: the total number of markers positioned onto all linkage groups was divided by the total genome units for all linkage groups, and then multiplied by the total genome units of the estimated linkage group. All types of markers (AFLPs, SSRs, and ESTPs) were considered together for each linkage map. In a second analysis, only AFLP markers were considered. SSR and ESTP markers could not be considered independently of AFLPs because they were not frequent enough on each linkage group.

Coefficients of marker dispersion

When the distribution of markers was homogeneous, a coefficient of marker dispersion could be estimated for each map to evaluate the presence of marker clusters on the linkage groups and thus, to verify the presence of possible chromosome structures (Ma et al. 2004). The assessment of marker clusters on linkage groups was conducted by counting markers in all linkage groups of each parental map, using a sliding window of 10 cM (Haldane) and considering all types of markers (AFLPs, SSRs, and ESTPs). The size of the window was chosen to be higher than the average marker density, which was 6.6 cM (Haldane; see "Results" section). At the end of each linkage group, the last window was taken into account only when it was more than 7.5 cM (Cervera et al.

2001). For the composite map, the size of the window was reduced to 3 cM, due to higher marker density (see "Results" section). Observed and expected frequencies were compared with a chi-square test. The ratio between the variance and the mean of the number of markers per window provides a rapid method to test if observed frequencies are distributed following a Poisson distribution (Sokal and Rohlf 1998). This relationship can be expressed as the coefficient of dispersion (CD):

$$CD = s^2/\bar{Y} \quad (4)$$

where s^2 stands for the variance of the number of markers per window and $\bar{Y}$ is the mean number of markers per window. A CD value lower than 1 ($CD < 1$) indicates that there are less markers than expected in given intervals, whereas a CD value higher than 1 ($CD > 1$) indicates that more markers than expected are observed in given intervals, so that the markers are clustered.

Results

Genetic background of the parents

Out of the eight species-specific markers used to screen the genetic background of the three parents 11307-03 [83×425], 425, and 422 used in the crosses, three out of the six species-specific RAPD markers ($OPA06_{900}$, $OPA19_{1250}$, and $OPD14_{590}$) and both ESTP markers (*Sb62* and *Sb70*; Perron et al. 2000) revealed the presence of *P. rubens* genetic background, in addition to that of *P. mariana* (Table 1). The three additional RAPD markers ($OPE09_{603}$, $OPF17_{800}$, and $OPG12_{1000}$) did not indicate the presence of *P. rubens* elements into the genome of the three parents. These results show that the parents are of introgressed nature.

Marker screening

AFLP markers

Out of 113 AFLP primer combinations screened, 87 and 86 primer combinations were retained for crosses BC1 and F1, respectively (Table S1, electronic supplementary material). The total number of markers segregating 1:1 was 549 (64%) out of a total of 862 and 615 (71%) out of a total of 863 AFLP markers for crosses BC1 and F1, respectively (Table 2). Both crosses shared a total of 112 AFLP markers (56 in configuration 1:1 for both crosses, 9 in configuration 1:1 for BC1 and 3:1 for F1, 28 in configuration 3:1 for BC1 and 1:1 for F1, and 19 in configuration 3:1 for both crosses), which were considered as homologous because the same primer combinations amplified fragments of identical size and the two crosses had one parent in common.

SSR markers

Out of a total of 43 primer pairs tested, 31 and 26 SSR primer pairs resulted in polymorphic banding patterns in each of crosses BC1 and F1, respectively. Out of these primer pairs, 71% (22) and 54% (14) corresponded to single loci, respectively. The remaining primer pairs resulted in multilocus amplification patterns, such that 16% (five primer pairs) and 31% (eight primer pairs) produced double loci, 3% (one primer pair) and 4% (one primer pair) produced triple loci, and 10% (three primer pairs) and 11% (three primer pairs) produced

Table 1 Genetic background of three parents 11307-03, 425, and 422 involved in two crosses analyzed in the species complex *P. mariana* × *P. rubens*

Species-specific markers	Fragment or allele frequencies in allopatric populations		Parental genotypes		
	P. mariana	*P. rubens*	♀11307-03 [♀83×♂425]	♂425	♀422
Dominant RAPD markers[a]					
$OPA06_{900}$	0	1.0	*Aa*	*Aa*	*aa*
$OPA19_{1250}$	0.89	0	*Aa*	*aa*	*AA*
$OPD14_{590}$	1.0	0	*A-*	*A-*	*A-*
$OPE09_{603}$	0	1.0	*aa*	*aa*	*aa*
$OPF17_{800}$	0	1.0	*aa*	*aa*	*aa*
$OPG12_{1000}$	0	1.0	*aa*	*aa*	*aa*
Codominant ESTP markers[b]					
$Sb62_{681}$	0	0.973	*681/681*	*681/681*	*681*
$Sb62_{689}$	0.143	0	–	–	*689*
$Sb62_{691}$	0.976	0.027	–	–	–
$Sb62_{706}$	0.131	0	–	–	–
$Sb70_{404}$	0.024	0	–	–	–
$Sb70_{410}$	0.690	0.014	–	–	*410*
$Sb70_{417}$	0.036	0.986	*417/417*	*417/417*	*417*

[a] From Perron et al. (1995)
[b] From Perron et al. (2000)

Table 2 Number of markers genotyped for each of two crosses in the species complex *P. mariana* × *P. rubens*

Type of marker	Cross BC1			Cross F1			Shared markers
	Segregating 1:1 or 1:1:1:1	Segregating 3:1 or 1:2:1	Total	Segregating 1:1 or 1:1:1:1	Segregating 3:1 or 1:2:1	Total	
AFLPs (% mds[a])	549 (2.7)	313 (5.8)	862 (3.8)	615 (1.5)	248 (8.5)	863 (3.5)	112 (6.3)
SSRs (% mds[a])	38[b] (0)	9[c] (0)	47 (0)	39[d] (0)	6[e] (0)	45 (0)	33 (0)
ESTPs (% mds[a])	30 (0)	11 (0)	41 (0)	46 (0)	10 (0)	56 (0)	40 (0)
Total (% mds[a])	617 (2.4)	333 (5.4)	950 (3.5)	700 (1.3)	264 (8.0)	964 (3.1)	185 (3.8)

[a] % markers with distorted segregation, significant at $P \leq 0.01$/number of tests (*Bonferroni correction*)
[b] Including 13 dominant SSR markers
[c] Including 7 dominant SSR markers
[d] Including 14 dominant SSR markers
[e] Including 5 dominant SSR markers

quadruple loci, respectively for BC1 and F1. In total, 47 and 45 mappable polymorphic SSR markers were obtained for crosses BC1 and F1, respectively (Table 2). Out of these, 20 and 19 markers were dominant, including seven and five markers segregating 3:1, respectively. The remaining codominant markers corresponded to 27 and 26 loci, respectively for BC1 and F1, segregating 1:2:1, 1:1, or 1:1:1:1. A total of 33 SSR markers were shared between the two crosses.

ESTP markers

Out of 113 ESTP markers previously developed and screened for spruces (Guillet-Claude et al. 2004; Pelgas et al. 2004; Table S3, electronic supplementary material), 41 and 56 were polymorphic for the crosses BC1 and F1 (Table 2), respectively. Each of the polymorphic markers was codominant and segregated 1:2:1, 1:1, or 1:1:1:1. Of these, 40 were polymorphic in both crosses (shared markers).

Individual linkage maps

Out of 118 and 85 genotyped individuals for the crosses BC1 and F1, 109 and 80 individuals with less than 10% missing data were retained. Out of a grand total of 950 and 964 markers available for each of crosses BC1 and F1, respectively, 96.5% and 96.9% of the analyzed markers fitted the expected Mendelian ratios ($P \leq 0.01/n$; Table 2). Only three markers, all from cross F1, exhibited a similarity of one with another marker, thus showing the same genotypes for the whole progeny. These markers were removed from the data set to prevent bias in estimating genome lengths. Depending on the parent and cross analyzed, between 349 and 412 Mendelian markers were available for map construction (Table 3). Despite the elimination of markers with similarity of one before analyses, four and two markers were placed exactly at the same position on one linkage group for the parents 425 and 422 of the cross F1, respectively. These additional markers were also removed to facilitate the estimation of expected map lengths for the same reason as explained above.

By counting only once the markers that were assigned in common among two, three, or four individual maps, a total number of 1,137 distinct marker loci were assigned to the four individual linkage maps, including 1,023 AFLP segregating 1:1, 3 RAPD, 55 SSR, and 56 ESTP markers. Linkage groups were generally stable for a LOD of 4.0 and a minimum recombination fraction (θ) of 0.35. A LOD value up to 7.0 was sometimes applied, to obtain comparable groups from one parent to the other. The results obtained using JoinMap 3.0 or CarthaGene were similar (data not shown). For the remaining analyses, only results from JoinMap 3.0 are reported. A major linkage group resulting from the junction between two minor groups was constructed and ordered with a LOD of 2.5 for the female parent 422 in order to be consistent with the results obtained for other parental maps. Depending on the cross, between 8 (2.3%) and 43 (10.7%) markers were found to be unlinked (Table 3).

In the next paragraphs, genome lengths (G_F) will be reported in Haldane distances for each map, followed by the density of markers in brackets.

Cross BC1

The analyses of the female map 11307-03 were conducted with a total of 350 assigned Mendelian markers (Table 3). Out of these, 313 markers were mapped (251 AFLPs, 2 RAPDs, 30 SSRs, and 30 ESTPs), including 16 accessory markers. They resulted in 12 major and 2 minor linkage groups covering 2,223.1 cM, with an average marker spacing of 7.1 cM (Table 3; Fig. 1). The map for the male parent 425 was obtained from a total of 337 assigned Mendelian markers, from which 281 markers (227 AFLPs, 1 RAPD, 28 SSRs, and 25 ESTPs), including four accessory markers, were mapped on 12 major and 5 minor linkage groups representing 1,937.2 cM (6.9 cM average marker spacing; Table 3). Observed genome lengths were 3,314.3 and 2,394.8 cM for the parents 11307-03 and 425, respectively, when unlinked markers were considered. However, observed genome lengths estimated without unlinked markers covered 2,328.7 and 2,113.2 cM, respectively, for each parent. These values correspond to observed map

Table 3 Parameters of individual, reference, and composite linkage maps from two crosses in the species complex *P. mariana* × *P. rubens*

Mapping parameter	Crosses/Parents				Reference map for male parent 425 from BC1 and F1	Composite map
	Cross BC1		Cross F1			
	♀11307-03 [♀83×♂425]	♂425	♀422	♂425		
Total number of available markers[a]	393	349	412	377	1,097	1,729[b]
Number of distorted markers (*Bonferroni correction;* $P \leq 0.01$/number of loci)	15	4	9	4	44	56[b]
Total number of markers without segregation distortion	378	345	403	373	1,053	1,673
Number of assigned marker loci	350	337	360	343	674	1,188
Number of AFLP loci	284	279	290	276	567	1,074
Number of RAPD loci	2	1	0	1	1	3
Number of SSR loci	33	31	31	33	53	55
Number of ESTP loci	31	26	39	33	53	56
Number of positioned marker loci[c] (%)	313 (89.4)	281 (83.4)	326 (90.6)	303 (88.3)	626 (92.9)	1,124 (94.6)
Number of AFLP loci	251	227	260	238	548	1,014
Number of RAPD loci	2	1	0	1	1	3
Number of SSR loci	30	28	29	33	37	53
Number of ESTP loci	30	25	37	31	40	54
Number of positioned accessory marker loci	16	4	16	0	51	97
Number of major linkage groups ($n > 10$ markers)	12	12	12	12	12	12
Number of minor linkage groups ($3 \leq n \leq 10$ markers)	2	5	3	8	1[d]	0
Number of doublets	1	0	2	1	0	0
Number of unlinked markers (%)	28 (7.4)	8 (2.3)	43 (10.7)	30 (8.0)	143 (13.6)	276 (16.5)
Map length G_F, cM (Haldane)	2,223.1	1,937.2	1,849.8	2,123.4	2,106.0	2,319.1
Map length G_F, cM (Kosambi)	1,819.5	1,573.6	1,489.3	1,724.6	1,704.8	1,845.5
Average map density, cM (Haldane)	7.1	6.9	5.7	6.9	3.4	2.1
Average map density, cM (Kosambi)	5.8	5.6	4.6	5.6	2.7	1.6
Average size for major linkage groups, cM (Haldane)	182.4	150.3	143.4	150.4	162.0	193.3
Average size for major linkage groups, cM (Kosambi)	149.0	121.6	115.0	122.1	131.1	153.8
Observed map length G_o, cM (Haldane)	3,314.3	2,394.8	3,549.0	3,468.6	–	–
Observed map length G_o, cM (Haldane) without unlinked markers	2,328.7	2,113.2	2,026.8	2,406.6	–	–
Observed map length G_o, cM (Kosambi)	2,681.3	1,935.0	2,833.3	2,784.8	–	–
Observed map length G_o, cM (Kosambi) without unlinked markers	1,902.9	1,712.6	1,666.3	1,947.8	–	–
Expected map length G_e, cM (Haldane)	3,227.7	2,772.3	2,968.2	3,330.7	–	–
Confidence interval	3,176.4–3,277.4	2,726.8–2,816.3	2,925.4–3,009.8	3,275.5–3,384.1	–	–
Expected map length G_e, cM (Kosambi)	2,549.1	2,189.5	2,347.7	2,695.1	–	–
Confidence interval	2,508.6–2,588.3	2,153.6–2,224.3	2,313.8–2,380.6	2,650.4–2,738.3	–	–
Observed map coverage C_o (G_o/G_e)[e]	>99.9%	86.4%	>99.9%	>99.9%	–	–
Observed map coverage C_o (G_o/G_e)[e] without unlinked markers	72.2%	76.2%	68.3%	72.3%	–	–
Observed map coverage C_o (G_F/G_e)[f]	68.9%	69.9%	62.3%	63.8%	–	–
Expected map coverage C_e	93.5%	94.2%	95.7%	91.9%	–	–

[a] For individual linkage maps, only markers segregating 1:1 or 1:1:1:1 were used. Exceptionally, markers in configuration 3:1 or 1:2:1 for one cross and 1:1 for the other cross were also used (see "Materials and methods" section)

[b] Between the two crosses BC1 and F1, 185 markers were in common, from which seven were distorted (see Table 4)

[c] Unpositioned markers correspond to markers with a recombination frequency higher than 0.35 or having a $-\log_{10}(P)$ value higher than 1 that could affect marker order (see "Materials and methods" section)

[d] For the male reference map, one linkage group with $n = 13$ markers was considered minor because it did not correspond to any major linkage group on the female individual linkage maps (see Fig. 1)

[e] According to the method of Tani et al. (2003)

[f] According to the method of Cervera et al. (2001)

"–": Could not be calculated, due to the merging of data

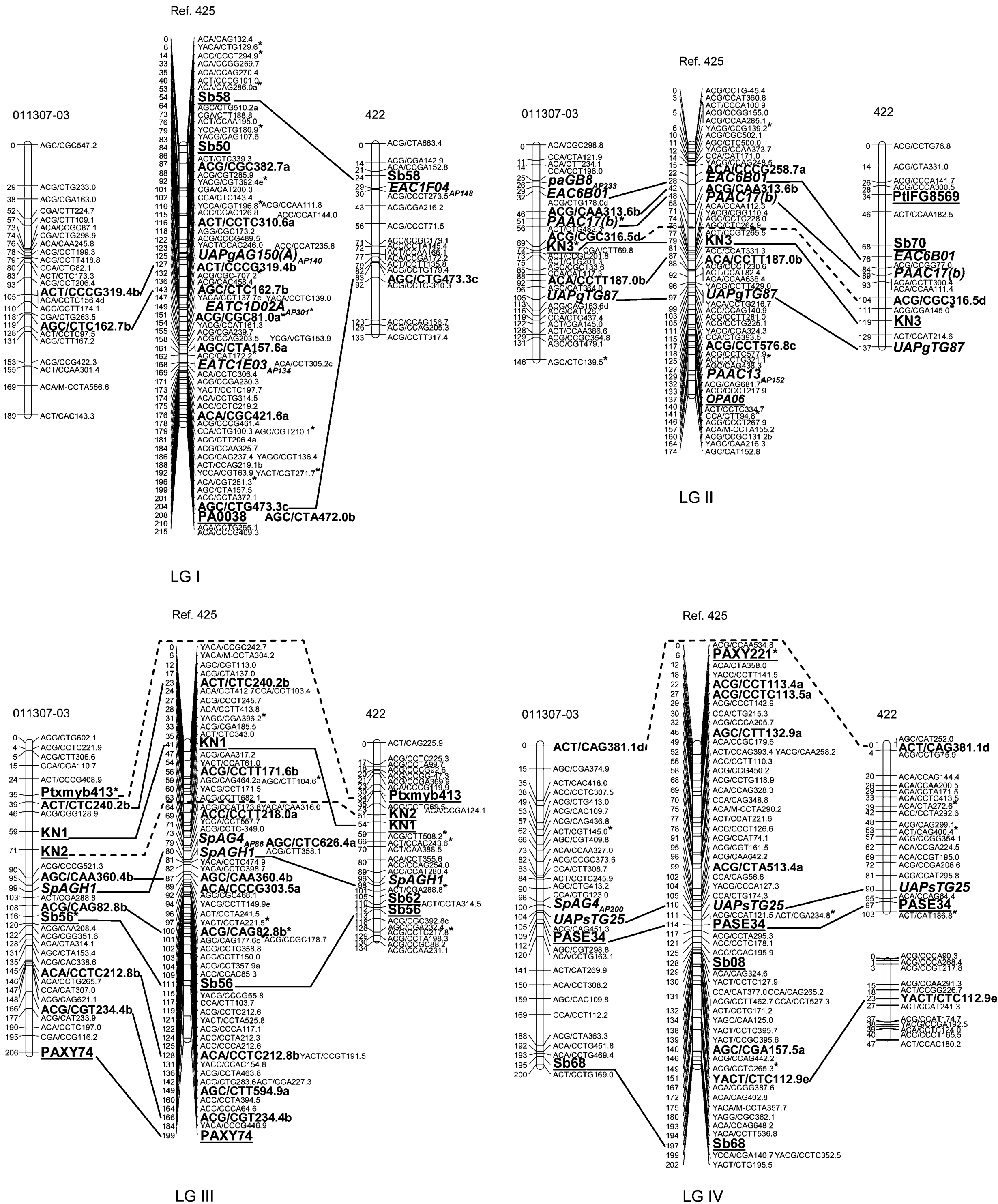

Fig. 1 Parental genetic maps obtained from the two crosses BC1 and F1 for the species complex *P. mariana* × *P. rubens*. On the left, linkage groups from the female individual map 11307-03 [83×425] from the cross BC1; in the middle, linkage groups from the reference map of male parent 425 involved in both crosses BC1 and F1; on the right, linkage groups from the female individual map 422 from the cross F1. Markers in *bold* and *underlined* are ESTPs, markers in *bold* and *italics* are SSRs, markers in *bold* only are homologous AFLPs, which were defined as those AFLPs revealing migration products of same size and similar positioning onto homologous linkage groups from one parent to the other, and markers in *bold*, *underlined*, and *italics* are homologous RAPDs, defined as above for homologous AFLPs. All remaining markers are other AFLPs. Accessory markers (jump value >5) are indicated with an *asterisk*. Two markers positioned side-by-side correspond to the same linkage map bar. Homologous loci used as bridges between female individuals are connected with a *dotted line*, whereas other homologous loci are connected with a solid line. Centimorgan distances (Haldane) are indicated on the left of each linkage group

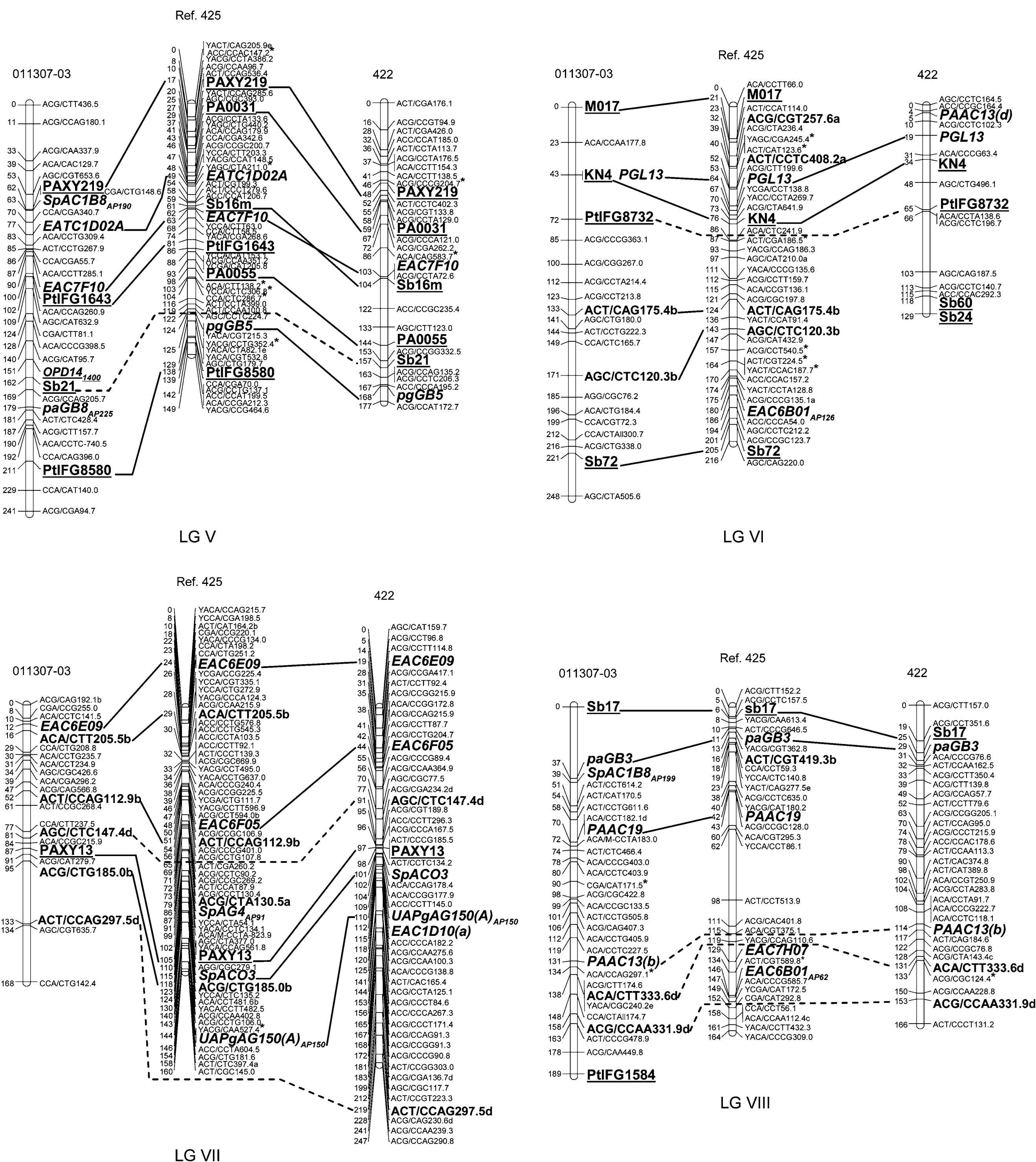

Fig. 1 (Contd.)

coverage values of 72.2 and 76.2% (Table 3) following the method of Tani et al. (2003).

Cross F1

The female linkage map (parent 422) was determined from a total number of 360 assigned Mendelian markers (Table 3). Map construction resulted in 12 major and 3 minor linkage groups based on a total of 326 mapped markers (260 AFLPs, 29 SSRs, and 37 ESTPs), including 16 accessory markers and covering 1,849.8 cM (5.7 cM average marker spacing; Table 3; Fig. 1). For the male linkage map (parent 425), 343 Mendelian markers were assigned (Table 3). Out of these, 303

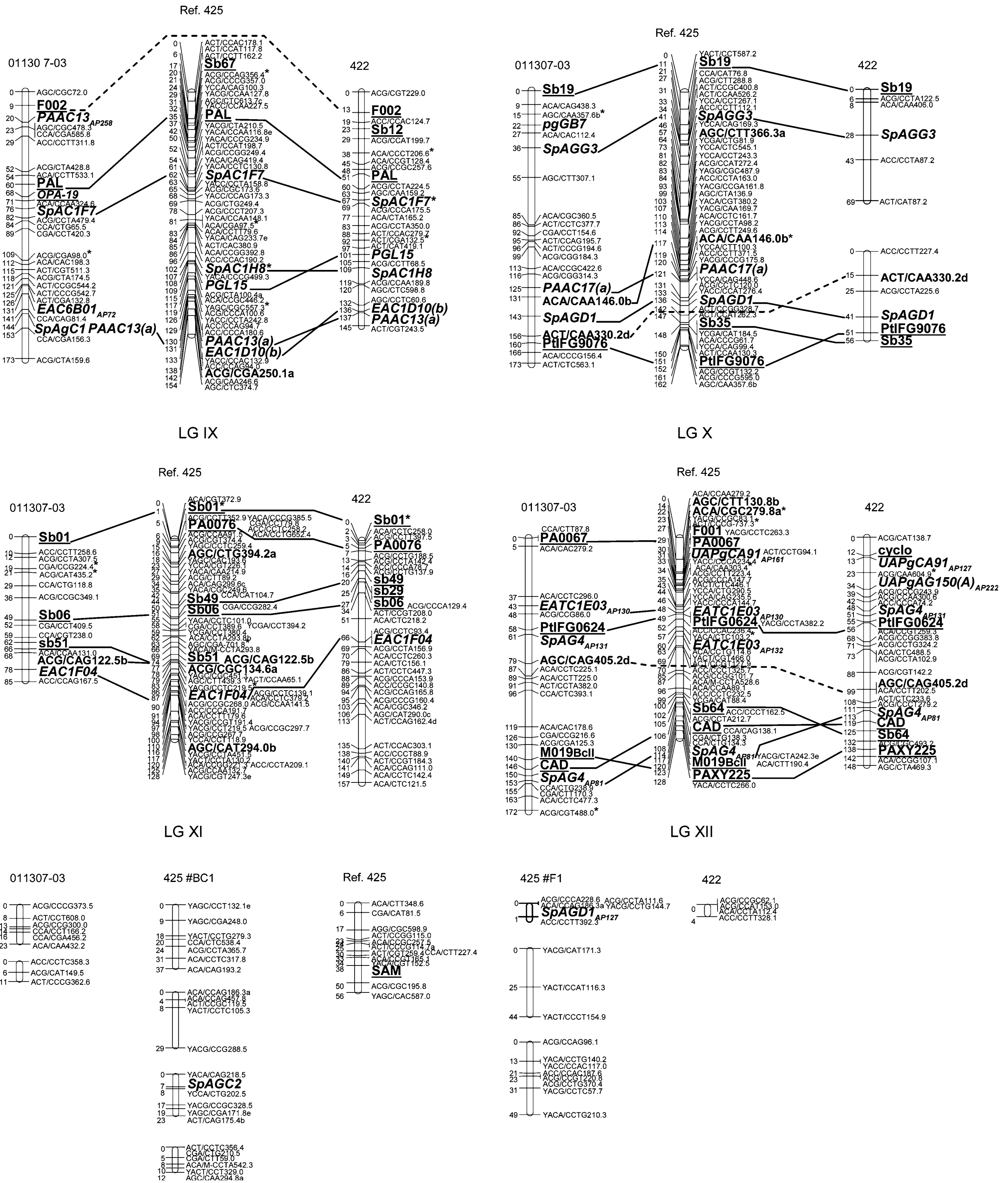

Fig. 1 (Contd.)

markers (238 AFLPs, 1 RAPD, 33 SSRs, and 31 ESTPs) were mapped and distributed over 12 major and 8 minor linkage groups covering 2,123.4 cM (6.9 cM average marker spacing; Table 3). When all markers were considered, lengths of observed maps for parents 422 and 425 were 3,549.0 and 3,468.6 cM, respectively. However, lengths of observed maps calculated without unlinked markers were 2,026.8 and 2,406.6 cM, respectively, for

each parent, which correspond to observed map coverage values of 68.3 and 72.3% (Table 3) following the method of Tani et al. (2003).

Genome lengths and map coverages

Observed genome lengths for each individual linkage maps were in the same range, except for the parent 425 of BC1 for which the length was smaller (Table 3). This tendency was also observed with the map coverage values estimated by taking into account unlinked markers, and thus varying from 86.4% to 99.9% (method of Tani et al. 2003). However, when considering the observed genome lengths estimated without unlinked markers, no significant difference was observed between the parents of each cross. The mean values of observed map coverage were 72.3% when estimated with the method of Tani et al. (2003) without considering unlinked markers, and 66.2% when estimated following the method of Cervera et al. (2001) based on G_F estimate. On average, for the four individual maps, expected map coverage values were 22% and 28% higher than observed map coverage values estimated following the methods of Tani et al. (2003) and Cervera et al. (2001), respectively (Table 3).

Reference linkage map of male parent 425

The marker grouping of the reference linkage map for the male parent 425 was conducted with 674 assigned Mendelian markers (including markers 3:1) obtained from both paternal data sets (Table 3). Before merging homologous linkage groups, the two data sets for the parent 425 were composed of 613 and 593 markers from the crosses BC1 and F1, respectively. Out of these two data sets, 34 and 131 unlinked markers were removed from the analysis, including 28 and 101 markers segregating 3:1. Before marker ordering, 22 heterogeneous linkage pairs of loci were excluded from analyses, out of a total of 324 common linkage pairs between both 425 individual linkage maps, because they showed significant differences ($P < 0.01$) in their recombination frequencies. After integration of each homologous linkage group, a total of 626 markers (548 AFLPs, 1 RAPD, 37 SSRs, and 40 ESTPs) were positioned onto the reference map (Table 3). This reference linkage map consisted of 12 major linkage groups comparable with linkage groups of individual maps of both female parents, and one additional minor linkage group not comparable with female individual maps (Fig. 1). The reference map covered 2,106.0 cM with an average marker density of 3.4 cM (Table 3). The length of linkage groups varied from 55.8 to 216.4 cM (Fig. 1).

Composite linkage map of the species complex *P. mariana* × *P. rubens*

Out of a total of 1,673 Mendelian markers combined from the two data sets used to construct the two ‘‘sub-composite’’ maps, 276 markers were considered as unlinked, of which 51% were AFLPs segregating 3:1. They were simply removed from the analysis. After merging homologous linkage groups of each ‘‘sub-composite’’ map obtained for the two crosses, 1,188 assigned markers were distributed among the composite linkage groups, and 1,124 were positioned (Table 3). However, before marker ordering on each composite linkage group, a test of heterogeneity of recombination frequencies was conducted between 533 homologous pairs of loci merged from both ‘‘sub-composite’’ maps. Out of these homologous pairs of loci, 451 showed no significant difference between their recombination frequency estimates ($P > 0.01$). For the remaining 82 homologous pairs of loci, recombination frequency estimates were significantly heterogeneous, so they were excluded from analyses to avoid erroneous marker positioning. Out of the 1,124 markers (1,014 AFLPs, 3 RAPDs, 53 SSRs, and 54 ESTPs) ordered onto the composite map, 128 markers (52 AFLPs, 1 RAPD, 33 SSRs, and 42 ESTPs) were homologous between at least two parents. These markers were assembled in 12 linkage groups covering 2,319.1 cM (Haldane; Table 3), with a length for linkage groups varying from 157.8 to 228.4 cM (Fig. 2). A total of 107 anchor markers (53 SSRs and 54 ESTPs), useful in comparative mapping studies between species, were placed onto the composite map, for an average of nine markers per linkage group. In comparison to the number of anchor markers derived from only one cross, the use of two crosses resulted in an increase of 19 (22%) or 22 (26%) additional anchor markers positioned on the composite map, if the additional cross was BC1 or F1, respectively, for an average increase of 24%. The average marker density was of 2.1 cM, more than twofold the marker density obtained for individual linkage maps, which varied from 5.7 to 7.1 cM (Table 3).

Synteny and macro-colinearity

Comparing reference and individual linkage maps

The alignment between each linkage group of the two female maps (11307-03 and 422) with their homologous linkage groups from the reference map of male parent 425 could be conducted with 92 homologous marker loci: 26 AFLPs of same size, 28 SSRs, and 38 ESTPs (Fig. 1). The homologous markers of the female individual linkage maps represented 23.3% and 19.9% of all type of markers mapped onto parents 11307-03 and 422, respectively. For the reference linkage map constructed for male parent 425, these homologous markers represented 12.3% of all placed markers.

Between the reference linkage map of the male parent 425 and both individual linkage maps of female parents 11307-03 and 422, 59 (16 AFLPs, 19 SSRs, and 24 ESTPs) and 48 (2 AFLPs, 22 SSRs, and 24 ESTPs) homologous markers were shared, respectively. When considering only the two female individual linkage

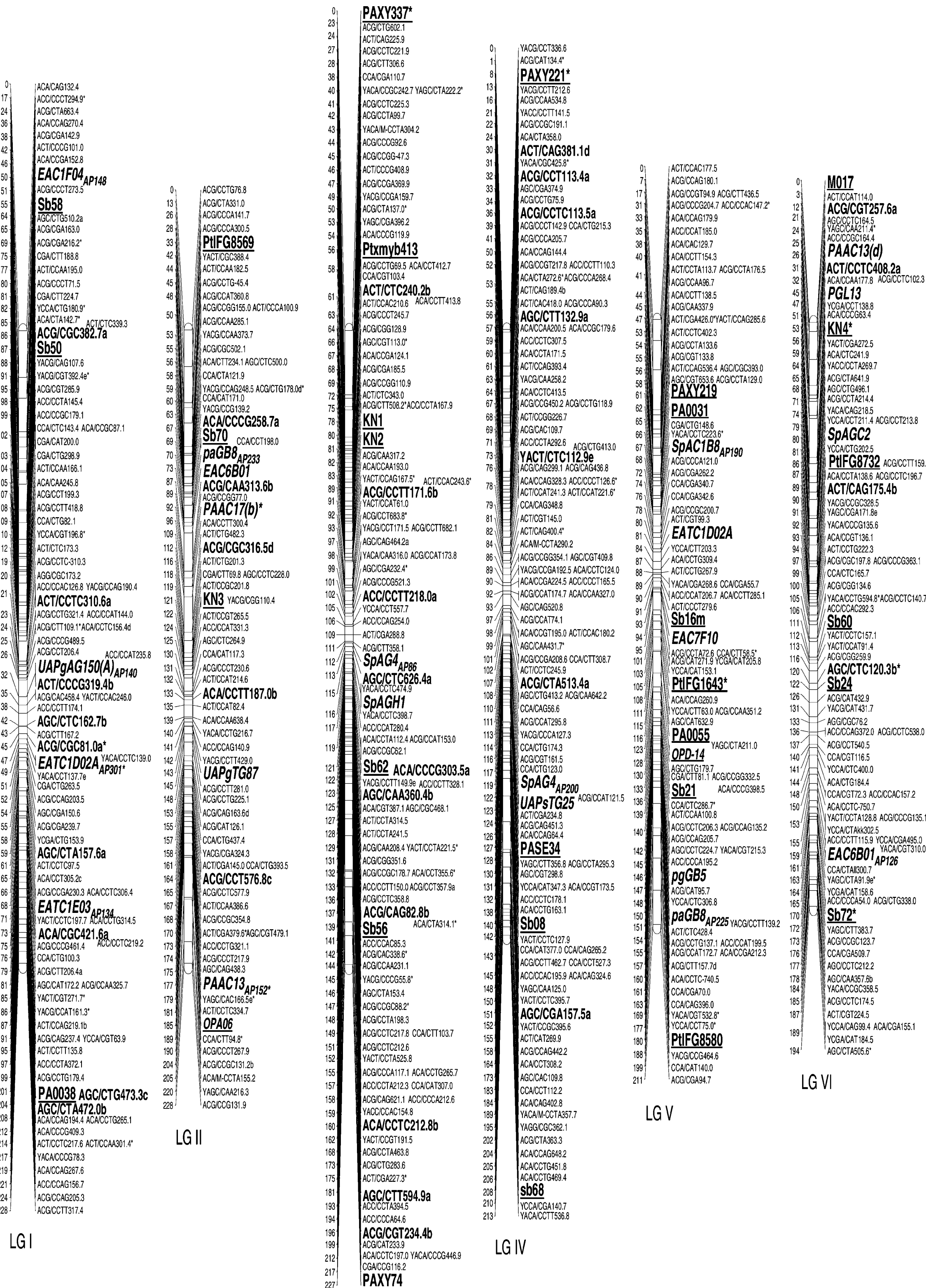

Fig. 2 Composite linkage map for the species complex *P. mariana* × *P. rubens*. The composite map was obtained by assembly of data sets of both crosses BC1 and F1, and using JoinMap 3.0 (Stam 1993; Van Ooijen and Voorrips 2001). Genetic distances are on the left of each linkage group (Haldane). Markers in *bold* and *underlined* are ESTPs, markers in *bold* and *italics* are SSRs, markers in *bold* only are homologous AFLPs, and markers in *bold*, *underlined*, and *italics* are homologous RAPDs. All remaining markers are other AFLPs. Accessory markers (jump value >5) are indicated with an *asterisk*. Two markers positioned side-by-side correspond to the same linkage map bar

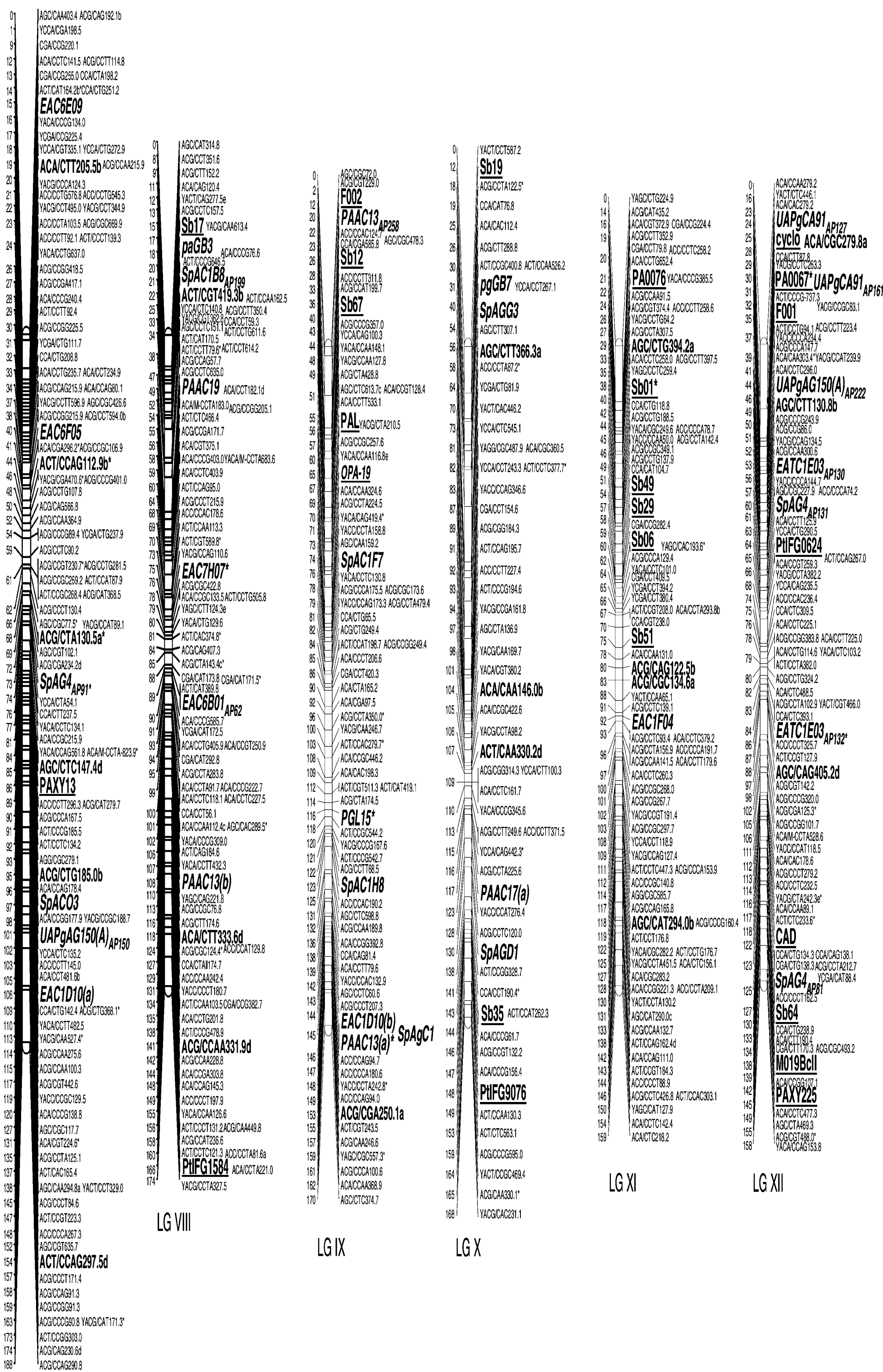

Fig. 2 (Contd.)

maps, 45 homologous markers (8 AFLPs, 17 SSRs, and 20 ESTPs) were in common. Synteny was well conserved among the three linkage maps, both female individual maps, and the male reference linkage map (Fig. 1), since all homologous markers were placed onto the same homologous linkage groups. Macro-colinearity was also

well conserved among homologous linkage groups. Indeed, marker order was the same for 56 out of 59 (95%) homologous markers between the reference linkage map of male parent 425 and the linkage map of female parent 11307-03, and for 42 out of 48 (88%) homologous markers between the male reference linkage map and the linkage map of female parent 422. Between the two female linkage maps, marker order was the same for 41 out of 45 (91%) homologous markers.

Minor changes in marker order involved three AFLP (3:1), nine SSR, and five ESTP markers, which were implicated in various small inversions within linkage groups III, V, IX, X, and XII (Fig. 1). Overall, seven homologous linkage groups had the same marker order (LG I, II, IV, VI, VII, VIII, and XI) between male reference and the two female individual linkage maps. When small inversions involved only codominant markers, the localization of the marker segregating 1:1:1:1, which were more informative compared to the other implied markers segregating 1:1 or 1:2:1, was considered as the most certain. In one case, the inversion was presumably the result of the doubtful positioning of one marker, considered as an accessory marker on the reference map. On average, 90% of the 92 homologous markers (AFLPs, SSRs, and ESTPs) shared between the male reference and the two female individual linkage maps were mapped at similar positions, such that 92%, 86%, and 92% of markers were in the same order, respectively. To obtain the best representation of the genome before assembling the composite map, the marker order on the male reference map was considered as the most certain.

Comparing composite versus male reference and female individual linkage maps

Comparisons between the reference and the composite maps were conducted with the help of the markers segregating simultaneously for the male parent 425 of both crosses. No difference was observed in marker synteny for the 87 comparable homologous markers (32 AFLPs of same size, 1 RAPD of same size, 26 SSRs, and 28 ESTPs). Of these, 82 (94%) were in the same order between the reference and the composite maps. Between the composite and both individual linkage maps of female parents 11307-03 and 422, no difference in marker synteny was observed for the 74 and 63 comparable homologous markers, respectively. Of these, 95% and 92% were in the same order, respectively. Differences in marker order involved inversions between markers 1 or 2 cM apart and were similar to those observed between the male reference and female individual linkage maps (see above). When considering only SSRs and ESTPs, 107 of these anchor markers were placed onto the composite map and their positioning was nearly in the same order as that observed on the male reference or on the two female individual maps. Respectively, 98%, 97%, and 91% of anchor markers were colinear between the composite map and the male reference map and the two female individual maps 11307-03 and 422.

Marker distribution

Heterogeneity of marker distribution

G-tests for goodness-of-fit applied to assess the distribution of markers revealed a relatively homogeneous distribution among linkage groups of each individual (Fig. S1, electronic supplementary material) and composite maps (the detailed results are available from the journal web site as electronic supplementary material).

Coefficients of dispersion

When considering all markers together and a sliding window of 10 cM, values of the coefficient of dispersion (*CD*) varied between 0.76 and 0.96 for each of the four individual linkage maps. As such, these values below 1.0 do not indicate the presence of clusters of markers but rather that the markers are well dispersed over the whole genome. For the composite map, marker density was 2.1 cM, thus windows of 3 cM were considered, and the distribution of markers was found to be statistically homogeneous with a *CD* value of 1.1. Similar results were also obtained for a composite linkage map derived from one cross in *P. abies* (Acheré et al. 2004).

Discussion

Genetic background of parental trees

In this study, a large number of AFLPs segregating 3:1 was observed, indicating a high level of heterozygosity in the genome of the parents. This high level of heterozygosity could be explained at least partly by the mixed genetic background of the parents, as detected by markers specific to *P. mariana* or *P. rubens*. Such an observation is in agreement with recent studies where the use of interspecific backcross pedigrees resulted in a large number of polymorphic AFLP markers (Myburg et al. 2003; Yin et al. 2004). In our study, the number of AFLP markers segregating 3:1 was on average 32.5% for the two crosses investigated (respectively, 36.3 and 28.7% for BC1 and F1). Such a proportion of 3:1 AFLP markers is about three times as high as the levels observed in *P. abies* (9.9%, Acheré et al. 2004) and *P. glauca* (14.6%, Pelgas et al., in preparation). The proportion of 3:1 AFLP markers obtained in the present study was also much higher than the proportion expected by chance alone when using dominant anonymous markers. Indeed, the probability to get two heterozygous individuals for a same locus would be 12.3%, on average, in allopatric populations of *P. mariana* (Isabel et al. 1995), which is in the range of values

reported for *P. abies* and *P. glauca*. Therefore, it is likely that the introgressive nature of the parents involved has increased their genome-wide heterozygosity. Consequently, about twice as many primer combinations had to be used to obtain approximately the same number of AFLP markers segregating 1:1 than in *P. abies* or *P. glauca* mapping studies (Acheré et al. 2004; Pelgas et al. in preparation).

Marker transferability

The transferability of AFLP fragments between crosses has already been documented, indicating that a certain number of AFLPs could be used as homologous bridging markers across populations (Qi et al. 1998; Waugh et al. 1997; Rouppe van der Voort et al. 1997; Vuylsteke et al. 1999). However, the comparison of conspecific maps using AFLPs appears to be difficult: though a large number of AFLP markers were obtained in our study, only 112 markers (13.0%) were polymorphic in both crosses. Of these, 83 (9.6% of the total number of AFLPs) could be mapped in both crosses, since they were neither distorted nor unlinked. A similar proportion of polymorphic AFLP markers shared between two crosses (10.2%) have been reported in barley (Waugh et al. 1997). In our study, out of the 83 AFLP markers that could be mapped in both crosses, 52 markers were assigned at the same position on 11 homologous linkage groups and could therefore be considered as homologous marker loci. These AFLP markers are useful to help bridging individual linkage maps within one species and estimate composite linkage maps. However, due to possible homoplasy, they are not appropriate to construct highly reliable composite linkage maps and should not be considered as a primary source of transferable anchor markers between species or genera, as was recommended by Remington et al. (1999) and Lerceteau et al. (2001).

In comparative mapping studies, numerous sequence-based markers such as SSR and ESTP markers are necessary to anchor maps adequately. However, in spite of large EST sequencing efforts in conifers, these markers are still not available in large numbers (e.g. Acheré et al. 2004; Krutovsky et al. 2004). Because these markers are usually locus-specific, multi-allelic and highly informative, they can be mapped in several crosses and across species, and act as useful bridges. They are not always transferable from one species to the next, although the rate of transfer is very high for ESTPs, at least between congeneric species. The rate of transfer includes gene amplification and the ability to detect Mendelian polymorphism in the recipient species. With respect to these two factors, the rate of transfer of ESTPs between spruce species has been reported to be about 90% when several methods for detecting polymorphism are used (Perry and Bousquet 1998b; Pelgas et al. 2004). As for SSRs, their rate of transfer between congeneric species appears to be lower. For instance, out of the 43 SSR primer pairs tested in this study that were previously developed for various *Picea* taxa, about 24% resulted in the amplification of more loci than expected, even when the annealing temperature was adjusted. This result is quite similar to that obtained in mapping of *P. abies* where 36% of SSR primer pairs resulted in the amplification of more than one locus (Pfeiffer et al. 1997). The amplification of more SSR loci than expected is likely a consequence of the presence of duplicated loci in the large conifer genome (Karhu et al. 2000). Also, the presence of null alleles as observed for some SSR loci in our study could indicate primer mismatch problems, which should lower the transferability of SSR primer pairs between spruce species, as also observed in *Pinus* (Devey et al. 1999). New SSR markers developed from EST libraries are likely to show increased transferability across taxa (Jany et al. 2003; Chagné et al. 2004; Liewlaksaneeyanawin et al. 2004; Rungis et al. 2004).

Segregation distortion and unlinked markers

In our study, all linkage maps were estimated using only Mendelian markers, as we removed markers that were significantly distorted in a Bonferroni-corrected chi-square test. After correction, less than 4% of markers were significantly distorted at $P < 0.01$ for individual and composite linkage maps (Table 3). Without correction for multiple testing, 12% of the markers would have been scored as distorted, which is a value slightly higher than that previously reported in *P. sylvestris* (about 9%; Yin et al. 2003), but similar to the results obtained with *Pinus palustris* and *P. elliottii* (12%; Kubisiak et al. 1995) and *C. japonica* (about 12%; Tani et al. 2003). Often, distorted markers are integrated in linkage maps although they can weaken map structure, especially when they are aggregated (e.g. Kubisiak et al. 1995). In our study, the 8% of markers declared significantly distorted with the uncorrected chi-square test (the difference between 12 and 4%) were randomly distributed over the three parental maps and the composite map, and often they were reported as accessory markers. Among these markers, the percentage of unlinked loci was not higher than that for markers with no significant segregation distortion. Thus, unlinked markers were apparently not the result of segregation distortion.

A large number of unlinked markers was observed for the cross F1. Because of the mixed genetic background detected in the parents, one explanation could be that some chromosomal regions showing less homology have interfered with meiotic recombination events (Paterson et al. 1990; Orr 1996; Roeder 1997), thus preventing the production of recombinant gametic classes (Hanson 1959; Rieseberg et al. 1995; Livingstone and Rieseberg 2003). On the other hand, the majority of unlinked markers, 80 and 51%, were represented by AFLPs segregating 3:1 during the construction of the reference and the composite maps, respectively. Indeed, some AFLP markers segregating 3:1 were not informa-

tive because they did not allow distinguishing between recombinant and non-recombinant gametic classes within each cross. The presence of these numerous unlinked markers segregating 3:1 may be directly related to the high heterozygosity of the parents of both crosses.

Genome lengths and marker density

The high levels of heterozygosity observed for the parents of both crosses might have resulted in an overestimation of observed map lengths, as previously observed in pea by Knox and Ellis (2002). However, the average estimate of observed map length without unlinked markers [G_o, 1,807 cM (K), Kosambi] and the average map length [G_F, 1,652 cM (K)] of *P. mariana* × *P. rubens* individual linkage maps were slightly smaller than those estimated in other mapping studies in the genus *Picea*, such as for *P. glauca* [2,033 cM (K), Gosselin et al. 2002] and *P. abies* [1,856 cM (K), Acheré et al. 2004]. Between the composite linkage map estimated in this study for *P. mariana* × *P. rubens* from two crosses and that estimated for *P. abies* from a single cross (Acheré et al. 2004), a comparable difference was notable in map length (G_F) with estimated values of 1,846 cM (K; Table 3) and 2,035 cM (K), respectively. Several factors might account for these differences in genome length, such as variations in recombination rate among the parents (Liu 1998) or different conditions used to build linkage maps, such as the number and types of markers, the number of progeny, or the algorithms used to estimate map distances (Plomion and O'Malley 1996; Liu 1998; Gosselin et al. 2002). In the present study, the use of two crosses and the large number of markers used have likely increased the accuracy of the estimates of linkage between markers, especially for the composite map, resulting in smaller map lengths. Regarding the expected map length (G_e) estimated for each individual linkage map, the results obtained for *P. mariana* × *P. rubens* varied from 2,190 to 2,695 cM (K), in agreement with results from previous mapping studies in *Picea*, where a similar range was observed for *P. glauca* [2,363–2,706 cM (K), Gosselin et al. 2002] and where slightly larger values were noted for *P. abies* [2,793–2,886 cM (K), Paglia et al. 1998].

High values of observed map length (G_o) were obtained for three of the four individual linkage maps. This pattern may result from the numerous unlinked markers observed for these three maps, as was reported for *P. palustris* (Nelson et al. 1994). In our study, when the unlinked markers were not considered for the calculation of the map coverage values with the method of Tani et al. (2003), the observed values of map coverage (C_o) were similar among the four individual linkage maps and comparable to the values obtained with the method of Cervera et al. (2001) (Table 3), thus indicating the estimation bias incurred by unlinked markers. The difference obtained between the expected and the observed map coverage values may result from a lack of markers in some regions of the genome. Thus, additional markers may increase map coverage and allow the integration of the minor and major linkage groups obtained for each of the four individual linkage maps. Indeed, for each of these, more linkage groups were obtained than the haploid number of chromosomes ($n = 12$) for *Picea* (Table 3), which indicates a non-saturation of the genome. The assembly of the composite map allowed the integration of the minor linkage groups of individual maps to major groups, resulting in 12 major composite linkage groups and no minor groups (Table 3). The length of the composite map, which presented a higher value than that of individual linkage maps, indicates that some regions of the genome were probably covered by markers positioned onto only one individual map. Such a trend was also observed in *Helianthus* (Gentzbittel et al. 1995).

Marker density was three times as high for the composite map as for the individual linkage maps (Table 3). This value of marker density, one marker every 1.6 cM (K), is among the highest ever achieved in conifers. In other conifer composite maps constructed with at least one pedigree, the average distance between two markers was 2.6 cM in *P. abies* (Acheré et al. 2004) and 3.1 cM and more in other conifers (e.g. Sewell et al. 1999; Tani et al. 2003). The marker density of the *P. mariana* × *P. rubens* composite map obtained here is comparable to that obtained in a few studies on barley, for which average marker densities between 1.2 and 1.4 cM were achieved (Qi et al. 1996; Hori et al. 2003). Dense linkage maps are useful for more precise gene localization and colocalization with QTLs, which are themselves useful for the identification and/or validation of candidate genes, or for the integration of genetic and physical maps (Causse et al. 2004; Ma et al. 2004). Moreover, with dense linkage maps, the analysis of genome structure between distant taxa is likely to be more accurate (Salse et al. 2002).

Marker ordering

The reference linkage map for the common male parent of the two crosses was built to verify the consistency of marker ordering between both mapping populations and to obtain the best representation of the genome before assembling the composite map. The merging of both individual linkage maps of the common male parent 425 allowed the integration of data derived from two independent sets of meiosis events into a single reference linkage map. This integration could be conducted because of the colinearity of markers and the homogeneous recombination rates between homologous pairs of loci of both individual male linkage maps (few cases of heterogeneous recombination rates were noted). The integration of locus pairs of homologous linkage groups increases the confidence level in marker positioning, because twice as many recombination informations are provided for the same individual from two crosses than

from only one cross (Kowalski et al. 1994). For example, 98% of anchor markers (SSRs and ESTPs) shared between both individual maps of male parent 425 were positioned in similar order onto the reference map of the same parent. In several other cases, such a mapping approach with a common parent between two crosses has been an essential first step to the generation of reliable composite maps at the intraspecific level, allowing further genomic comparisons within and between species (e.g. Beavis and Grant 1991; Kowalski et al. 1994; Lan et al. 2000; Pelgas et al. in preparation). Thus, in our study, the reference map provided a foundation for the marker order, on which the macro-colinearity of the composite map could be efficiently validated.

Because synteny was well conserved between homologous linkage groups of the male reference and the two female individual linkage maps, macro-colinearity could be analyzed between homologous linkage groups. Inversions in marker order occurred for proximal markers and were most often caused by changes in the distribution of less informative dominant AFLP markers segregating 3:1 or codominant ESTP or SSR markers segregating 1:2:1. The lack of information of AFLP markers segregating 3:1 has also been observed by others (Maliepaard et al. 1997). Comparison of macro-colinearity between parental and composite linkage maps revealed that homologous markers were approximately positioned in the same order (on average 94%). Our results are similar to those obtained for two other conifer species, *P. taeda* and *C. japonica*, where 93% and 90% of homologous markers were found colinear between individual and composite linkage maps respectively (Sewell et al. 1999; Tani et al. 2003). In all cases, the discrepancies observed in marker order were small and most likely the result of analytical artifacts, as reported in *Pinus* and *Eucalyptus* (Sewell et al. 1999; Brondani et al. 2002), rather than the consequence of true chromosomal inversions.

Informativeness of the composite linkage map

Combining data from several crosses to estimate composite maps is considered as the most effective way to obtain an accurate and representative view of the genome (e.g. Gentzbittel et al. 1995; Sewell et al. 1999). Compared to model agronomic crop species, the estimation of composite maps relies frequently on a single cross for most tree species (e.g. Jermstad et al. 1998; Acheré et al. 2004). In our study, the primary goal for assembling a composite map was to estimate the relative positioning of as many anchor markers as possible onto a single map. To this end, the use of multiple crosses provides several advantages, such as a larger number of anchor loci that can be positioned on an assembled composite map (Hauge et al. 1993; Causse et al. 1996; Tani et al. 2003). In our study, about 60% of anchor markers (SSRs and ESTPs) deriving from only one cross will be useful for comparative mapping at the interspecific level with *P. glauca* (Pelgas et al. in preparation). However, information deriving from the combination of two crosses contributed to an increase of about 24% in the number of anchor markers (SSRs and ESTPs) positioned on the composite map, corresponding to an average of 21 additional markers in comparison with the use of a single cross. Out of these additional anchor markers, it was observed that about 50% will be useful for interspecific comparative studies with *P. glauca* (Pelgas et al. in preparation). Other studies involving two crosses with one parent in common and implicating anchor markers (mostly RFLPs) revealed that between 48% (in maize, Beavis and Grant 1991) and 77% (in *Arabidopsis*, Kowalski et al. 1994) additional anchor markers could be positioned by the use of a second cross. Differences between studies in the efficiency of using a second cross to position anchor markers are likely to be caused by the inherent genome-wide heterozygosity of the taxon under study, the type of anchor markers and the sensitivity of the method for detecting DNA polymorphisms. The efficiency of using a second cross will in general diminish, as a larger proportion of loci are found polymorphic within a single cross. For instance, in our study, we used DGGE to reveal ESTPs (Pelgas et al. 2004), which is more sensitive than agarose gel electrophoresis or CAPS (cleaved amplified polymorphic sites), thus reducing the probability that a marker will be monomorphic in a single cross. A similar argument could be made for hypervariable SSR loci compared to other types of markers.

The composite map obtained for *P. mariana* × *P. rubens* contains more than 100 anchor markers, which should ensure more reliable studies of synteny and macro-colinearity at the interspecific level. The first published composite maps in other conifers such as *P. taeda* and *C. japonica* had smaller numbers of anchor markers (Sewell et al. 1999; Tani et al. 2003) but additional anchor markers have been mapped since then (e.g. Brown et al. 2001; Temesgen et al. 2001). Recently, a similar number of orthologous gene-specific markers was used to study the conservation of genome macrostructure among several species (e.g. *Medicago truncatula*, *Lotus japonicus* and *Pisum sativum*) pertaining to two major tribes of legumes, Phaseoleae, and Galegeae (Choi et al. 2004). The trend is toward even larger numbers of anchor markers, as more than 600 presumed orthologous genes have recently been used to compare the genome architecture of rice and maize (Salse et al. 2004). Such high numbers appear necessary to identify adequately homoeologous segments in comparative mapping studies involving distant taxa (see also below).

Inter-generic comparative mapping in the Pinaceae

The composite map of *P. mariana* × *P. rubens* presented in this paper should serve as a useful starting point for comparative mapping studies among widely divergent

Picea species, including *P. glauca* and *P. abies* (Pelgas et al. in preparation), and with other taxa from the Pinaceae. With the anchor markers already positioned on the composite linkage map of *P. mariana* × *P. rubens*, a preliminary comparison could be made between the present composite map and linkage maps obtained for *Pinus* spp. and *P. menziesii* (Brown et al. 2001; Chagné et al. 2003; Komulainen et al. 2003; Krutovsky et al. 2004). The current comparison involves 16 ESTP anchor markers found in common between at least two of the three genera (Table 4). A number of these anchor markers correspond to genes implicated in the lignin pathway (*CAD, PAL*, and *PtIFG8732*) or in abiotic stress response (*PtIFG8569, PAXY13, PtIFG0624, Sb64, PtIFG1643*, and *PtIFG8580*). With these markers, four homoeologous linkage groups with at least two common anchor markers could be identified, corresponding to LG XI, IX, XII, and V in *P. mariana* × *P. rubens* (this study), LG 3, 6, 9, and 10 in *Pinus* spp., and LG 9 and 10 in *P. menziesii* (Table 4).

Apparent differences in synteny were observed between the homoeologous linkage groups of *Picea* and *Pinus*, with two instances of putative inter-chromosomal gene translocation. First, the anchor markers *PtIFG8569* and *PAXY13* positioned onto LG 2 in *Pinus* spp. were located onto two different linkage groups (LG II and VII) in *P. mariana* × *P. rubens* (Table 4). Second, when considering LG 3 of *Pinus* spp., out of three anchor markers (*Sb29*, *Sb49*, and *Sb72*), the two first ones mapped onto the same linkage group (LG XI) in *P. mariana* × *P. rubens*, whereas *Sb72* was positioned onto a distinct linkage group (LG VI) (Table 4). To validate these putative cases of inter-chromosomal gene translocation and to rule out paralogy as a possible cause of apparent differential gene positioning, PCR amplifications from haploid megagametophyte DNA of *P. taeda* and *P. mariana* were undertaken for each of these markers, followed by sequencing of the obtained PCR products (see Pelgas et al. 2004 for PCR and sequencing conditions) and comparison of these sequences with *Picea* and *Pinus* EST libraries (http://www.ccgb.umn.edu/cgi-bin/spruce/blastsubform and http://www.ccgb.umn.edu/cgi-bin/biodata2/blastsubform, respectively) and the National Center for Biotechnology Information database (NCBI; http://www.ncbi.nlm.nih.gov/BLAST/). Sequence comparisons indicated that *PAXY13* of *Picea* and *Pinus* that mapped onto two different linkage groups had 93% sequence homology on 213 bp. However, polymorphisms were observed in the chromatogram analyses of the haploid *Pinus* sequence, suggesting that paralogous sequences had been amplified from a same gene family. Comparison of these sequences with EST libraries resulted in the identification of several contigs belonging to the alpha-tubulin gene family. Thus, most likely, *PAXY13* of *Picea* and *Pinus* do not represent orthologous gene loci. Conclusive validation should be sought from phylogenetic analyses of the gene family involved. Sequence analyses also suggested that the differential positioning of *Sb72* between *Picea* and *Pinus* could result from positioning different members of the 60S ribosomal protein L27a gene family (contig#11,497 in *Picea*-CCGB EST library), which is represented by at least three paralogs in *Arabidopsis thaliana* (*At1g70600*, *At1g12960*, and *At1g23290*). Sequence homology between *Sb72* of *Picea* and *Pinus* was 100% but only on a length of 23 nucleotides. For the rest of the sequence, the homology decreased to 30% (337 nucleotides). Thus, spurious inferences of inter-chromosomal translocations may emerge from the mapping of different paralogous genes in different taxa. Recently, paralogous loci in the *knox* gene family of *Picea* and *Pinus* have been shown to be translocated to different chromosomes, following gene duplication (Guillet-Claude et al. 2004). If primers used for gene amplification are designed without sequence evidence from the various paralogs, it is possible that different paralogs are amplified and positioned in different taxa. The problem is exacerbated when different taxa are analyzed in different laboratories with different conditions for PCR. In addition, sequence analyses may not resolve all questionable cases of orthology if the region amplified is well conserved among paralogs.

A presumed difference in genome structure was observed among the four homoeologous linkage groups that could be identified between *Picea, Pinus*, and *Pseudotsuga* (Table 4). The anchor markers *Sb29* and *Sb49*, which were positioned onto two different linkage groups in *P. menziesii* (LG 2 and 13), were located on the same linkage group in *P. mariana* × *P. rubens* (LG XI) and in *Pinus* (LG 3). The sequence chromatograms of haploid megagametophytes indicated that no positions were polymorphic in either species for each of the two loci, suggesting that only one ortholog had been amplified in each species. *Sb29* codes for an ATAF1-like protein (Perry and Bousquet 1998a) and the sequence identity between *Picea* and *Pseudotsuga* was 88% on a length of 493 bp. For *Sb49*, pairwise sequence identities between the three taxa were from 89% to 90% for stretches ranging between 171 bp and 325 bp and all BLASTN analyses pointed to a gene (contig#2225 in *Picea*-CCGB EST library) similar to the rice gene coding for a YGL010w-like protein. Both genes (ATAF1-like and YGL010w-like proteins) do not pertain to sizeable gene families. Although paralogy cannot be entirely ruled out, the sequence comparisons above indicate that the genes amplified in the different taxa are likely to represent orthologous loci. Thus, a chromosomal fission may have occurred in *Pseudotsuga*, leading to the two linkage groups 2 and 13. This chromosomal fission may have played a role in generating the difference in basic chromosome number between *Pseudotsuga* ($n = 13$) and other Pinaceae ($n = 12$). For conclusive evidence on this subject, additional genome-wide comparisons must be undertaken in the Pinaceae family with a much larger number of common anchor markers and robust tests for gene orthology. Because sequence similarity corresponding to gene orthology is variable from one gene family to the next, simple sequence comparisons may

Table 4 Map synteny information and marker sharing in the Pinaceae between *Picea, Pinus, and Pseudotsuga*

Markers in common[a]	Linkage groups		
	P. mariana × *P. rubens*	*Pinus* spp.[b]	*P. menziesii*[c]
Sb58	I	1	–
PtIFG8569	II	2	–
PAXY13	VII	2	–
Sb29	XI	3	13
Sb49	XI	3	2
Sb72	VI	3	–
PtIFG8732	VI	5 or 6 or 8[d]	8
PtIFG1584	VIII	4	–
Sb12	IX	6	–
PAL	IX	6	–
PtIFG0624	XII	9	–
CAD	XII	9	9
Sb64	XII	–	9
PtIFG1643	V	10	–
PtIFG8580	V	10	10[e]
Sb21	V	–	10

[a] Markers in common among the three genera
[b] Including results previously published for *P. taeda*, *P. pinaster*, *P. sylvestris*, and *P. elliottii* (Brown et al. 2001; Chagné et al. 2003; Komulainen et al. 2003; Krutovsky et al. 2004)
[c] Krutovsky et al. (2004)
[d] Corresponding to loci *PtIFG8732a* on LG 8 for *P. taeda* and *P. sylvestris*, and loci *PtIFG8732b*, and *PtIFG8732c* on LG 5 and 6 for *P. sylvestris*, respectively
[e] Corresponding to locus *pmIFG_128D06a* in *P. menziesii*"–": not positioned

not be enough to guarantee gene orthology when translocations or inversions are inferred from the sole position of single gene markers. Whenever possible, phylogenetic information should be sought for the underlying gene families involved (e.g. Guillet-Claude et al. 2004). Otherwise, translocations and other genome rearrangements should only be identified when they involve several gene loci. Such a requirement imposes large numbers of anchor markers in genome comparisons, much larger than the numbers of anchor markers currently positioned on conifer maps.

Prospects

With little prospect in the near future for the complete sequencing of one or few conifer genomes, the mapping of genes appears essential to further increase the number of anchor markers necessary to compare maps between widely divergent species. The localization of such gene loci will also lead to a better understanding of the structure of the conifer genome and the physical distribution of expressed genes. In this study, numerous anchor markers were positioned onto a *P. mariana* × *P. rubens* composite linkage map, many of them representing candidate genes for various quantitative traits. The mapping of additional genes will provide an essential tool for estimating co-localization with genomic regions involved in quantitative trait variation and for guiding association mapping efforts (Causse et al. 1996; Neale and Savolainen 2004). Moreover, with such a composite map having a mixed genetic background, possibly resulting from many generations of hybridization, it will be possible to localize genes involved in adaptation and speciation (Rieseberg and Buerkle 2002).

Colocalization studies should also be facilitated by combining information from multiple crosses, because the number of mapped genes and the number of segregating phenotypic traits can be maximized in such conditions. Recent developments in this direction include the development of computer programs to estimate the colocalization of QTLs and candidate genes based on information from composite linkage maps (BioMercator, Arcade et al. 2004).

Acknowledgements We thank F. Gagnon, I. Gosselin, M. Lamothe, S. Plante, and S. Senneville (Univ. Laval and Canadian Forest Service) for their help and support in the laboratory, and with mapping and sequence analyses. We also thank R.K. Rutledge (Canadian Forest Service) for the gift of *HAP3a* and *HAP3b* from *P. mariana*. Finally, we thank P. Meirmans (Canadian Forest Service), O. Savolainen (University of Oulu, Finland) and two anonymous reviewers for their constructive comments on previous drafts of this manuscript. Pedigree material was developed by N.I. and J. Beaulieu (Canadian Forest Service) through a Canadian Biotechnology Strategy grant. This research was supported by a National Sciences and Engineering Research Council of Canada grant (Genomics Program) to J.B. and N.I.

References

Acheré V, Faivre-Rampant P, Jeandroz S, Besnard G, Markussen T, Aragones A, Fladung M, Ritter E, Favre JM (2004) A saturated consensus linkage map of *Picea abies* including AFLP, SSR, STS, 5S rDNA and morphological markers. Theor Appl Genet 108:1602–1613
Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, Joets J (2004) BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. Bioinformatics 20:2324–2326
Beavis WD, Grant D (1991) A linkage map based on information from four F2 populations of maize (*Zea mays* L.). Theor Appl Genet 82:636–644

Besnard G, Acheré V, Faivre-Rampant P, Favre JM, Jeandroz S (2003) A set of cross-species amplifying microsatellite markers developed from DNA sequence databanks in *Picea* (Pinaceae). Mol Ecol Notes 3:380–383

Brondani RPV, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. Mol Gen Genet 267:338–347

Brown GR, Kadel EE III, Bassoni DL, Kiehne KL, Temesgen B, van Buijtenen JP, Sewell MM, Marshall KA, Neale DB (2001) Anchored reference loci in loblolly pine (*P. taeda* L.) for integrating pine genomics. Genetics 159:799–809

Bucci G, Kubisiak TL, Nance WL, Menozzi P (1997) A population consensus partial linkage map of *Picea abies* Karst based on RAPD markers. Theor Appl Genet 95:643–654

Cato SA, Corbett GE, Richardson TE (1999) Evaluation of AFLP for genetic mapping in *Pinus radiata* D. Don. Mol Breed 5:275–281

Causse M, Santoni S, Damerval C, Maurice A, Charcosset A, Deatrick J, De Vienne D (1996) A composite map of expressed sequences in maize. Genome 39:418–432

Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M, Rothan C (2004) A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. J Exp Bot 55:1671–1685

Cervera MT, Storme V, Ivens B, Gusmao J, Liu B-H, Hostyn V, van Slycken J, Van Montagu M, Boerjan W (2001) Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. Genetics 158:787–809

Chagné D, Lalanne C, Madur D, Kumar S, Frigerio JM, Krier C, Decroocq S, Savouré A, Bou-Dagher-Kharrat M, Bertocchi E, Brach J, Plomion C (2002) A high density genetic map of maritime pine based on AFLPs. Ann For Sci 59:627–636

Chagné D, Brown G, Lalanne C, Madur D, Pot D, Neale D, Plomion C (2003) Comparative genome and QTL mapping between maritime and loblolly pines. Mol Breed 12:185–195

Chagné D, Chaumeil P, Ramboer A, Collada C, Guevara A, Cervera M-T, Vendramin GG, Garcia V, Frigerio JM, Echt C, Richardson T, Plomion C (2004) Cross species transferability and mapping of genomic and cDNA SSRs in pines. Theor Appl Genet 109:1204–1214

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

Choi HK, Mun JH, Kim DJ, Zhu H, Baek JM, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND, Cook DR (2004) Estimating genome conservation between crop and model legume species. Proc Natl Acad Sci USA 101:15289–15294

Costa P, Pot D, Dubos C, Frigerio J-M, Pionneau C, Bodénès C, Bertocchi E, Cervera M, Remington DL, Plomion C (2000) A genetic map of maritime pine based on AFLP, RAPD and protein markers. Theor Appl Genet 100:39–48

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

Devey MD, Sewell MM, Uren TL, Neale D (1999) Comparative mapping in loblolly pine and radiata pine using RFLP and microsatellite markers. Theor Appl Genet 99:656–662

Devey ME, Carson SD, Nolan MF, Matheson AC, Te Riini C, Hohepa J (2004) QTL associations for density and diameter in *Pinus radiata* and the potential for marker-aided selection. Theor Appl Genet 108:516–524

Gentzbittel L, Vear F, Zhang Y-X, Bervillé A, Nicolas P (1995) Development of a consensus linkage RFLP map of cultivated sunflower (*Helianthus annuus* L.). Theor Appl Genet 90:1079–1086

Gosselin I, Zhou Y, Bousquet J, Isabel N (2002) Megagametophyte-derived linkage maps of white spruce (*Picea glauca*) based on RAPD, SCAR and ESTP markers. Theor Appl Genet 104:987–997

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

Guillet-Claude C, Isabel N, Pelgas B, Bousquet J (2004) The evolutionary implications of *knox-I*gene duplications in conifers: correlated evidence from phylogeny, gene mapping, and analysis of functional divergence. Mol Biol Evol 21:2232–2245

Hackett CA, Broadfoot LB (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity 90:33–38

Haldane JBS (1919) The combination of linkage values, and the calculation of distances between loci of linked factors. J Genet 8:299–309

Hanson WD (1959) The Breakup of initial linkage blocks under selected mating systems. Genetics 44:857–868

Harry DE, Temesgen B, Neale D (1998) Codominant PCR-based markers of *Pinus taeda* developed from mapped cDNA clones. Theor Appl Genet 97:327–336

Hauge BM, Hanley SM, Cartinhour S, Cherry JM, Goodman HM, Koornneef M, Stam P, Chang C, Kempin C, Medrano L, Meyerowitz EM (1993). An integrated genetic/physical map of the *Arabidopsis thaliana* genome. Plant J 3:745–754

Hayashi E, Kondo T, Terada K, Kuramoto N, Goto Y, Okamura M, Kawasaki (2001) Linkage map of Japanese black pine based on AFLP and RAPD markers including markers linked to resistance against the pine needle gall midge. Theor Appl Genet 102:871–875

Hodgetts RB, Aleksiuk MA, Brown A, Clarke C, Macdonald E, Nadeem S, Khasa D (2001) Development of microsatellite markers for white spruce (*Picea glauca*) and related species. Theor Appl Genet 102:1252–1258

Hori K, Kobayashi T, Shimizu A, Sato K, Takeda K, Kawasaki S (2003) Efficient construction of high-density linkage map an dits application to QTL analysis in barley. Theor Appl Genet 107:806–813

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, 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

Jany JL, Bousquet J, Khasa DP (2003) Microsatellite markers for *Hebeloma* species developed from expressed sequence tags in the ectomycorrhizal fungus *Hebeloma cylindrosporum*. Mol Ecol Notes 3:659–661

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

Karhu A, Dieterich JH, Savolainen O (2000) Rapid expansion of microsatellite sequences in pines. Mol Biol Evol 17:259–265

Knox MR, Ellis THN (2002) Excess heterozygosity contributes to genetic map expansion in pea recombinant inbred populations. Genetics 162:861–873

Komulainen P, Brown GR, Mikkonen M, Karhu A, Garcia-Gil MR, O'Malley D, Lee B, Neale DB, Savolainen O (2003) Comparing EST-based genetic maps between *Pinus sylvestris* and *Pinus taeda*. Theor Appl Genet 107:667–678

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

Kowalski S, Lan T, Feldmann K, Paterson A (1994) Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. Genetics 138:499–510

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

Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB (2004) Comparative mapping in the Pinaceae. Genetics 168:447–461

Kubisiak TL, Nelson CD, Nance WL, Stine M (1995) RAPD linkage mapping in a longleaf pine × slash pine F_1 family. Theor Appl Genet 90:1119–1127

Kumar S, Spelman RJ, Garrick DJ, Richardson TE, Lousberg M, Wilcox PL (2000) Multiple marker mapping of wood density loci in an outbred pedigree of radiata pine. Theor Appl Genet 100:926–933

Lan TH, DelMonte TA, Reischmann KP, Hyman J, Kowalski SP, McFerson J, Kresovich S, Paterson AH (2000) An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*. Genome Res 10:776–788

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 genes will it take to span the human genome? Am J Hum Genet 24:842–845

Lerceteau E, Plomion C, Andresson B (2001) AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in *Pinus sylvestris*: a preliminary study. Mol Breed 6:451–458

Liewlaksaneeyanawin C, Ritland CE, El-Kassaby YA, Ritland K (2004) Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. Theor Appl Genet 109:361–369

Liu B-H (1998) Statistical genomics: linkage mapping and QTL analysis. CRC Press, Boca Raton, FL

Livingstone K, Rieseberg L (2003) Chromosomal evolution and speciation: a recombination-based approach. New Physiol 161:107–112

Ma H, Moore PH, Liu Z, Kim MS, Yu Q, Fitch MM, Sekioka T, Paterson AH, Ming R (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Genetics 166:419–436

Maliepaard C, Jansen J, Van Ooijen JW (1997) Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. Genet Res 70:237–250

Marques CM, Araujo JA, Ferreira JG, Whetten R, O'Malley DM, Liu B-H, Sederoff R (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. Theor Appl Genet 96:727–737

Mukai Y, Suyama Y, Tsumura Y, Kawahara T, Yoshimaru H, Hondo T, Tomaru N, Kuramoto N, Murai M (1995) A linkage map for sugi (*Cryptomeria japonica*) based on RFLP, RAPD, and isozyme loci. Theor Appl Genet 90:835–840

Myburg AA, Remington DL, O'Malley DM, Sederoff RR, Whetten RW (2001) High-throughput AFLP analysis using infrared dye-labeled primers and an automated DNA sequencer. Biotech 30:348–357

Myburg AA, Griffin AR, Sederoff RR, Whetten RW (2003) Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. Theor Appl Genet 107:1028–1042

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. Trends Plant Sci 9:325–330

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

Nikaido AM, Ujino T, Iwata H, Yoshimura K, Yoshimura H, Suyama Y, Murai M, Nagasaka K, Tsumura Y (2000) AFLP and CAPS linkage maps of *Cryptomeria Japonica*. Theor Appl Genet 100:825–831

Orr HA (1996) Dobzhansky, Bateson, and the genetics of speciation. Genetics 144:1331–1335

Paglia G, Morgante M (1998) PCR-based multiplex DNA fingerprinting techniques for the analysis of conifer genomes. Mol Breed 4:173–177

Paglia P, 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

Paterson AH, DeVerna JW, Lanini B, Tanksley SD (1990) Fine mapping of Quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. Genetics 124:735–742

Pelgas B, Isabel N, Bousquet J (2004) Efficient screening for expressed sequence tag polymorphisms (ESTPs) by DNA pool sequencing and denaturing gradient gel electrophoresis (DGGE) in spruces. Mol Breed 3:263–279

Perron M, Bousquet J (1997) Natural hybridization between black spruce and red spruce. Mol Ecol 6:725–734

Perron M, Gordon AG, Bousquet J (1995) Species-specific RAPD fingerprints for the closely related *Picea mariana* and *P. rubens*. Theor Appl Genet 91:142–149

Perron M, Perry D, Andalo C, Bousquet J (2000) Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. Proc Natl Acad Sci USA 97:11331–11336

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

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

Plomion C, O'Malley DM (1996) Recombination rate differences for pollen parents and seed parents in pine. Heredity 77:341–350

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

Qi X, Stam P, Lindhout P (1998) Use of locus-specific AFLP markers to construct a high-density molecular map in barley. Theor Appl Genet 96:376–384

Rajora OP, Rahman MH, Dayanandan S, Mosseler A (2001) Isolation, characterization, inheritance and linkage of microsatellite DNA markers in white spruce (*Picea glauca*) and their usefulness in other spruce species. Mol Gen Genet 264:871–882

Remington DL, Whetten RW, Liu BH, O'Malley DM (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theor Appl Genet 98:1279–1292

Rieseberg LH, Buerkle CA (2002) Genetic mapping in hybrid zones. Am Nat 159:S36–S50

Rieseberg LH, Linder CR, Seiler GJ (1995) Chromosomal and genic barriers to introgression in *Helianthus*. Genetics 141:1163–1171

Roeder GS (1997) Meiotic chromosomes: it takes two to tango. Genes Dev 11:2600–2621

Rouppe van der Voort JN, van Zandvoort P, van Eck HJ, Folkertsma RT, Hutten RC, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol Gen Genet 255:438–447

Rungis D, Bérubé Y, Zhang J, Ralph S, Ritland CE, Ellis BE, Douglas C, Bohlmann J, Ritland K (2004) Robust simple sequence repeat markers for spruce (*Picea spp.*) from expressed sequence tags. Theor Appl Genet 109:1283–1294

Salse J, Piégu B, Cooke R, Delseny M (2002) Synteny between *Arabidopsis thaliana* and rice at the genome level: a tool to identify conservation in the ongoing rice genome sequencing project. Nucl Acids Res 30:2316–2328

Salse J, Piégu B, Cooke R, Delseny M (2004) New *in silico* insight into the synteny between rice (*Oriza sativa* L.) and maize (*Zea mays* L.) highlights reshuffling and identifies new duplications in the rice genome. Plant J 38:396–409

Schiex T, Gaspin C (1997) CARTHAGENE: constructing and joining maximum likelihood genetic maps. Proc Int Conf Intell Syst Mol Biol 5:258–67

Scotti I, Magni F, Fink R, Powell W, Binelli G, Hedley PE (2000) Microsatellite repeats are not randomly distributed within Norway spruce (*Picea abies* K.) expressed sequences. Genome 43:41–46

Scotti I, Magni F, Paglia G, Morgante M (2002a) Efficient development of dinucleotide microsatellite markers in Norway spruce (*Picea abies* Karst.) through dot-blot selection. Theor Appl Genet 104:1035–1041

Scotti I, Magni F, Paglia G, Morgante M (2002b) Trinucleotide microsatellites in Norway spruce (*Picea abies* K.): Their features and the development of molecular markers. Theor Appl Genet 106:40–50

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

Shepherd M, Cross M, Dieters MJ, Henry R (2003) Genetic maps for *Pinus elliottii* var. *elliottii* and *P. caribaea* var. *hondurensis* using AFLP and microsatellite markers. Theor Appl Genet 106:1409–1419

Sokal RR, Rohlf FJ (1998) Biometry, 4th edn. W. H. Freeman, New York

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

Tani N, Takahashi T, Iwata H, Mukai Y, Ujino-Ihara T, Matsumoto A, Yoshimura K, Yoshimaru H, Murai M, Nagasaka K, Tsumura Y (2003) A consensus linkage map for sugi (*Cryptomeria japonica*) from two pedigrees, based on microsatellites and expressed sequence tags. Genetics 165:1551–1568

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

Tsumura Y, Suyama Y, Yoshimura K, Shirato N, Mukai Y (1997) Sequence-tagged-sites (STSs) of cDNA clones in *Cryptomeria japonica* and their evaluation as molecular markers in conifers. Theor Appl Genet 94:764–772

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

Van Ooijen JW, Voorrips RE (2001) Joinmap 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands, Website: http://www.joinmap.nl

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. Nucl Acids Res 23:4407–4414

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Luebberstedt T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Waugh R, Bonar N, Baird E, Thomas B, Graner A, Hayes P, Powell W (1997) Homology of AFLP products in three mapping populations of barley. Mol Gen Genet 255:311–321

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

Yin TM, Wang XR, Andersson B, Lerceteau-Köhler E (2003) Nearly complete genetic maps of *Pinus sylvestris* L. (Scots pine) constructed by AFLP marker analysis in a full-sib family. Theor Appl Genet 106:1075–1083

Yin TM, DiFazio SP, Gunter LE, Riemenschneider D, Tuskan GA (2004) Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. Theor Appl Genet 109:451–463