Theor Appl Genet (2003) 106:1075–1083
DOI 10.1007/s00122-003-1194-3

T.-M. Yin · X.-R. Wang · B. Andersson · E. Lerceteau-Köhler

Nearly complete genetic maps of *Pinus sylvestris* L. (Scots pine) constructed by AFLP marker analysis in a full-sib family

Received: 14 December 2001 / Accepted: 16 October 2002 / Published online: 21 February 2003

Abstract We have constructed nearly complete linkage maps of *Pinus sylvestris* (L.) using AFLP markers based on a two-way pseudo-testcross strategy in a full-sib family founded in an advanced breeding program. With 39 primer combinations, a total of 737 markers (320 from the mother and 417 from the father) segregated in a 1:1 ratio, corresponding to DNA polymorphism: heterozygous in one parent and null in the other. In the maternal parent, 188 framework markers were mapped in 12 linkage groups, equivalent to the *Pinus* haploid chromosome number, with a total coverage of 1,695.5 cM. In the paternal parent, 245 framework markers established a map with 15 linkage groups, spanning a genome length of 1,718.5 cM. The estimated total map length was $\hat{L}_F = 1,681$ cM for the female and $\hat{L}_M = 1,645$ cM for the male using a modified method-of-moment estimator. Combining these values with those estimated from the observed map lengths in both parents, we estimated the genome length in Scots pine to be between 1,600 and 2,100 cM. Our genome coverage was estimated to be more than 98% with a framework marker interval of 20 cM for both parents. Most of the female and male linkage groups were associated through the analysis of the intercross markers.

Keywords *Pinus sylvestris* · AFLPs · Linkage map · Genome length

Communicated by D.B. Neale

T.-M. Yin · E. Lerceteau-Köhler
Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, S-901 83 Umeå, Sweden

X.-R. Wang
Institute for Working Life, S-907 13 Umeå, Sweden

X.-R. Wang
Laboratory of Systematics and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, 100093 Beijing, China

B. Andersson
The Forestry Research Institute of Sweden, P.O. Box 3, S-918 21, Sävar, Sweden

E. Lerceteau-Köhler (✉)
Centre Interrégional de Recherche et d'Expérimentation de la Fraise, Lanxade, F-24130 Prigonrieux, France,
e-mail: elercete@bordeaux.inra.fr
Fax: +33-5-57-12-24-39

Introduction

Forests are, collectively, the most widespread type of terrestrial ecological system and provide some of the most important sustainable resources for humankind. However, our knowledge of forest trees is limited in comparison with that of many other organisms, and a better understanding of the genetic mechanisms influencing tree adaptation and productivity is essential for the management of the world's forest resources (O'Malley and Whetten 1997).

Genetic mapping is a powerful approach for analyzing the structure of the genome and its evolution in forest trees (Yin et al. 2001a). It offers new opportunities to assist traditional tree-breeding programs through the identification of quantitative trait loci (Bradshaw and Stettler 1995) and their integration into marker-assisted selection programs (Plomion et al. 1996). In perennial species, QTL mapping is a useful tool for studying gene expression at different developmental stages (Wu et al. 1998; Lerceteau et al. 2001), which is highly desirable for carrying out early stage-selection in forest trees. Moreover, when data from highly conserved and orthologous markers such as Expressed Sequence Tag markers are added, it is possible to develop consensus maps (Temesgen et al. 2001), which can help the identification of candidate genes affecting important traits.

The genus *Pinus* comprises about 100 or more species, including many commercially important taxa, and has a wide distribution all over the world. In many areas of the globe, pine trees dominate the landscape. Pine, like other conifers, has haploid megagametophyte tissue, which is ideal material for mapping studies. Although the quantity

of this tissue contained in a single seed is usually very limited, there is enough to generate sufficient PCR-based anonymous markers for mapping purposes. Most of the linkage maps in pine have been constructed from megagametophytes (Nelson et al. 1993; Travis et al. 1998; Remington et al. 1999; Costa et al. 2000; Hayashi et al. 2001). However, a few studies have been based on diploid progenies derived from different generations of crosses, including: loblolly pine (e.g. Devey et al. 1994), radiata pine (e.g. Devey et al. 1996) and longleaf pine × slash pine (Kubisiak et al. 1995). In the long term, linkage maps constructed from temporary mapping materials, such as megagametophytes, will probably have limited future applicability, unless the seedlings from which the megagametophytes originated are maintained for further studies. To fulfill the promising possibilities offered by molecular breeding, mapping materials produced in practical breeding programs with comprehensive traditional goals and modes of operation are needed.

Scots pine breeding programs are at an advanced stage in Sweden due to its economic and ecological importance in Scandinavia. New methods based on molecular markers are being evaluated for further integration in selection. By integrating molecular strategies into the ongoing breeding programs, we expect genetic improvement in timber quality, growth rate, cold tolerance and wood production to be accelerated. However, to fulfill all these expectations, high-density linkage maps are highly desirable. To-date, maps in Scots pine have been constructed using megagametophytes with allozyme markers (Rudin and Ekberg 1978; Szmidt and Muona 1989), or RAPD markers (Yazdani et al. 1995). Unfortunately, these maps have been of little use for further application since the numbers of markers analyzed were limited and/or the corresponding progenies of the mapping materials were not maintained.

The Forestry Research Institute of Sweden (SkogForsk) has developed a full-sib family with about 1,000 progenies, which have been kept in the field trial since 1988 and used to develop preliminary maps (Lerceteau et al. 2000). In this paper we report further saturation of the maps with molecular markers and supply framework maps for future QTL investigations. These will provide more-detailed genomic information related to Scots pine, which could be integrated in the future into the practical Scots pine breeding programs.

Materials and methods

Mapping pedigree

The F_1 full-sib family used in this study is part of the breeding program carried out by the Forestry Research Institute of Sweden in Sävar. About 1,000 progenies have been maintained in the field trial since 1988. The parental trees (AC3065 as mother and Y3088 as father) belong to the most advanced Scots pine breeding materials developed for Northern Sweden. The progenies display clear morphological variations in many traits, such as growth rate, cold hardiness and branch characteristics. Ninety-five individuals were randomly selected for map construction in this study.

Genotyping

DNA extraction and AFLP procedures were carried out as described in Lerceteau and Szmidt (1999). In this study, we used data generated by 13 *Eco*RI + 3/*Mse*I + 3 selective primer pairs from our preliminary study (Lerceteau et al. 2000) together with information from another 26 new primer pairs (18 *Eco*RI + 4/*Mse*I + 3 primer pairs from 128 primer combinations and eight *Eco*RI + 3/*Mse*I + 4 primer pairs from a test of 48 primer combinations selected after screening parental DNA samples). All primers used in this study are standard AFLP *Eco*RI (E) and *Mse*I (M) primers (Vos et al. 1995). The AFLP markers were named after the selective primer set used and the order of the bands on the gels. The scoring process was performed visually three times to help ensure correct genotyping.

Map construction and marker distribution

Chi-square tests were performed to check whether individual markers were segregating in a 1:1 ratio ($\alpha \leq 0.01$). Framework maps including only the non-distorted markers were constructed using MAPMAKER 3.0b software (Lander et al. 1987). The data sets were duplicated and inverted to test the repulsion-phase linkage. Linkage groups were assigned with thresholds set at a minimum LOD score of 5.0 and a maximal recombination fraction (θ) of 0.30. For each assigned linkage group, a set of framework markers was selected according to the information from the LOD table generated by the 'two point' command, and then ordered using the First-Order command. The ordered marker sequences were then confirmed using the Ripple command. Markers ordered with low confidence were placed using the Place and Try command. When trying a marker, if LOD scores increased from the position tried towards both ends of the linkage group, the marker was added to the framework map. Tightly linked markers tended to be placed in the same interval on the framework and competed for the same position with very close LOD support. In such cases, only one of the markers was added to the framework, while the others were considered accessory markers of Type I. Distorted and loosely linked markers were taken as an accessory markers of Type II. When the framework maps were established, we decreased the minimum LOD score and increased the maximal θ criterion to see if the different linkage groups could be merged by testing the linkage of the markers at group ends. Linkage maps were generated with the Map command using the Kosambi map function. DrawMap Version 1.1 (Van Ooijen 1994) was used to generate the graphs of the framework maps. Pairwise linkage between the markers segregating in a ratio 1:1 and the intercross markers present in both parents and segregating in a ratio 3:1 was tested using JOINMAP 2.0 (Stam 1993) at a LOD of 3.0. An intercross marker was considered as an accessory marker when linked to more than one marker from the framework map. Female and male linkage groups were declared similar when at least two intercross markers were mapped on both groups. Marker distribution among linkage groups and clustering of accessory markers within linkage groups was tested by the method described by Remington et al. (1999), and all the parameters in the test were set up the same as in the paper of Remington et al. (1999). All the undistorted markers, both framework and accessory, were included.

Genome length and coverage estimates

The recombination length of the Scots pine genome was estimated assuming random marker distribution using Hulbert et al.'s method-of-moment estimator (Hulbert et al. 1988), as modified by Chakravarti et al. (1991) to correct for upward bias related to chromosome ends. The modified equation for estimating genome length ($\hat{L}$) is:

$$\hat{L} = \frac{n(n-1)d}{2k}\left[1 + \sqrt{1 - \frac{2Ck}{n(n-1)}}\right],$$

where n is the total number of markers; d is the map distance corresponding to the LOD threshold, Z, for declaring linkage; k is the number of marker pairs having LOD values at or above Z; and C is the haploid chromosome number. The values tested for Z were 3, 5 and 7. The values of d and k were obtained using the LOD function of MAPMAKER with the minimum LOD threshold equal to Z and the maximum θ set at 0.5. The confidence interval for $\hat{L}$ was calculated from the equation:

$$I_\alpha(\hat{L}) = \hat{L}\left[1 \pm \frac{n_\alpha}{\sqrt{k}}\right]^{-1},$$

where n_α = 1.96 for a significance level α = 5% (Gerber and Rodolphe 1994).

We also made estimates of genome length based on the observed linkage groups under the assumptions that the average distances between the framework markers could be taken as the distance between the terminal markers and the ends of the respective linkage groups, and the distance between the small linkage segments and the main linkage groups was 50 cM (set at two times the maximal map distance corresponding to the LOD threshold of 5.0 adopted for the framework map construction in this paper). The genome coverage was estimated by the function given by Lange and Boehnke (1982), assuming random marker distribution.

$$c = 1 - e^{-2nd/\hat{L}},$$

where c is the percentage proportion of the genome within d cM of a marker, $\hat{L}$ is the estimated genome length and n is the number of markers.

Results

Marker polymorphism and segregation

From 39 primer combinations, 737 bands generated were testcross markers, of which 320 (43.4%) were heterozygous for the mother and 417 (56.6%) heterozygous for the father. On average, each primer pair yielded about 18 visually distinguishable testcross markers. According to the χ^2 ($\alpha \leq 0.01$) tests, 34 testcross markers deviated significantly from the Mendelian 1:1 segregation ratio on both the maternal (10.6%) and paternal (8.1%) sides. All of these were analyzed as accessory markers. Therefore, 285 and 383 undistorted markers were tested in the framework map-construction for maternal and paternal parents, respectively. From 29 primer combinations out of the 39 tested, 128 intercross markers present in both parents were found fitting a 3:1 segregating ratio.

Map construction

Two framework linkage maps, one for each parent, were constructed. Only the markers that could be ordered with high confidence were included. For the maternal parent, AC3065, setting a minimum LOD score of 5.0 and a maximal recombination fraction (θ) of 0.30, 190 framework markers were associated into 18 groups, namely 12 major linkage groups with seven or more markers per group, three small groups with four markers per group, one triplet and two doublets. Setting a LOD score of 3.0 and a θ value of 0.50, the three small groups and the triplet merged with the major groups. When the LOD score was reduced to 2.0, one of the doublets merged with the shortest group, but the other remained unlinked. The 188 framework markers associated in the 12 linkage groups covered a genome size of 1,695.5 cM (not including the unlinked doublet), with an average map distance between adjacent markers of 8.9 cM (Fig. 1). The size of the linkage groups ranged from 103.9 cM to 195.8 cM. The linkage group number obtained for the maternal tree corresponded exactly to the chromosome number in *Pinus*. In the mapping process, 25 accessory markers of Type I and 86 of Type II were identified. Eighteen markers remained unlinked to the framework markers when the linkage criteria consisted of a minimum LOD of 2.0 and a maximal recombination fraction of 0.5 in the maternal tree. For the paternal parent, Y3088, 252 framework markers were associated into 18 different linkage groups, three triplets and one doublet at a LOD score of 5.0 and a maximal recombination fraction (θ) of 0.30. When linkage criteria were relaxed to 3.0 for the LOD score and 0.50 for θ, three small groups and two of the triplets merged with other groups. Fifteen linkage groups (12 major groups longer than 80 cM and three small groups ranging from 35.2 to 52.5 cM in length) consisting of 245 framework markers, one triplet and one doublet were finally obtained for the paternal tree (Fig. 2), covering a genome size of 1,759 cM with an average map distance of 7.0 cM. The length of the 12 major linkage groups ranged from 80.9 cM to 170.4 cM. In the mapping process for the paternal tree, 24 accessory markers of Type I and 133 of Type II were identified. Eight markers remained unlinked under the LOD value of 2.0 and a maximal recombination fraction of 0.5. Accessory markers of Type I could be used as alternative framework markers (see Figs. 1 and 2).

Among the 128 intercross markers, 76 (59%) and 72 (56%) were mapped as accessory markers on the female and male maps respectively (Figs. 1 and 2). Forty-five markers (35%) were mapped on both maps, 62 (48%) on one and 21 (16%) on none. Eleven female groups could be associated with 12 male groups (F2 with M1, F3 with M8, F4 with M5, F5 with M9, F6 with M2, F7 with M7, F8 with M6 and M13, F9 with triplet, F10 with M4, F11 with M3 and F12 with M12). M6 and M13 should thus be considered as one linkage group, which is not in contradiction with the size of the linkage groups. M14, a small linkage group, may also be linked with M9 forming a longer linkage group (association of M14 with F5 through only one marker) and M15 may be linked with M12. F1, M10, M11 and the doublet remained unassociated, but one of the male groups should be associated with F1 and the other should be linked to the triplet and F9.

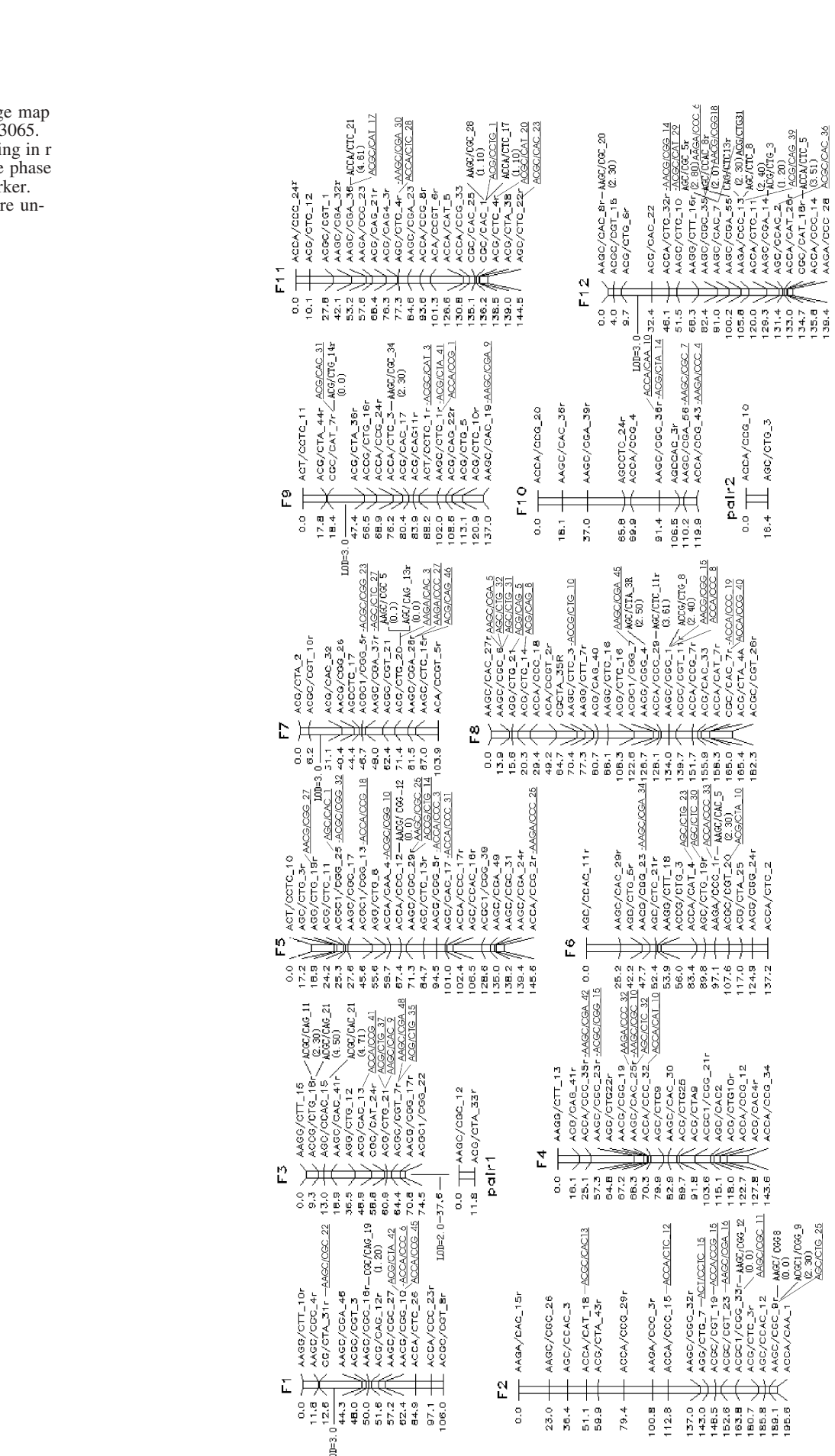

Fig. 1 Framework linkage map of the maternal tree, AC3065. Markers with names ending in r were in repulsion linkage phase to the corresponding marker. The intercross markers are underlined

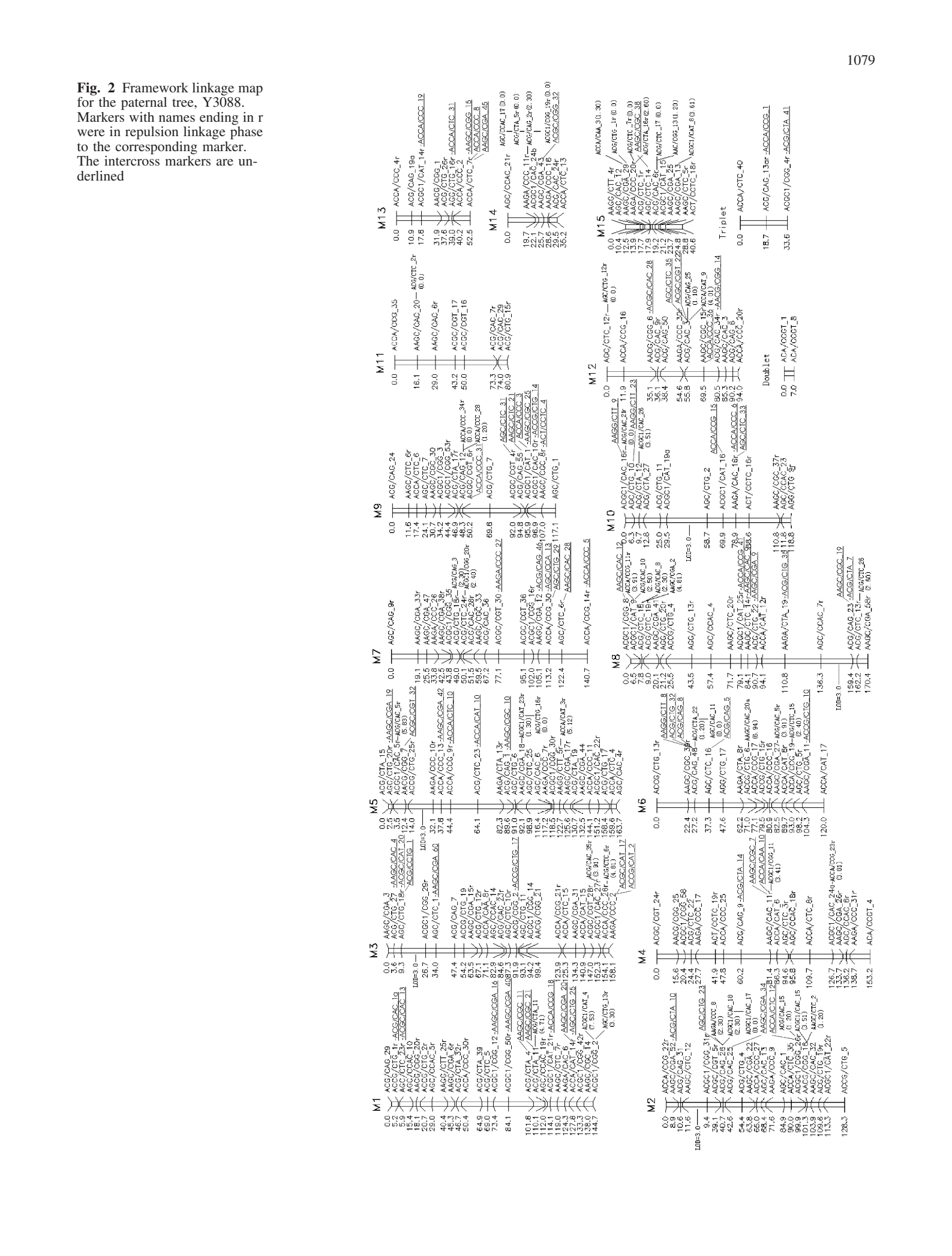

Fig. 2 Framework linkage map for the paternal tree, Y3088. Markers with names ending in r were in repulsion linkage phase to the corresponding marker. The intercross markers are underlined

Table 1 Genome length (L) estimated by the adjusted Hulbert et al.'s method-of-moment estimator in both parents, where *n* is the total number of markers, *d* is the map distance corresponding to the LOD threshold, *Z*, for declaring linkage and *k* is the number of marker pairs

Z value	K	n	d	L	Confidence interval
Maternal tree (distorted markers not included)					
3	1,711	286	36	1,460	1,393–1,532
5	1,222	286	28	1,681	1,591–1,780
7	906	286	22	1,835	1,722–1,962
Average				1,659	1,569–1,758
Maternal tree (distorted markers included)					
3	2,134	320	36	1,468	1,408–1,533
5	1,624	320	28	1,572	1,499–1,652
7	1,056	320	22	1,964	1,852–2,090
Average				1,668	1,586–1,758
Paternal tree (distorted markers not included)					
3	3,047	383	36	1,475	1,384–1,486
5	2,236	383	28	1,645	1,579–1,716
7	1,710	383	22	1,739	1,660–1,825
Average				1,619	1,541–1,676
Paternal tree (distorted markers included)					
3	3,455	417	36	1,556	1,504–1,608
5	2,494	417	28	1,761	1,692–1,831
7	1,955	417	22	1,809	1,729–1,889
Average				1,696	1,642–1,776

Genome length and coverage

The observed framework map lengths in the combined linkage groups, including the triplets and doublets, equaled 1,695.5 cM for the maternal tree and 1,759 cM for the paternal tree. If we assume that the 12 linkage groups for the maternal tree and the 12 major linkage groups for the paternal tree correspond to the 12 chromosomes of *Pinus*, with the assumptions described in **Material and Methods**, the estimated map length would be 1,900 cM and 2,100 cM for the mother and father, respectively. Based on female and male genome lengths, *L*, of 1,900 cM and 2,100 cM, respectively, the estimated coverage when only including the framework markers would be 86.2% and 98.1% of the genome, respectively, within 10 cM and 20 cM of a framework marker of the maternal parent and 91.3% and 99.2 % within 10 cM and 20 cM of a framework marker of the paternal parent.

We also estimated the genome lengths of both parents by the adjusted Hulbert et al.'s method-of-moment estimator. For $Z = 5$, the value corresponding to the threshold for the framework map's construction, the estimated total map length was $\hat{L}_F = 1,681$ cM for the female and $\hat{L}_M = 1,645$ cM for the male (Table 1). For $Z = 3$, 5 and 7, the estimated genome length ranged from 1,460 to 1,835 cM and 1,475 to 1,739 cM for the female and male, respectively, with averages of $\hat{L}_F = 1,659$ cM and $\hat{L}_M = 1,619$ cM. Average estimated values appeared very close to those estimated under $Z = 5$ in both parents. These estimates were based only on the undistorted markers. We also tested the estimated map length when the distorted markers were included (Table 1). For the female map, the results were very close to the previous estimates under $Z = 3$, whereas when $Z = 5$ or 7 the difference in the estimates was greater. However, the average values were still very close. The differences were larger for the male map.

Marker density and clustering of accessory markers

Tests of the marker distribution among linkage groups compared the total number of markers, *mi*, for each linkage group with its expected value, $\lambda i = mGi/\Sigma iGi$. These tests were done on the 12 linkage groups obtained for the maternal tree and the 12 major linkage groups derived for the paternal tree. The results showed that Poisson probabilities for deviations of *mi* from λi in either direction were greater than 0.025 for all linkage groups in both parental trees. Therefore, we did not detect significant differences in marker density among linkage groups at the 0.05 probability level. When we tested for clustering of accessory markers within linkage groups, two out of 188 intervals of the maternal tree showed deviations from the Poisson expectation at the 0.05 level (i.e. $P(x \leq b_{ij}) < 0.025$ or $P(x \geq b_{ij}) < 0.025$) and the 0.01 level. In the paternal tree, six out of 245 intervals showed deviations from the Poisson expectation at the 0.05 level and three deviated at the 0.01 level. In both parents, these deviated intervals were distributed on different linkage groups.

Discussion

Estimates of genome coverage (Lange and Boehnke 1982) indicate that 300 randomly distributed markers should cover 95% of the Scots pine genome at an evenly spaced distance of 10 cM between adjacent markers. In

our study, these figures have been almost reached in the paternal map using the AFLP technique. The bands generated by AFLP markers have a higher multiplex ratio and reproducibility than RAPDs (Vos et al. 1995; Remington et al. 1999; Yin et al. 2001b). Therefore, AFLP analysis is the most efficient and economic technique for identifying markers for mapping (Paglia et al. 1998). A highly reliable framework map constructed with random markers could be used as the starting point for the addition of new codominant markers developed in other species of the *Pinus* genus (Devey et al. 1999; Temesgen et al. 2001).

In linkage mapping, it is better to choose a subset of available markers that can be ordered reliably to construct a 'framework map' rather than to construct a 'comprehensive map' by placing all the markers on the maps (Keats et al. 1991). In our study, we selected 188 and 245 markers to construct the framework maps for the maternal and paternal trees, respectively. In an analysis of loblolly pine (Remington et al. 1999) only 184 markers, which could be ordered uniquely at specified support levels, were selected from a set of 521 markers to construct the framework map. In the mapping process we followed, we designated some markers as type II accessory markers. If markers from this set had been placed on the maps, the map length would have been substantially increased in a great proportion. Thus, we subsequently checked the bands of these markers. For the undistorted markers, there was no evidence to suggest that these markers were subject to more genotyping error than those placed on the framework maps, so we deduced that they may be involved in more complex recombination events. In contrast, the type I accessory markers could serve as alternative markers for the most closely linked framework markers. If type I accessory markers had been placed on the framework maps, the map length would not have increased significantly. With a larger population size, this set of markers could be mapped confidently. The marker order in the new map did not differ from that in the preliminary maps (Lerceteau et al. 2000). However, since only a limited number of markers were available when constructing the preliminary maps, the linkage groups were relatively small. The small linkage segments were easily integrated in the new maps.

Our results indicated that the proportion of deviated intervals was low and no clustering of those intervals was observed in both parents. Most of the AFLP markers were therefore randomly generated and evenly distributed along the chromosomes of the whole genome even within this large set of data. The presence of a higher proportion of deviated intervals in the paternal tree suggested that when a large number of markers are generated, the markers are less evenly distributed along the genome. This tendency was also observed in Remington et al. (1999).

The proportion of distorted testcross markers (10.6% in the mother and 8.1% in the father) in our study was similar or lower than corresponding proportions found in other examples of pseudo-testcross mapping: e.g. 9% in *Eucalyptus urophylla* and 8% in *Eucalyptus grandis* (Verhaegen and Plomion 1996); 13% and 12%, respectively, in *Pinus palustris* and *Pinus elliottii* (inter-specific cross) (Kubisiak et al. 1995); 18% for both parents of a *Quercus robur* full-sib family (Barreneche et al. 1998); 18% in *Populus alba* and 13% in *Populus adenopoda* (Yin et al. 2001a). Zamir and Tadmor (1986) suggested that the deviation should increase with the level of divergence among the parents. The low level of distortions observed in our intra-specific cross may therefore be correlated with the close genetic proximity of the parents. Departures from expected Mendelian ratios often indicate a linkage between molecular markers and distorting factors (Zamir and Tadmor 1986). For example, Bradshaw and Stettler (1994) reported a pollen lethal allele that caused segregation distortion in *Populus trichocarpa*. In many species, such as pines, eucalyptus and oaks, there is a general trend for distorted markers to cluster on particular linkage groups (Kubisiak et al. 1995; Verhaegen and Plomion 1996; Barreneche et al. 1998). Similarly, in this study, we found linkage group F12 in the maternal parent and linkage groups M8 and M11 in the paternal parent contain clusters of distorted markers (data not shown).

Genome organization in conifers appears to be highly conserved in terms of both ploidy level and chromosome number. The haploid chromosome number for the genus *Pinus* is 12 (Sax and Sax 1933). However, most of the maps so far constructed have more than the 12 linkage groups expected according to the chromosome number, except for those of maritime pine (Plomion et al. 1995a; Costa et al. 2000) and loblolly pine (Remington et al. 1999), which were constructed using megagametophytes. To our knowledge, the female map presented here is the first map based on diploid material for which 12 linkage groups were obtained in *Pinus*. In contrast, more than 12 linkage groups have been generated on the male side. Several reports have indicated that greater recombination occurs in male meiosis than in the female (Moran et al. 1983; Robertson 1984; Groover et al. 1995; Yin et al. 2001b) and this tendency may be inherent to the genus *Pinus* (Groover et al. 1995). It may also be true in Scots pine since, although there were more markers available for the male parent than in the female parent, more gaps were left unfilled in the male map.

According to the published linkage maps for *Pinus* spp., the observed genome lengths have ranged from about 1,300 cM to more than 3,000 cM (Nelson et al. 1993, 1994; Plomion et al. 1995a; Yazdani et al. 1995; Travis et al. 1998; Costa et al. 2000). Based on partial linkage data and Hulbert's estimator, Gerber and Rodolphe (1994) estimated the genome size of maritime pine to be about 2,000 cM and Echt and Nelson (1997) obtained similar estimates for three species of *Pinus* using a set of standardized criteria. Remington et al. (1999) derived another estimate of the genome length of loblolly pine, of approximately 1,700 cM Kosambi. Estimates based on chiasma frequency have suggested that the total map distance in *Pinus* should be approximately 1,500 cM

(Plomion et al. 1995b). From the nearly complete framework maps of Scot pine presented in this study, our estimates of genome length based both on the observed map length and the adjusted Hulbert's estimator would range from 1,600 to 2,000 cM Kosambi. The discrepancies in the results from the different studies may be due in part to differences in the choice of map function, the criteria used and the differences in recombination rates between pollen and seed parents (Plomion and O'Malley 1996; Echt and Nelson 1997; Remington et al. 1999). Overestimates of linkage map length may also be due to genotyping errors in the marker data: it has been shown, for instance, that a 3% error rate in genotyping can double the genetic map length (Brzustowicz et al. 1993). Genome size can be calculated from complete linkage maps only when confidently ordered markers are included. The genome of pine species is composed of 10–20×10^6 kilobases (kb) (Rake et al. 1980). According to the estimated genome length, one centiMorgan represents between 5,000 to 10,000 kb in *Pinus*. The large number of kilobases per centiMorgan indicates that the recombination frequency is very low per unit length of DNA in pine species, since in some other tree species, such as eucalyptus (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996) and oak (Barreneche et al. 1998), the number of kilobases per centiMorgan is much lower.

Chakravarti et al. (1991) have proposed a new adjusted method for estimating the genome length. Using computer simulations and experimental data, they compared their estimator with the one published by Hulbert et al. (1988), and showed that the latter tended to overestimate the genome length (partly, they suggested, because of ignorance of the chromosome end effect). Moreover, they observed that Hulbert et al.'s estimator appeared to be affected by the sample size, whereas their adjusted estimator was not, as later confirmed by Gerber and Rodolphe (1994) with data from maritime pine. However, Gerber and Rodolphe (1994) obtained results concerning the genome length that conflicted with these hypotheses, since the derived size was lower according to Hulbert et al.'s estimator than when the adjusted estimator was used. Both studies showed that the adjusted estimator was affected more by the threshold chosen for the LOD score than that of Hulbert et al. However, since the adjusted estimator is insensitive to the mapping size, we used it to estimate the genome length in this study. Our results show that the estimated length varied with the chosen LOD score threshold and increased as the Z value was increased (Table 1). The same trend was observed in studies reported by Gerber and Rodolphe (1994). Our results show that the average lengths estimated at three different Z values were similar to the results obtained under Z = 5 (Table 1). Both the average estimates and those obtained under a Z value of five were very close to Remington et al.'s estimates under Z equal to seven in loblolly pine (Remington et al. 1999). When we set the Z value to Remington et al's as one, our estimates were about 200 cM longer than theirs. We may overestimate the genome length under the same Z value, since the estimates based on megagametophytes tend to be more reliable.

The maps we constructed in this study constitute a comprehensive basis for further genomic studies and provide a useful framework for consensus map construction by adding orthologous anchor markers developed in *Pinus* (Temesgen et al. 2001). This work should facilitate quantitative trait loci identification and marker-aided selection in Scots pine due to the following qualities of the material tested. The mapping pedigree was founded in advanced breeding programs of Scots pine. A large progeny size is maintained in the field trial and annual investigations are carried out on various morphological traits. The progenies are already 14 years old and most of them have reached an age where they produce fruit, which enables more intensive QTL analysis by progeny testing or the addition of more individuals to the test program. It is also good material in which to study the expression of QTLs as a function of time, which will provide highly reliable information for carrying out earlier-stage-selection in Scots pine. The association of most of the female and male linkage groups is a plus in the QTL studies when comparing QTL location on the two parental maps.

Acknowledgements We thank Dr. Wan Song for her excellent help with the laboratory work, Drs. David O'Malley, Rongling Wu and Nengjun Yi for enlightening discussions and suggestions, and Dr. Alfred Szmidt for valuable comments on the manuscript. The study was sponsored by the Swedish Council for Forestry and Agricultural Research (SJFR/Sweden), the Centre for Forest Biotechnology and Chemistry (Sweden), the Carl Tryggers fund (Sweden) and the Erik-Philip Sörensens fund.

References

Barreneche T, Bodenes C, Lexer C, Trontin JF, Fluch S, Streiff R, Plomion C, Roussel G, Steinkellner H, Burg K, Favre JM, Glossl J, Kremer A (1998) A genetic linkage map of Quercus robur L. (Pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. Theor Appl Genet 97:1090–1103

Bradshaw HD, Stettler RF (1994) Molecular genetics of growth and development in Populus 2. Segregation distortion due to genetic load. Theor Appl Genet 89:551–558

Bradshaw HD, Stettler RF (1995) Molecular genetics of growth and development in Populus 4. Mapping QTLs with large effects on growth, form and phenology traits in a forest tree. Genetics 139:963–973

Brzustowicz LM, Merette C, Xie X, Townsend L, Gilliam TC, Ott J (1993) Molecular and statistical approaches to the detection and correction of errors in genotype databases. Am J Hum Genet 53:1137–1145

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

Costa P, Pot D, Dubos C, Frigerio JM, Pionneau C, Bodenes C, Bertocchi E, Cervera M-T, 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 BH, Knapp SJ, Neale DB (1994) An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. Theor Appl Genet 88:273–278

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

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

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

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

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

Groover AT, Williams CG, Devey ME, Lee JM, Neale DB (1995) Sex-related differences in meiotic recombination frequency in Pinus taeda. J Hered 86:157–158

Hayashi E, Kondo T, Terada K, Kuramoto N, Goto Y, Okamura M, Kawasaki H (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

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

Keats BJB, Sherman SL, Morton NE, Robson EB, Buetow KH, Cartwright PE, Chakravarti A, Francke U, Green PP, Ott J (1991) Guidelines for human linkage maps – an International system for Human Linkage Maps (ISLM 1990). Ann Hum Genet 55:1–6

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

Lander ES, Green P, Abrahamson J, Barlow A, Daly M, 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, Szmidt AE (1999) Properties of AFLP markers in inheritance and genetic diversity studies of Pinus sylvestris L. Heredity 82:252–260

Lerceteau E, Plomion C, Andersson B (2000) AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in Pinus sylvestris: a preliminary study. Mol Breed 6:451–458

Lerceteau E, Szmidt AE, Andersson B (2001) Detection of quantitative trait loci in Pinus sylvestris L. across years. Euphytica 121:117–122

Moran GF, Bell JC, Hilliker AJ (1983) Greater meiotic recombination in male vs female gametes in Pinus radiata. J Hered 74:62

Nelson CD, Nance WL, Doudrick RL (1993) A partial genetic linkage map of slash pine (Pinus elliottii Engelm. var. elliottii) based on random amplified polymorphic DNAs. Theor Appl Genet 87:145–151

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

O'Malley DM, Whetten R (1997) DNA markers: protocols, applications and overviews. John Wiley and Sons, New York

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

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

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

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

Plomion C, Durel CE, Verhaegen D (1996) Marker-assisted selection in forest tree breeding programs as illustrated by two examples: Maritime pine and eucalyptus. Ann Sci For 53:819–848

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

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

Robertson DS (1984) Different frequency in the recovery of crossover products fro the male and female gametes of plants hypoploid for B-A translocations in maize. Genetics 107:117–130

Rudin D, Ekberg I (1978) Linkage studies in Pinus sylvestris L. – using macrogametophyte allozymes. Silvae Genet 27:1–12

Sax R, Sax HJ (1933) Chromosome number and morphology in conifers. J Arnold Arbor Harvand Univ 14:356–357

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

Szmidt AE, Muona O (1989) Linkage relationships of allozyme loci in Pinus sylvestris. Hereditas 111:91–97

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

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

Van Ooijen JW (1994) Drawmap – a computer program for drawing genetic linkage maps. J Hered 85:66

Verhaegen D, Plomion C (1996) Genetic mapping in Eucalyptus urophylla and Eucalyptus grandis using RAPD markers. Genome 39:1051–1061

Vos P, Hogers R, Bleeker M, Reijans M, Vandelee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wu R, Bradshaw Jr HD, Stettler RF (1998) Developmental quantitative genetics of growth in Populus. Theor Appl Genet 97:1110–1119

Yazdani R, Yeh FC, Rimsha J (1995) Genomic mapping of Pinus sylvestris (L.) using random amplified polymorphic DNA markers. For Genet 2:109–116

Yin TM, Huang MR, Wang MX, Zhu LH, Zeng ZB, Wu RL (2001a) Preliminary interspecific genetic maps of the Populus genome constructed from RAPD markers. Genome 44:602–609

Yin TM, Zhang XY, Huang MR, Wang MX, Zhuge Q, Zhu LH, Zeng ZB, Wu RL (2001b) Molecular linkage maps of the Populus genome. Genome 45:541–555

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot Gaz 147:355–358