Annotated genetic linkage maps of *Pinus pinaster* Ait. from a Central Spain population using microsatellite and gene based markers

Marina de Miguel1,2, Nuria de María1,2, M Ángeles Guevara1,2, Luis Díaz1,2, Enrique Sáez-Laguna1,2, David Sánchez-Gómez1, Emilie Chancerel4,5, Ismael Aranda1, Carmen Collada2,3, Christophe Plomion4,5, José-Antonio Cabezas1,2 and María-Teresa Cervera1,2*

**Abstract**

**Background:** *Pinus pinaster* Ait. is a major resin producing species in Spain. Genetic linkage mapping can facilitate marker-assisted selection (MAS) through the identification of Quantitative Trait Loci and selection of allelic variants of interest in breeding populations. In this study, we report annotated genetic linkage maps for two individuals (C14 and C15) belonging to a breeding program aiming to increase resin production. We use different types of DNA markers, including last-generation molecular markers.

**Results:** We obtained 13 and 14 linkage groups for C14 and C15 maps, respectively. A total of 211 and 215 markers were positioned on each map and estimated genome length was between 1,870 and 2,166 cM respectively, which represents near 65% of genome coverage. Comparative mapping with previously developed genetic linkage maps for *P. pinaster* based on about 60 common markers enabled aligning linkage groups to this reference map. The comparison of our annotated linkage maps and linkage maps reporting QTL information revealed 11 annotated SNPs in candidate genes that co-localized with previously reported QTLs for wood properties and water use efficiency.

**Conclusions:** This study provides genetic linkage maps from a Spanish population that shows high levels of genetic divergence with French populations from which segregating progenies have been previously mapped. These genetic maps will be of interest to construct a reliable consensus linkage map for the species. The importance of developing functional genetic linkage maps is highlighted, especially when working with breeding populations for its future application in MAS for traits of interest.

**Keywords:** *Pinus pinaster*, Genetic linkage map, Functional annotation, Microsatellites, SNPs

**Background**

Maritime pine (*Pinus Pinaster* Ait.) is one of the most important species in the Mediterranean region for its ecology and wood productiveness. As other conifers, this long lived species dominates different landscapes and can withstand severe environmental conditions [1]. Several studies have revealed high levels of phenotypic variation [2-4] and genetic diversity [5-7] in maritime pine. This species has a fragmented geographic distribution that could be subdivided into different meta-populations based on its high level of genetic differentiation [8-10]. In the Iberian Peninsula different patterns of local adaptation have been identified [11]. Besides its ecological value, maritime pine is also a significant species for its economic importance. Particularly, *P. pinaster* is a major resin producing species in the Iberian Peninsula [12]. The resin is at the basis of many manufactured products such as turpentine, oils, varnishes, sealing wax, plastics and others. In 1990s resin tapping was reintroduced in Spain after a drastic reduction in 1970s due to the international crisis in this sector [13]. Many of the abandoned stands have been tapped again. In particular,
natural stands of Central Spain are one of the most important resin tapping regions [14]. As resin production shows high heritability [15] a breeding program is a useful strategy to improve productiveness [16]. Consequently several breeding programs have been implemented for resin production in maritime pine [17–19].

Genetic linkage mapping can facilitate marker-assisted selection (MAS) as it allows the identification of quantitative trait Loci (QTL) [20–23]. Furthermore, as genome organization is well conserved in conifers, comparative mapping is a useful strategy to find homologous chromosomal segments involved in the genetic control of economical and adaptive traits [24,25].

Traditional molecular makers, such as proteins, RFLPs (Restriction Fragment Length Polymorphisms), RAPDs (Random Amplified Polymorphic DNAs), AFLPs (Amplified Fragment Length Polymorphisms) and nSSRs (nuclear Simple Sequence Repeats) have help to build a first generation genetic linkage maps in forest trees [26,27]. The use of RAPDs and AFLPs, randomly distributed in the genome [28,29], has allowed the construction of genetic linkage maps from species with large genome sizes like conifers [30–32]. An alternative for species with extremely large genomes or for populations with low levels of polymorphism are SAMPL markers (Select Amplication of Microsatellite Polymorphic Loci). SAMPL combines the advantages of AFLPs and microsatellites resulting in higher percentage of polymorphic markers per assay and higher repeatability between assays [33].

In recent years, efforts have focused in sequencing genes of interest to build genetic linkage maps with direct functional information [34]. Functional genetic linkage maps have experienced a revolution with the availability of new sets of markers from coding regions such as: EST-Ps (Expressed Sequence Tags Polymorphisms), EST-SSRs (EST derived microsatellites) and SNPs (Single Nucleotide Polymorphisms) [35–37]. Functional genetic linkage maps based on annotated genes allow to assess redundant and paralogous EST markers and further improve the quality and utility of genetic maps [38]. Specifically, SNPs have several advantages for their use as molecular makers because they are very abundant in the genome, they show higher stability than SSRs, are usually bi-allelic and codominant [39,40]. Moreover, new technologies have been developed for high throughput detection and genotyping of SNPs reducing the cost of assays [41,42]. Thus, highly saturated genetic linkage maps can be constructed even for species with large and un-sequenced genomes like conifers [21,43–46].

As other pines, P. pinaster is a diploid organism characterized by a large and complex genome with high low-copy fraction [47,48]. Particularly, maritime pine has $2n = 24$ chromosomes and its genome size is estimated between 51–62 pg/2C [49,50]. Several genetic linkage maps have been developed for maritime pine based on proteins [51–54], RAPDs [54–57], AFLPs [44,49,54,58], SSRs [44,58–60], EST-Ps [44,58,61] and SNPs [44]. Also, comparative mapping have been performed with Pinus taeda L. [44,61]. None of the genetic linkage maps available for P. pinaster has been derived from individuals belonging to Spanish populations. These populations show high levels of genetic divergence with the French populations used to design mapping progenies in previous genetic linkage maps [8]. As maritime pine shows a fragmented geographic distribution with high levels of population genetic structure and variation [6,8] it is important to explore the genetic organization in a representative population from the Castilian Plateau (Central Spain) and thus better cover the natural distribution of the species.

Thus, the main objective of this work was to construct saturated genetic linkage maps for P. pinaster using controlled crosses between two trees that take part in a breeding program for resin production in a natural population from Central Spain, as a first step to the genetic dissection of this trait. Combining different kind of molecular markers we aim to construct a map with annotated gene functions and homologous markers with previous maps for contributing to the development of a consensus map for the species. A second objective was to identify candidate genes overlapping with QTL already detected in this species [62–64].

**Methods**

**Mapping populations**

Two outbred full-sibs families of P. pinaster were used for genetic linkage mapping. Progenies were originated from two reciprocal controlled crosses between two progenitors (C14 and C15) belonging to a natural population in Coca (Segovia) located in Central Spain (41° 12’ N 4° 31’ W). Previous studies on this population have showed a differential genetic structure when compared with other populations of the natural distribution of the species [8]. Progenitors took part in a breeding program for resin production started in 1994 and they were selected for their contrasting resin production, low for C14 and higher for C15. Controlled crosses were carried out in 1999 for C14xC15 and in 2000 for C15xC14. F1 seeds were collected and germinated in controlled conditions at Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, INIA (Madrid, Spain). Then they were planted in semi-controlled conditions at Dirección Nacional de Biodiversidad– Madrid (40° 27’ N 3° 44’ W). A paternity test analysis was performed with 13 SSRs. Finally, once the contaminants were removed, the mapping population comprised 161 individuals: 106 from family C14xC15 and 55 individuals from C15xC14.
Molecular markers

Genomic DNA was extracted from needles using a modified protocol from Dellaporta et al. [65] for all marker analyses, but for the 1536 and 384 GoldenGate assays (Illumina Inc., San Diego, CA, USA), for which a commercial Invitrogen DNA plants HTS 96kit (Invitek GmbH, Berlin, Germany) was used. Four types of molecular marker were used for genotyping the mapping populations: nSSRs, EST-Ps, SAMPLs and SNPs.

nSSRs: Forty seven primer pairs designed for amplification of nSSR loci in P. pinaster and P. taeda [60,66,67] were tested for segregation in the mapping populations. Thirteen loci were polymorphic, 27 were monomorphic, and seven resulted in muti-banding or non-clear patterns. Amplification of A6F03, A5B01, A5A11, A6F10, A6D04, A5B07 loci was performed as in Guevara et al. [66]. Amplification of NZPR823, NZPR413,NZPR114, NZPR544, SsrPt_ctg64, SsrPt_ctg275 loci was performed as described by Chagné et al. [60] and the amplification of PtTX3116 followed the protocol described by Auckland et al. [68] with modified touchdown profile, using 55°C and 45°C as starting and final temperatures [69]. A Perkin-Elmer GenAmp 9700 thermal cycler (Perkin Elmer Inc., Waltham, Massachusetts, USA) was used to carry out PCR reactions. Amplified products were separated in denaturing gels containing 6% acrylamide / bisacrylamide (19:3), 7 M urea and 1x TBE. Amplified products were visualized in a DNA Analyzer System (4300, LI-COR Biosciences, Lincoln, NE, USA). Fragments were scored visually as codominant markers.

EST-Ps: EST-P genotyping was carried out by Tilling (Targeting Induced Local Lesions in Genomes) as described by Till et al. [70]. A set of 14 EST-P primer pairs (PtIFG_893, PtIFG_9136, PtIFG_9034, PtIFG_1955, PtIFG_8429, PtIFG_8702, PtIFG_3C8E, PtIFG_228B, PtIFG_1CA6C, PtIFG_9044, PtIFG_2253, PtIFG_8436, PtIFG_8887, PtIFG_C6H11) derived from cDNA sequences of P. taeda and P. pinaster [61,72,73] were tested, in order to identify the most informative combinations. Amplification of NZPR823, NZPR413, NZPR114, NZPR544, SsrPt_ctg64, SsrPt_ctg275 loci was performed as described by Chagné et al. [60] and the amplification of PtTX3116 followed the protocol described by Auckland et al. [68] with modified touchdown profile, using 55°C and 45°C as starting and final temperatures [69]. A Perkin-Elmer GenAmp 9700 thermal cycler (Perkin Elmer Inc., Waltham, Massachusetts, USA) was used to carry out PCR reactions. Amplified products were separated in denaturing gels containing 6% acrylamide / bisacrylamide (19:3), 7 M urea and 1x TBE. Amplified products were visualized in a DNA Analyzer System (4300, LI-COR Biosciences, Lincoln, NE, USA). Fragments were scored visually as codominant markers.

SAMPLs: SAMPL genotyping was performed as indicated by Vos et al. [28] with several modifications [74]. Preamplifications were carried out using three primer combinations (EcoRI + A/ Msel + G; EcoRI + A/ Msel + C; EcoRI + A/ Msel + T). For the selective amplification a SAMPL primer [CATA: (CA)$_5$(TA)$_2$; GATA: (GA)$_5$(TA)$_2$ [75]], was used in combination with an EcoRI + 3 primer. In order to select the most informative combinations (those with a higher level of polymorphism) different combinations were tested using template DNA from the parental lines and 9 offspring. Progenitor C14 revealed lower levels of polymorphisms than C15 (see Results section), thus primer combinations were chosen in order to equilibrate the number of markers segregating from each progenitor. A total of 31 CATA/EcoRI and 26 GATA/EcoRI primer combinations were used for the selective amplification. Selective PCR reaction were performed in 10 μl of 1x PCR Buffer (10 mM Tris–HCl, 50 mM KCl, pH 8.3), 0.1 mM of each dNTP, 2.5 mM MgCl$_2$ (Roche, Basel, Switzerland), 3 ng IRDye 800 5’end labeled CATA or GATA primers, 15 ng EcoRI + 3 primer, 0.2U Taq DNA polymerase (Invitrogen, Grand Island, NY, USA) and 5 μl of 10-fold diluted pre-amplification DNA fragments using classical AFLP cycling parameters [12]. Samples were loaded into denaturing gels containing 8% Long Ranger polyacrylamide (Cambrex, East Rutherford, NJ, USA), 7 M urea and 1x TBE. Fragments detection was carried out on a DNA Analyzer System (4300, LI-COR Biosciences, Lincoln, NE, USA). Fragments were scored visually as dominant markers.

SNPs: two SNP genotyping assays were used in this study; a 1,536 BeadArray™ and a 384 BeadXpress®
Golden Gate assays (Illumina Inc., San Diego, CA, USA). SNPs selected for 1,536 Golden Gate assay corresponded to three different sets (see Chancerel et al. [44] for further details): in vitro polymorphisms from 35 candidate genes for cell wall formation and drought stress resistance; in silico SNPs from a maritime pine EST assembly; and in silico polymorphism from re-sequenced amplicons of the species. In this genotyping assay, 95 DNA samples of the mapping progenies were genotyped (73 for C14xC15 and 22 for C15xC14). In order to increase the number of genotyped individuals for a set of genes of interest, another genotyping assay was developed. This genotyping assay (384 SNPlex) consisted in a sub-sample of SNPs selected from the 1,536 genotyping assay and 14 additional SNPs from candidate genes for drought resistance [76]. It was carried out at Center for Genomic Regulation (CRG, Barcelona, Spain) for a total of 119 DNA samples (79 for C14x15 and 40 for C15xC14). Both genotyping assays were realized according to the manufacturer’s instructions (Illumina Inc., San Diego, CA, USA) and SNPs clusters revised manually with Illumina Bead Studio v2.0 Software. When the same SNP was successfully genotyped in both assays priority was given for the 384 Vera Code data because of the higher number of DNA samples genotyped in this assay. Contig and gene sequences containing the polymorphic SNPs are presented in Additional file 1.

**Linkage map construction**

For each progenitor we assembled three different linkage maps belonging to datasets of C14xC15 (106 individuals), C15xC14 (55 individuals) and a dataset with the information of the individuals of both reciprocal crosses (161 individuals). Since no relevant differences were found as a consequence of merging both progenies (see Results section), further linkage analyses were developed using only the data set with the merged information of both progenies. Parental maps were constructed using the “two-way-pseudo-testcross” mapping strategy [77]. Markers with more than 70% of missing data were excluded from further analysis. Linkage analyses and map estimations were performed using the regression mapping algorithm implemented in the software JoinMap v4.0 [78] with the CP population type and using a recombination fraction <0.35 and a LOD >3 as mapping parameters. Map distances were calculated using Kosambi mapping function [79]. When difficulties in estimating marker order are found, two additional maps are constructed (map2 and map3). In map2, new markers are added because more pairwise data are available. In map3, the remaining loci are added by decreasing statistical support. In these cases we kept map2 for further analyses. When a pair of markers was considered identical, only one of the markers was selected for mapping. In order to assign unlinked loci to selected linkage groups (LG), the strongest cross link was employed with a LOD value of 3 (JoinMap command “assign ungrouped loci to SCL-groups”). Segregation ratios were tested using $\chi^2$ test ($P \leq 0.01$).

**Evaluation of homogeneity of recombination rate between female and male meiosis**

In order to evaluate whether the male and female gametes presented different levels of recombination, we tested departure from homogeneity of recombination fraction following Plomion et al. [57]. Since the statistical power of homogeneity depends largely on the sample size, the test was performed for all markers pairs in common in the three genetic maps for each progenitor and having a recombination fraction lower than 0.1 (Additional file 2).

**Comparative mapping**

The linkage maps of both progenitors were compared based on common markers. Besides genetic maps were compared with previously developed P. pinaster maps [44] based on common SSRs, ESTPs and SNPs. LGs were named according to Chancerel et al. [44] using loblolly pine nomenclature, as it is the reference pine species.

**Genome length and map coverage**

Total genome length was calculated as the sum of all mapped marker intervals. Estimated genome length ($G_e$), was determined from the partial linkage data according to Hulbert et al. [80] modified by Chakravarti et al. [81] (Method 3). A minimum LOD score of three was chosen to estimate genome length using framework maps constructed following the methodology previously described in order to avoid overestimation of genome size because of clustered markers. Observed map coverage was calculated as the ratio of total genome length to estimated genome length [82].

**Marker distribution**

To evaluate whether markers were randomly distributed, we tested the procedure explained in Echt et al. [38]. A Kolmogorov-Smirnov test for two populations was implemented to compare the observed marker distribution frequencies with expected distribution frequencies under the assumption of randomness. SAMPLs and SNPs distribution were also analyzed by calculating Pearson correlation coefficient between the number of SAMPLs and SNPs in the LGs and the size of the LGs as in Cervera et al. [82].
Heterozygosity levels
The average heterozygosity was estimated for each progenitor and for each molecular marker type independently. Heterozygosity levels based on SNPs were calculated as the ratio between polymorphic and total number of tested SNPs, discarding those with multibanding and non-clear patterns. Three SSR primer pairs resulted in the amplification of two different loci with clearly different segregation patterns and were scored as different markers, but they were considered as only one for heterozygosity estimations. Heterozygosity levels based on SNPs were calculated as the ratio of polymorphic SNPs and total number of SNPs successfully genotyped. Heterozygosity estimates for SAMPLs were calculated for the first primer combination tested; since the following ones were selected in order to maximize the number of polymorphic markers in C14 (see Molecular markers subsection). Heterozygosity based on EST-Ps was not calculated because we only analyzed markers that had been found polymorphic in previous studies in other pine species [61,72,73], therefore a bias could be introduced.

Functional annotation
Functional annotation of the mapped SNP-based genes was carried out using sequence information from the Oligo Pool Assay-OPA (60 nucleotides in length at both sides of the SNP position). In order to obtain homology with longer sequences a BLAST-N search was performed using the pine Gene Index [83] and GeneBank [84]. We retained sequences showing the highest homology (e-value lower than $10^{-20}$ were considered significant). Then, these longer sequences were annotated using Blast2GO software [85]. Gene Ontology (GO) annotation terms for molecular function at ontology level equal to 3 were placed in the map in order to search for clusters of genes with similar function. For sequences where GO annotation for level 3 was not available we selected the GO annotation terms for level 2. To evaluate whether similar GO terms were clustered or randomly distributed along the genome we performed for each GO term at level 2 a Kolmogorov-Smirnov test for two populations as explained in Marker Distribution subsection.

Marker nomenclature
Marker nomenclature for SSRs and SNPs were maintained according to their original publications (see Molecular Markers subsection). EST-Ps also conserved original nomenclature, but the size of the amplified band was added to the marker name. SAMPLs were named with the differential selective nucleotide used in the preamplification (C, G or T), followed by the targeted microsatellite (CATA or GATA), and the selective EcoRI + 3 primer employed, ending by the size of the amplified band fragment.

Results and discussion
The paternity test analysis revealed seven contaminants for C14xC15 and three for C15xC14 that were removed for further analyses. The final number of individuals per progeny, 106 for C14xC15 and 55 for C15xC14, was in the limit for reliable estimations and as we did not observe significant differences in recombination fraction between female and male meiosis (see next subsection) we constructed the genetic linkage maps by pooling all individuals of both reciprocal crosses.

Evaluation of homogeneity of recombination between female and male meiosis
Ninety-six marker pairs for C14 and 42 for C15, with a recombination fraction below 0.1, were available in all three maps (C14xC15, C15xC14 and pool map) (see Methods section). Eight marker pairs out of the 96 analyzed, showed significant differences between female and male meiosis for C14 (data not shown). Five of them showed a higher recombination rate for male meiosis and three for female meiosis. No marker pair resulted in significant differences in recombination rate for C15. The low level of differences detected in recombination fraction between female and male meiosis supports the merging of both progenies in order to obtain a higher number of offspring in the mapping population and thereby establishes more precise parental maps. No evidence of heterogeneity of recombination was previously reported for P. pinaster [56] and other conifer species [87]. However, Plomion and O’Malley [57] suggested that recombination fraction could be higher in male meiosis for P. pinaster. The important differences in number of individuals between our mapping progenies (106 versus 55) compelled us to perform the analyses with a narrow window of markers (only those with a recombination fraction lower than 0.1), while in Plomion and O’Malley [57] analyses were performed with a wider window (markers pairs with a recombination fraction lower than 0.3). This could explain the difference in results obtained. Nevertheless, further research in testing homogeneity of recombination between female and male meiosis is needed to clarify whether or not female and male gametes exhibit similar recombination rate, which can have some implications for MAS.
Individual linkage maps and comparative mapping

Previous analysis of the mapping population with four AFLP primer combinations revealed very low levels of polymorphism (data not shown). Therefore, we decided to use SAMPL technique to increase the number of polymorphic fragments. SAMPL analysis was performed using the most informative primer combinations (see Methods section). This result validates the use of SAMPLs as an alternative for genotyping low polymorphic populations.

Out of the total set of molecular markers available (Table 1), four and five markers were excluded from C14 and C15 datasets respectively, because of their identical segregation profiles with other markers. All of them were SNPs belonging to the same gene or contig. Markers with more than 70% of missing data were also excluded. Most of them were SAMPLs genotyped only in the C15xC14 pedigree. In addition, in a “two-way-pseudo-test-cross” mapping strategy, intercross markers i.e. markers with the same heterozygous allelic configuration in both progenitors, are less informative. Because of that, several SAMPLs, SNPs and one microsatellite marker were excluded. However, when it was possible we kept a number of intercross markers because they allow to align homologous LGs between both parental maps.

Near 5% of markers used for linkage analysis were unlinked, which is in the same range of what has been observed in other conifer maps [24,45]. Most of them presented more than 35% of missing data and corresponded to SAMPLs genotyped only in the C14xC15 pedigree. Several SNPs were also unlinked. Near 94% of the markers could be assigned to LGs and 60% could be positioned in the final maps (Table 1, Figure 1, Figure 2 and Additional file 3). The lower percentage of markers positioned when compared with other highly saturated maps [46] is due to the use of SAMPLs only scoring in one of the mapping progenies, as revealed by the low percentage of positioned SAMPLs (Table 1). When we discard SAMPLs scored only in one or the two mapping progenies, the percentage of SAMPL markers positioned in the parental maps increases to 65.3% for C14 and 72.1% for C15. These results are very similar to those obtained with positioned SNPs (Table 1) indicating that both type of markers are suitable for the construction of linkage maps. Even more, for a complete coverage of the genome it is interesting to use markers with different target sequences, since coding and non-coding regions seems not to be randomly distributed along the genome [87,88].

In a first phase, before aligning on the reference P. pinaster linkage map, 22 LGs were obtained for C14 and 20 for C15 (Table 1, Figure 1 and Figure 2). The smallest LGs were similar in size between C14 and C15 maps. However, the largest LG were higher in C15 than in C14. Also, average size of LGs was slightly higher for C15 than for C14. This was explained because average distance and maximum distance between two adjacent markers was larger in C15 than in C14. Thirty-one intercross markers between both parental maps allowed the identification of homologous LGs (Table 2). Eleven markers with segregation 1:2:1 (same heterozygous combination in both parents) could only be positioned in one parental map (Table 2). Five of them could not be mapped in the other parent because they were ungrouped and the remaining six markers because of the increase in the goodness of fit calculated for the order of markers when were included in the map.

The alignment with maps described by Chancerel et al. [44], based on common SSRs, EST-Ps and SNPs (Table 2), made it possible to bring together some LGs resulting into 13 LGs for C14 and 14 LGs for C15 (Table 1), close to the 12 chromosomes of the haploid P. pinaster genome [89]. In general, similar size of LGs for parental maps was obtained except for LGs 1, 3 and 10 that were larger in C15 map and LG12 that was larger in C14 map. The fact that we could not assemble the markers in 12 LGs, the differences in size of homologous LGs and the presence of common markers only positioned in one parental map, are probably related with the presence of homozygous regions in the genome of these individuals that prevent mapping markers in these areas. This effect was partially expected because previous studies of the population of origin of both parental trees, Coca, revealed a high coefficient of endogamy [90]. As a result of endogamy we would expect a loss of polymorphisms in the individuals coming from this population, strictly confirmed by the low levels of polymorphism detected by AFLPs genotyping (data not shown) and the low levels of heterozygosity found in the parental trees (Table 3) compared with observed heterozygosity in other provenances of P. pinaster [5].

In this respect, it is important to point out to the difference in heterozygosity between C14 and C15 parental trees as revealed by the estimation obtained from SAMPLs (Table 3). This difference was overcome by further genotyping using selected SAMPL primer combinations with a higher number of polymorphic markers in C14 (see Methods section). Percentage of heterozygosity calculated from SSRs and SNPs yielded lower values than those obtained from SAMPLs and differences in heterozygosity between C14 and C15 could not be appreciated. One possible explanation is that analyzed SNPs were selected from coding regions where the level of polymorphism is lower than in non-coding regions [91].
Twenty four contigs with several SNPs (from two to seven) were included in the linkage maps. SNPs belonging to the same gene or contig mapped always in the same position or less than 3 cM away (Figure 1 and Figure 2), except m682 and m127 (LG 5), separated by 26.8 cM. Marker m682 was distorted at the 0.1% significance level, which could affect the accurateness of its position. Alternatively, both SNPs may be associated to different loci at the same LG. The fact that nearly all SNPs belonging to the same contig were mapped in the same position supports the accuracy of the genotyping method used, as previously reported [41,44].

Alignments with the linkage maps developed by Chancérel et al. [44] pointed out that marker order was highly conserved, excepting small inversions of less than 5 cM (data not shown). The only major inconsistency in data was found for marker PtIFG_8436_200, which was amplified using the same primer combination as in Chancérel et al. [44], but subjected to different detection techniques, tilling versus SSCP (Figure 2). This EST-P marker was mapped in LG 10 in our mapping progeny in agreement with previous developed maps in P. taeda [72]. However, in other published linkage maps of P. pinaster this gene was mapped in LG 7 [44,61]. Chagné et al. [61] discussed the possibility that PtIFG_8436 in P. pinaster targeted a paralogous gene as they found low similarity at the DNA sequence level. Our result suggests the existence of an orthologous sequence between P. pinaster and P. taeda genomes for the region amplified by PtIFG_8436 marker in LG 10 and a paralogous sequence in LG 7 of P. pinaster genome.

**Table 1** Mapping parameters of parental linkage maps constructed by merging two reciprocal crosses: C14xC15 and C15xC14 (Continued)

| Mapping parameter | C14   | C15   |
|-------------------|-------|-------|
| Total number of available markers | 402   | 410   |
| Number of SSR loci | 11    | 13    |
| Number of ESTP loci | 13    | 12    |
| Number of SAMPL loci | 228   | 237   |
| Number of SNP loci | 150   | 148   |
| Total number of distorted markers (p ≤ 0.01) | 39    | 52    |
| Number of excluded markers | 62    | 72    |
| Number of SSR loci | 1     | 1     |
| Number of ESTP loci | 0     | 0     |
| Number of SAMPL loci | 46    | 60    |
| Number of SNP loci | 15    | 11    |
| Number of markers not excluded | 340   | 338   |
| Number of assigned markers | 321   | 319   |
| Number of SSR loci | 9     | 10    |
| Number of ESTP loci | 11    | 11    |
| Number of SAMPL loci | 174   | 166   |
| Number of SNP loci | 127   | 132   |
| Number of positioned markers | 215   | 211   (62.4%) |
| Number of SSR loci | 6     | 7     (58.3%) |
| Number of ESTP loci | 10    | 7     (58.3%) |
| Number of SAMPL loci | 98    | 98    (55.4%) |
| Number of SNP loci | 101   | 99    (72.3%) |
| Number of distorted markers (p ≤ 0.01) | 14    | 14    |
| Unlinked markers (%) | 19(6.3%) | 19 (5.6%) |
| Number of LG >3 before making alignments | 22    | 20    |
| Number of LG >3 after making alignments | 13    | 14    |
| Smallest LG (cM) before making alignments | 17.5  | 13.4  |
| Largest LG (cM) before making alignments | 81.1  | 155.3 |
| Average length (cM) LG ± SD before alignments | 53.7 ± 20.6 | 69 ± 35.6 |
| Smallest LG (cM) after making alignments | 52    | 42.3  |
| Largest LG (cM) after making alignments | 142.2 | 155.3 |
| Average length (cM) of a LG ± SD after alignments | 90.8 ± 29.14 | 98.5 ± 38 |
| Maximum distance (cM) between 2 adjacent markers | 24.9  | 35.3  |
| Average distance (cM) between 2 adjacent markers ± SD | 6.12 ± 5.8 | 7.22 ± 6.4 |

*a The 25 ESTPs correspond to 11 gene loci.

b The SNPs markers correspond to 47 gene loci and 143 contigs.

c Markers with more than 70% of missing data (see Methods section) and identical markers.

d Assigned markers correspond to markers linked with more than 2 other markers.

e Unpositioned markers correspond to markers with a recombination frequency higher than 0.35 with the nearest linked marker (unlinked markers) or markers which position could not be reliably estimated. Percentage of unpositioned markers was calculated over the number of not excluded markers.

f Percentage of unlinked markers was calculated over the number of not excluded markers.

SD Standard deviation.

Segregation distortion

A χ² test (d.f. = 1) was performed to test Mendelian segregation of each marker. We detected 9.7% of markers showing distorted segregation ratios at 1% significance level for C14 and 12% for C15 linkage maps (Table 1).
These results are very similar to those obtained in other pine [92,93] and conifer species [36]. The number of distorted markers excluded and unlinked was similar to the number of distorted markers finally assigned to LGs (Additional file 4). Besides, among the unlinked and excluded loci the number of distorted markers was not higher than those showing no segregation distortion (Table 1). Thus, in this case, unlinked and excluded loci seem not to be the result of segregation distortion, as previously reported in other linkage studies [24]. Distorted markers assigned to a LG were randomly distributed (Additional file 4). Only 13 distorted markers could be positioned in each map indicating the difficulty to estimate an accurate position for these distorted markers. Distorted markers positioned in the maps did not appear clustered in specific regions of the genome (Table 1, Figure 1 and Figure 2) suggesting that segregation distortion was probably related with genotyping errors rather than the effect of pre or post-zygotic selection. As they were not clustered they did not compromise map structure [93].

**Marker distribution**

Markers were randomly distributed along the genome as no significant differences were found between distribution of markers along the LGs and expected distribution under the hypothesis of randomness (Kolmogorov-Smirnov for two populations, \( D = 0.55, p\text{-value} = 0.124 \) for C14 map and \( D = 0.5, p\text{-value} = 0.474 \) for C15 map), in accordance to other conifer maps [24,32]. Besides, largest LG had more SNPs (Pearson correlation, \( r = 0.53, p\text{-value} = 0.01 \) for C14 map and \( r = 0.62, p\text{-value} = 0.003 \) for C15 map) than smaller LG. Same results were obtained for SAMPLs (Pearson correlation, \( r = 0.69, p\text{-value} < 0.001 \) for C14 map and \( r = 0.75, p\text{-value} < 0.001 \) for C15 map) indicating that they are also randomly distributed along the genome, as expected for this kind of multiband markers [28].
 Genome length and map coverage
Observed genome length ranged from 1,180.4 (C14) to 1,379.5 cM (C15), 200 cM larger for C15 map than for C14 map (Table 1). In other *P. pinaster* maps observed genome length ranged from 869 to 1,860 cM depending on the density of markers [44,55,58]. The higher genome length observed in C15 map agrees with its higher heterozygosity estimation compared to C14 map (Table 3). Estimated genome length ranged from 1,870.2 to 2,166.6 cM, in line with what has been obtained in previous *P. pinaster* maps (1,223 to 3,252 cM depending on the method of estimation [44,56,88]). The last generation maps estimated *P. pinaster* genome size to be 2,500 cM [44], a value which is near our estimates and close to other pine species [94]. Estimated genome length was higher for C15
linkage map although the observed map coverage (near 65%, Table 1) was similar for both parental maps. High density genetic linkage maps usually report map coverage over 90% [46,95]. However, previously published P. pinaster genetic linkage maps also reported map coverage near 65% [44,57], indicating the difficulty to achieve a complete coverage for such a large and complex genome.

Functional annotation

We validated and improved the functional annotation information for mapped SNPs. Significant sequence homology was found in pine Gene Index database for 160 out of the 171 mapped SNPs in both parental linkage maps [83] (Additional file 5). Sequence homology was found for several species with the top hit homologies for Picea sitchensis, Pinus taeda, Pinus radiata, Vitis vinifera and Picea glauca (Additional file 5). Nine sequences over the 171 sequences showed no match with the InterPro database [96] and we did not find any GO term for seven sequences. Thus, a total of 144 sequences were annotated, 132 of them for molecular function and 101 sequences with annotation for molecular function levels 2 or 3. Most of the mapped SNPs were associated to cDNA belonging to the GO terms: binding, catalytic activity and hydrolase activity (Additional file 6). As expected, SNPs belonging to the same contig reported identical GO annotation terms. However, our results could not confirm statistically (Kolmogorov-Smirnov test for two populations not significant, data not shown) if neighboring SNPs belonging to different genes or contigs exhibited GO terms for the same molecular function. Denser genetic maps with deeper functional annotation are required to evaluate if genes with similar functions are clustered or not.

The comparison of our annotated linkage maps and linkage maps reporting QTL information revealed candidate genes for several QTLs for wood properties or isotopic composition of C13 (δC13) [61,62,64]. δC13 is a character closely related with water use efficiency [97]. In our study, SNPs annotated for water-stress inducible proteins, AQUAPORINs and DEHYDRINs were positioned in the same region as QTLs for δC13 [62] (Table 4). This outcome reinforces the hypothesis that the genomic regions identified by QTL analysis [62] might play a key role in the genetic control of water use efficiency. Also SNPs associated with a CELLULOSE SYNTHASE CESA3, a PEROXIDASE (enzyme involved

### Table 3 Heterozygosity

| Marker | C14 Poly. | C14 Mono. | C14 Heterozygosity (%) | C15 Poly. | C15 Mono. | C15 Heterozygosity (%) |
|--------|-----------|-----------|------------------------|-----------|-----------|------------------------|
| SSR    | 9         | 31        | 0.23                   | 10        | 30        | 0.25                   |
| SAMPL  | 133       | 191       | 0.41                   | 251       | 191       | 0.57                   |
| SNP    | 150       | 672       | 0.18                   | 96        | 726       | 0.12                   |

Percentage of heterozygosity calculated as the ratio of polymorphic markers segregating in each parental map over total (polymorphic – Poly. and monomorphic- Mono.-) markers.

Heterozygosity estimates for SAMPLs were calculated for the first primer combination tested in the family.

### Table 4 Co-localizations of SNPs and QTLs

| SNP ID | Sequence description | e-value | LG | Trait | QTL | Reference |
|--------|----------------------|---------|----|-------|-----|----------|
| m157   | CELLULOSE SYNTHASE   | 1.5e-48 | 3  | wood  | Pot et al. [64] |
| m264   | PEROXIDASE           | 2.9e-147 | 8  | wood  |     |
| m941   | ENDO- BETA-XYLANASE A-LIKE | 0 | 4  | wood  |     |
| m1542  | WATER-STRESS INDUCIBLE PROTEIN 1 | 9.5e-26 | 2  | δC13  | Brendel et al. [62] |
| m1543  | WATER-STRESS INDUCIBLE PROTEIN 3 | 7.3e-34 | 2  | δC13  |     |
| m426   | WATER DEFICIT INDUCIBLE LP3-LIKE PROTEIN | 3.2e-62 | 3  | δC13  |     |
| m295   | AQUAPORIN            | 9.4e-3  | 5  | 5  | δC13  |     |
| m965   | THYLAKOID LUMENAL 19 KDA PROTEIN | 7.1e-110 | 5  | ring growth |     |
| m712   | DEHYDRIN 9 PROTEIN   | 1.2e-71 | 8  | δC13  |     |
| m716   | DEHYDRIN 2           | 1.9e-43 | 8  | δC13  |     |
| m859   | THYLAKOID LUMENAL PROTEIN | 9.8e-145 | 12 | δC13  |     |

Co-localization of mapped SNPs with QTLs detected in previously published maps of P. pinaster. δC13 stands for isotopic composition of C13. Wood stands for wood chemical composition and fibre properties.
in lignin polymerization [98]) and a **ENDO-1,4BETA-XYLANASE A-LIKE** gene (Additional file 5) co-localized with QTLs for wood chemical composition and fiber properties [64] (Table 4). This result increases the evidence of function assigned to these genes and has special relevance when we consider that orthologous QTLs for wood properties were also found in other *Pinus* species [61]. This finding highlights the importance of developing functional genetic linkage maps to be used as useful tools to look for favorable allelic variants to be implemented in MAS.

**Conclusions**

Our study demonstrates the importance of developing different genetic backgrounds in order to generate an accurate consensus linkage map of the same species. Comparative mapping is a key process to facilitate the understanding of genome organization and evolution in conifers. For that purpose it is essential to correctly identify orthologous versus paralogous genes. New efforts in detecting orthologous markers as well as progress in sequencing conifer genomes will improve comparative mapping studies in the future. Here we also confirm the importance of developing functional genetic linkage maps, especially when working with breeding populations for its future application in MAS for traits of interest.

**Additional files**

- **Additional file 1:** Sequence of contigs and genes containing the analyzed SNPs. Nucleotide sequences in fasta format of contigs and genes containing the polymorphic SNPs identified in the mapping progenies. All sequences are the consensus sequence obtained from different individuals except for genes LP3-1 and LP3-3 which sequences obtained for one individual.
- **Additional file 2:** Standard error of recombination frequency. Representation of the standard error of recombination frequency for markers segregating 1:1 in the F1 mapping populations C14xC15 (N=106) and C15xC14 (N=55). Standard error calculated following Ritter et al. [99].
- **Additional file 3:** Polymorphic SNPs. List of the polymorphic SNPs in mapping progenies and associated features.
- **Additional file 4:** Distorted markers. Number of total (assigned, ungrouped and excluded) markers showing distortion in their segregation rates at p ≤ 0.01 per LG obtained for C14 and C15 linkage maps constructed by merging two reciprocal crosses: C14xC15 and C15xC14.
- **Additional file 5:** BLAST analysis of SNP-sequences. Results of BLASTN search in pine databases of the Gene Index Project and GeneBank for mapped SNPs.
- **Additional file 6:** GO annotation for mapped SNPs. Sequence distribution of GO terms for molecular function.

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

MdM: Analyzed the quality of SNP genotyping results, constructed the linkage maps and wrote the first draft of the manuscript. NdM, MAG and LD: Carried out SSRs, EST-Ps and SAMPAL analyses. DSG: Handled and managed plant material. EC and CP: helped MdM in analyzing SNPs from de 1,536 SNP array. ELC: Carried out DNA extraction. CC: Performed functional annotation. JAC: Participated in the construction of linkage maps, MTC, CP and IA: Designed experiments and coordinated projects. All authors contributed to writing the article and approved the final manuscript.

**Acknowledgements**

MdM was supported by an FPI-INIA PhD fellowship granted by the “Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria”. Es was supported by an FPI PhD fellowship granted by the “Ministerio de Ciencia e Innovación” (MICINN). DS-G was supported by the “Juan de la Cierva” program granted by MICINN. This work was supported by funding from the Spanish MICINN (AGL2009-10496, MAPINSEQ), the transnational Plant-KBBE program (PLE2009-0016, SUSTAINPINE), and the European Commission (FP7-211868, NOVELTREE). The authors are very grateful to S. González-Martínez for designing the 384 SNPlex and discussing quality of SNP genotyping results; Ricardo Alía and Salustiano Iglesias for providing seeds and land resources to install the mapping progenies, respectively. We also would like to thank two anonymous reviewers for their helpful comments that have improved the manuscript.

**Author details**

1INIA-CIFOR, Departamento de Ecología y Genética Forestal, Carretera de la Corona Km7.5, Madrid 28040, Spain. 2Unidad Mixta de Genómica y Ecología Forestal, INIA/UPM, Madrid, Spain. 3ETSIM, Departamento de Biotecnología, Ciudad Universitaria s/n, Madrid 28040, Spain. 4INRA, UMR1202 BIOECO, Cestas F–33610, France. 5Université de Bordeaux, UMR1202 BIOECO, Talence 33170, France.

**Received:** 9 June 2012 **Accepted:** 26 September 2012 **Published:** 4 October 2012

**References**

1. Blanco E, Casado MA, Costa M, Escobar R, García-Anton M, Génova M, Gómez-Manzaneque A, Gómez-Manzaneque F, Moreno JC, Moña C, Regato P, Sainz H: Los bosques ibéricos. Una interpretación geobotánica. 4th edition. Barcelona: Planeta; 2005.
2. Fernández M, Gil L, Pardos JA: Effects of water supply on gas exchange in *Pinus pinaster* Ait. provenances during their first growing season. *Ann For Sci* 2000, 57:9–16.
3. Sánchez-Gómez D, Majada J, Alía R, Feito I, Aranda I: Intraspécific variation in growth and allocation patterns in seedlings of *Pinus pinaster* Ait. submitted to contrasting watering regimes: can water availability explain regional variation? *Ann For Sci* 2010, 67:505.
4. Aranda I, Alía R, Ortega U, Dantas AK, Majada J: Intra-specific variability in biomass partitioning and carbon isotopic discrimination under moderate drought stress in seedlings from four *Pinus pinaster* populations. *Tree Genet Genomes* 2010, 6:169–178.
5. Mariette S, Chagné D, Lézец C, Pastuszka P, Raffin A, Plomion C, Kremer A: Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity* 2001, 86:469–479.
6. Ribeiro M, Mariette S, Vendramin G, Szmidt A, Plomion C, Kremer A: Comparison of genetic diversity estimates within and among populations of maritime pine using chloroplast simple-sequence repeat and amplified fragment length polymorphism data. *Mol Ecol* 2002, 11:869–877.
7. González-Martínez SC, Mariette S, Ribeiro MM, Burbán C, Raffin A, Chambel MR, Ribeiro CAM, Aguas A, Plomion C, Alía R, Gil L, Vendramin GG, Kremer A: Genetic resources in maritime pine (*Pinus Pinaster* Aiton): molecular and quantitative measures of genetic variation and differentiation among maternal lineages. *For Ecol Manag* 2004, 197:103–115.
8. Eveno E, Collada C, Guevara MA, Léger V, Soto A, Díaz L, Léger P, González-Martínez SC, Cervera MT, Plomion C, Garnier-Géré P: Contrast patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. *Mol Ecol* 2008, 17:417–437.
25. Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB: Population genetic structure in a Mediterranean pine (Pinus pinaster Ait.): a comparison of allozyme markers and quantitative traits. Heredity 2002, 89:199–206.

26. Arrabal C, García-Vallejo MC, Cadahía E, Cortijo M, de Simón BF: Comparing two characteristics of Pinus pinaster by their terpene and acid patterns in needles. Plant Syst Evol 2012, 288:511–522.

27. Devey ME, Nason N, Alia R, Uren TL, Garza JL, Pardos J: The development of high resi yielders of Pinus pinaster Ait. International Journal of Forest Genetics 2001, 8:271–278.

28. Pavy N, Bousquet J, Meirmans P, Ritland K, Isabel N: QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. BMC Genomics 2011, 12:415.

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

30. Plomion C, Yani A, Marpeau A: Genetic determinism of delta 3-carene in maritime pine using RAPD markers. Genome 1996, 39:1123–1127.

31. Pelgas B, Bousquet J, Meirmans P, Ritland K, Isabel N: QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. BMC Genomics 2011, 12:415.

32. Remington D, Whetten R, Liu BH, O’Malley D: Construction of an AFLP genetic map with nearly complete genome coverage in Pinus taeda. Theor Appl Genet 1999, 98:1279–1292.

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

34. Gonzalez-Martinez SC, Krutovsky KV, Neale DB: Tree-forest population genomics and adaptive evolution. New Phytol 2006, 170:227–238.

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

36. Tani N, Takahashi T, Iwata H, Muki Y, Ujino-Ihara T, Matsumoto A, Yoshimura K, Yoshimura H, Mura M, Nagasaka K, Tsumura Y: A consensus linkage map for Sugi (Cryptomeria japonica) from two pedigrees, base on microsatellites and Expressed Sequence Tags. Genetics 2003, 165:1551–1568.

37. Plomion C, Hurme P, Frigero J, Roidoli M, Pot D, Plonneau C, Avila C, Gallardo F, David H, Neutelings G: Developing SSR markers in two Pinus species. Mol Breeding 1999, 5:21–31.

38. Echt C, Sah S, Krutovsky K, Wimalanathan K, Epeleding L, Liang C: An annotated genetic map of loblolly pine based on microsatellite and cDNA markers. BMC Genomics 2011, 12:17.

39. Neale DB: Genomics to tree breeding and forest health. Curr Opin Genet Dev 2007, 17:359–344.

40. Cho RJ, Mindrinos M, Richards DR, Sappington RI, Anderson M, Drenkard E, Dewdney J, Reuber TL, Stammer M, Fedorin MP: Genome-wide mapping with biallelic markers in Arabidopsis thaliana. Nat Genet 1999, 23:203–207.

41. Lepoittevin C, Frigero JM, Gerner-Geré P, Salin F, Cervera MT, Vornam B, Hanvegnt L, Plomion C: In Vitro vs In Silico Detected SNPs for the Development of a Genotyping Array: What Can We Learn from a Non-Model Species? PLoS ONE 2010, 5:e11034.

42. Pavy N, Parsons J, Plomeuse P, Mackay J, Bousquet J: Automated SNP detection from a large collection of white spruce expressed sequences: contributing factors and approaches for the categorization of SNPs. BMC Genomics 2006, 7:174.

43. Eckert AJ, Pande B, Ersoz E, Wright MT, Rashbrook WN, Nicolet CM, Neale DB: High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (Pinus taeda L.). Tree Genet Genomes 2009, 5:225–234.

44. Chancerel L, Lepoittevin C, Le Provost G, Lin Y-C, Karlsson-Crome J, Eckert A, Wegryn J, Zelenika D, Boland A, Frigerio J-M, Chaumeil P, Gerner-Geré P, Bonne C, Grivet D, Gonzalez-Martinez S, Pouze P, Van de Peer Y, Neale D, Cervera M, Kemerer A, Plomion C: Development and implementation of a highly-multiplexed SNP array for genetic mapping in maritime pine and comparative mapping with loblolly pine. BMC Genomics 2011, 12:1586.

45. Pavy N, Pelgas B, Beaussiele S, Blais S, Gagnon F, Gosselin L, Lamothie M, Isabel N, Bousquet J: Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: application to the large and unsequenced genomes of white spruce and black spruce. BMC Genomics 2008, 9:21.

46. Moniguchi Y, Ujino-Ihara T, Futamura N, Saito M, Ueno S, Matsumoto A, Tani N, Taira H, Shinohara K, Tsumura Y: The construction of a high-density linkage map for identifying SNP markers that are tightly linked to a large and unsequenced genomes of white spruce and black spruce. BMC Genomics 2008, 9:21.

47. Gustafsson M, Cervera MT, Eto A, Collada C, Plomion C, Savolainen O, Neale DB, Gonzalez-Martinez SC: Genomics applied to the study of adaptation in pine species. Investigacion Agraria. Sistemas y Recursos Forestales 2005, 14:292–306.

48. Elsk C, Williams C: Retroelements contribute to the excess low-copy-number DNA in pine. Mol Gen Genet 2003, 264:47–55.

49. Changné D, Lalanne C, Madur K, Kumar S, Frigero JM, Krier C, Decrocq S, Savoureux A, Bou-Dagher-Kharraz M, Bertocchi E, Brach J, Plomion C: A high density genetic map of maritime pine based on AFLPs. Ann For Sci 2002, 59:627–636.

50. Grotkop E, Rejmenak M, Sanderson MJ, Rost TL: Evolution of genome size in pines (Pinus) and its life-history correlates: supertree analyses. Evolution 2004, 58:1705–1729.

51. Bahram N, Darmayvel C: Linkage relationships of loci controlling protein amounts in maritime pine (Pinus pinaster Ait.). Heredity 1989, 63:267–274.
97. Farquhar GD, Richards RA: Isotopic composition of plant carbon correlates with water-use efficiency of wheat genotypes. Aust J Plant Physiol 1984, 11:539–552.
98. Boerjan W, Ralph J, Baucher M: Lignin biosynthesis. Annu Rev Plant Biol 2003, 54:519–546.
99. Ritter E, Gebhardt C, Salamini F: Estimation of recombination frequencies and construction of RFLP linkage maps in Plants from crosses between heterozygous parents. Genetics 1990, 125:645–654.

doi:10.1186/1471-2164-13-527
Cite this article as: de Miguel et al.: Annotated genetic linkage maps of Pinus pinaster Ait. from a Central Spain population using microsatellite and gene based markers. BMC Genomics 2012 13:527.