Linkage map construction and QTL analysis for *Betula platyphylla* Suk using RAPD, AFLP, ISSR and SSR

By Kaixuan Zhang, Dan Wang, Chuanping Yang, Guanjun Liu, Guifeng Liu, Hanguo Zhang, Lianlian and Zhiang Wei

Key Laboratory of Forest Tree Genetic Improvement and Biotechnology (Northeast Forestry University), Chinese Ministry of Education

(Received 18th November 2010)

Abstract

A linkage map for *Betula platyphylla* Suk was constructed based on RAPD, ISSR, AFLP and SSR markers by a pseudo-testcross mapping strategy. A F1 segregating population including 80 progenies was obtained from the cross between two superior trees selected from Qinghai and Wangqing provenance, respectively. The paternal map was constructed with 282 markers consisting of 14 major and 15 minor (5 triplets and 10 doubles) linkage groups and spanning 1131 cM at an average distance of 4.0 cM between adjacent markers. The maternal map has 277 markers consisting of 15 major and 8 minor (5 triplets and 3 doubles) groups covering 1288 cM at an average distance of 4.6 cM between adjacent markers. In the same pedigree we investigated the association of genetic markers with seedling stem height and circumference. The composite interval mapping was used to detect the number of quantitative trait loci and their position on the genetic linkage maps. Three QTLs (one on the male map and two on the female map) were found explaining 13.4%, 17.5% and 18.8% of the trait variation, respectively.

Key words: silver birch, linkage map, RAPD, AFLP, ISSR, SSR, QTL.

Introduction

Silver birch (*Betula platyphylla* Suk) is widely distributed in northeast, north, northwest and southwest forest areas of China. The largest birch stand volume is found in northeast forest region of China. With the characteristics of rapid growth, strong adaptability, wide distribution, high content of cellulose with low content of lignin and white color (Li et al., 1995), silver birch is one of the most important commercial tree species for paper-making, furniture and plywood in China and also the pioneer species in north secondary broad-leaved forest of China.

A significant amount of researches have been done with Chinese silver birch, including green wood cutting propagation (Zhan et al., 1994), provenance division and genetic variation (Zhu et al., 2001; Jiang et al., 2001), intensive breeding (Yang et al., 2004), transgenic study (Zhan et al., 2006) and hybridization breeding (Li et al., 2006), fiber length experiment (Wei et al., 2006). Molecular markers such as RAPDs, AFLPs, ISSRs, and SSRs have been used extensively for the preparation of linkage maps of a number of tree species (Grattapaglia and Sederoff, 1994; Remington et al., 1999; Arcade et al., 2000; Hayashi et al., 2001). Owing to its advantages over the traditional breeding mode, genetic mapping emerges as a new tool to assist traditional tree breeding in the identification of quantitative trait loci (Bradshaw and Stettler, 1995).

Linkage maps have been recently constructed for several tree fruit and nut crops including pear (Yamamoto et al., 2002), apricot (Vilanova et al., 2003; Lambert et al., 2004), citrus (Sanekai and Moore, 2001), macadamia (Peace et al., 2003), apple (Conner et al., 1997), walnut (Fjellstrom and Parfitt, 1994) and pecan (Sudheer et al., 2005). Like many other tree fruit and nut crops, silver birch is an out-bred heterozygous forest tree. F2 or backcross population for mapping is generally not available due to the constraints of inbreeding depression and long generation time (Grattapaglia and Sederoff, 1994). Thus, a F1 mapping population has been created for hybrid silver birch allowing the analysis of linkage according to the pseudo-testcross mapping strategy. Pekkinen et al. (2005) presented the first silver birch linkage map with 82 AFLP and 19 SSR markers that covered 1561 cM. A mid-density linkage map based on 307 RAPD markers for *Betula pendula* Roth and *Betula platyphylla* Suk was constructed by Jiang (2007). To date, there is little information available on linkage map for Chinese silver birch, especially because the molecular markers resources are still limited foe the species. QTLs have been mapped for growth and yield, wood quality and bud phenology traits in forest trees, such as poplar (Bradshaw and Stettler, 1995), pine (Plomion et al., 1996; Sewell et al., 2000, 2002) and fir (Jermstad et al., 2001), while there is no QTL study performed in silver birch. The aim of this present work was to construct a high density linkage map of Chinese silver birch using a combination of genetic markers as well as locate the growth QTLs.

Materials and Methods

Plant materials

‘Wangqing’ provenance and ‘Qinghai’ provenance are two different provenance districts with the largest genetic distance among all the studied provenances (Jiang J et al., 2001). ‘Wangqing’ is considered the optimal provenance for *Betula platyphylla* Suk on the
seedling stage with better seedling characters than ‘Qinghai’ (Jiang J et al., 1999). The male and female parents of this study were originated from the ‘Qinghai’ and ‘Wangqing’ regions, respectively. A controlled cross was performed in the experimental field of Northeast Forestry University Experimental Forest Farm between the two parents. After germination, the seedlings were maintained in the greenhouse. The mapping population comprised 80 individuals of a full-sib F1 progeny derived from a cross between the two parents in the nursery.

**DNA extraction**

Total genomic DNA was extracted from fresh leaves using the Universal Genomic DNA Extraction Kit (TaKaRa, Japan) and stored at −20°C. DNA quality was checked by 0.8% gel electrophoresis. The purified DNA was quantified at 260 nm using a spectrophotometer, diluted at a concentration of 50 ng/μL in sterile water and stored at −20°C.

**RFLP analysis**

The random 10-mer primers for RAPD analysis were purchased from Sangon Biological Engineering Technology and Service Co. Ltd (Shanghai, China). A total of 1200 RAPD primers were screened using two parents and four F1 individuals. The 20 μL PCR reaction mix consisted of 12.3 μL H2O, 2 μL of a 50 ng/μL of template DNA, 0.5 μL dNTP (10 mmol/L), 2 μL MgCl2 (25 mmol/L), 1 μL primers (10 pmol/L), 2 μL 10× buffer and 0.2 μL Taq polymerase (5 U/μL, Fermentas). PCRs were performed with an initial 4 min at 94°C followed by 40 cycles at 94°C for 30 s, 36°C for 30 s and 72°C for 2 min. The last cycle was followed by a final extension at 72°C for 7 min. The reaction products were separated on 2.0% agarose gels and band sizes were estimated using a DL-2000 DNA Marker (MBI).

**AFLP analysis**

AFLP analysis was performed according to the system set up by Lianlian et al. (2007) based on a previous study (Vos et al., 1995). Total genomic DNA (145 ng) was digested with a combination of EcoRI and MseI restriction enzymes and specific double-stranded adapters were ligated to the ends of fragments. Pre-amplification was performed with 1-bp extension primers (MseI primer+C; EcoRI primer+A). Selective amplification of restriction fragments was conducted using primers with 3-bp selective nucleotides. A set of 64 primer combinations were screened by two parents and four F1 individuals to generate AFLP markers with a sample of 80 F1 progenies. The selective amplification products were separated on 6% polyacrylamide gel and the bands were visualized by silver staining.

**ISSR analysis**

One hundred and fifty primers were screened for polymorphism using the parents and four F1 individuals. The ISSR analysis was carried out in a 20 μL volume containing 0.3 μL Taq DNA polymerase (Fermentas), 2 μL 10× Buffer, 1 μL genomic DNA (50 ng/μL), 1.5 μL MgCl₂ (25 mmol/L), 1.2 μL dNTP (2.5 mmol/L), 0.6 μL oligonucleotide primer (50 pmol/L). DNA amplification was performed under the following conditions: an initial denaturation step at 94°C for 3 min, and 38 cycles of 94°C for 30s, 48°C for 30s, and 72°C for 90s, followed by a final extension at 72°C for 7 min. The amplified products were separated on 2.0% agarose gels and the bands were visualized with ethidium bromide staining.

**SSR analysis**

The sequences of SSR primer pairs were obtained from the sequence analysis of 2548 ESTs in Betula pendula deposited in NCBI (http://www.ncbi.nlm.nih.gov/dbEST/index.html). Three hundred and six SSRs were distributed in 260 ESTs out of 2548 ESTs sequences. The software Primer5 was used to design 176 pairs of SSR primers, and the software SSR primer was applied to design 100 pairs of primers. The other primers pairs were designed based on previous studies (Ogyu et al., 2003; Kulju et al., 2004; Truong et al., 2005). A total of 330 primers pairs were screened against two parents and four F1 individuals. SSR amplification was carried out according to methods of Wang et al. (2008). The SSR products were resolved on 6% denaturing polyacrylamide gel and visualized by silver staining.

**Linkage map construction**

The presence or absence of polymorphic bands was evaluated in the F1 individuals and two different types of segregation models were classified: testcross markers (Grattapaglia and Sederoff, 1994) with a 1:1 segregation and intercross markers (Verhaeegen and Plomion, 1996) with a 3:1 segregation. A χ²-test was performed for each locus to determine whether the inherited alleles of the offspring were in compliance with the Mendelian segregation ratios. Markers segregating in a Mendelian fashion (χ²≤χ²0.01) or deviating only slightly from expected ratio (χ²≤χ²=0.1) were used for map construction.

Mapmaker 3.0 (Lander et al., 1987) was employed for preliminary grouping of marker loci. Intercross markers were given an approximate location on both maps using JoinMap 3.0 (Van Ooijen and Voorrips, 2001). Map distance in CenitMorgans was calculated using Kosambi’s mapping function (Kosambi, 1994) and the linkage threshold parameters were set as a minimum LOD threshold of 4.0 and maximum recombination frequency of 0.4. The linkage map was drawn using the Map Chart 2.1 (Voorrips, 2002).

**Map length and coverage estimates**

The method-of-moments type estimator (Huibert et al., 1988) as proposed in ’method 3’ by Chakravarti et al. (1991) was used to estimate genome length (G) with the formula

\[
G = N(N-1)X/K
\]

where N is the total number of markers in major groups, X is the maximum distance between two adjacent markers (cM), and K is the number of marker pairs linked at the same minimum LOD score. A minimum LOD score of 4.0 was chosen to estimate the genome length.
Quantitative traits and QTL mapping

Quantitative traits were measured on the same progeny utilized for the construction of the linkage map. First-year seedling stem height (SH) and stem circumference (SC) above ground level were measured in 2006.

QTL analysis was performed using QTL Cartographer 2.5 and Composite Interval Mapping (CIM) analysis (Zeng, 1994) was performed on maternal map. Significant thresholds (with overall P = 0.05) were established using a permutation test (1,000 runs) (Churchill and Doerge, 1994).

Results

Molecular markers

Out of the 1200 RAPD arbitrary primers screened, 223 (18.6%) were selected to produce 440 polymorphic fragments with a mean of 2.0 markers per primer. The rest of the primers did not yield to any amplified polymorphic product. Four hundred and five loci (206 for the female parent and 199 for the male parent) were identified segregating in a testcross, 1:1 configuration. Thirty five loci segregating in an intercross configuration with a 3:1 ratio were identified (Table 1).

Thirty one (48.4%) out of 64 AFLP primer combinations were selected to obtain polymorphic fragments. 311 AFLP markers were heterozygous in one parent and null in the other, 43 showed a 3:1 segregation and 32 were anomalous loci which were absent in both parents while segregating in the progenies. The number of polymorphic AFLP markers produced per primer combination ranged from 4 to 20, with a mean of 11.5 markers per primer combination (Table 1).

Out of 150 ISSR primers examined, only 46 (30%) ISSR primers yielded efficient amplification and detection of polymorphism between the parental genotypes. They produced a total of 95 ISSR markers (55 in male parent and 40 in female parent) segregating 1:1 in the progeny (Table 1) and 7 markers fitted a 3:1 segregation ratio in the progeny. The average number of markers produced per primer was 2.2. Fragment sizes ranged from 200 bp to 1300 bp.

A total of 330 SSR primer pairs were screened. Twenty nine of these showed at least one polymorphic band, detecting 58 SSR markers. Out of the 58 SSR loci, 8 loci showed co-dominant segregation, 50 loci segregated 1:1 and were mapped with a high degree of confidence (Table 1). There were 3 markers showing anomalous patterns.

An analysis of genotype frequencies showed that 124 (13%) of 954 markers suffered from slight segregation distortion ($\chi^2_{1}=\chi^2_{2}-\chi^2_{0.01}$). A total of 46 (4.8%) markers were highly distorted ($\chi^2_{2}>\chi^2_{0.001}$) and were discarded prior to the construction of the linkage map.

Map construction

Two separate framework maps were built from the 908 polymorphic markers after the exclusion of 46 markers showing highly significant levels of distortion. Of these 908 markers, 784 markers segregated in a Mendelian fashion and 124 showed a slight deviation from the Mendelian fashion.

The paternal genetic linkage map consisted of 317 loci that were subdivided into 14 linkage groups (LGs) with a LOD threshold of 4.0. Each group contained a minimum of five markers (Figure 1A), and additional 35 markers were distributed as 5 triplets and 10 doublets. 136 testcross markers remained unlinked. The length of the LGs varied from 39.4 cM to 197.9 cM (mean 80.8 cM), comprising 5-63 loci per LG (mean 20.1). The average distance between adjacent markers was 4.0 cM and the largest gap (28.6 cM) was found on LG13 (Table 2). 22 intercross markers were distributed in six LGs (LG1, LG2, LG3, LG4, LG11, LG14).

For maternal linkage map, 277 markers were placed in 15 linkage groups, and 21 markers were arranged into 5 triplets and 3 doublets (Figure 1B) at a LOD threshold of 4.0. 154 testcross markers remained unlinked. The length of the linkage groups varied from 37.4 cM to 179.4 cM with an average of 85.9 cM. The number of markers for each linkage group varied from 4 to 36 with an average of 18.5. One marker was found on an average of 4.6 cM with the largest gap (25.7 cM) being on LG15 (Table 2). 23 intercross markers were distributed in 11 LGs.

Estimated map length

Map lengths were estimated as 1293.0 cM for the male parent and 1455.3 cM for the female parent. Considering that the obtained linkage groups spaned 1131.0 cM and 1288.0 cM in the male and female maps, respectively, our framework map covered 87.5% of the male genome and 88.5% of the female genome.

QTL analysis

By Composite Interval Mapping [P < 0.05] set by the permutation test] and Multiple Interval Mapping we found three QTLs under the likelihood ≥3.3: one on the

<table>
<thead>
<tr>
<th>Table 1. – Total numbers of RAPD, AFLP, ISSR and SSR markers obtained from parents and F1 individuals.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAPD</td>
</tr>
<tr>
<td>Number of testcross markers (1:1) present in “Qinghai”</td>
</tr>
<tr>
<td>Number of testcross markers (1:1) present in “Wangqing”</td>
</tr>
<tr>
<td>Number of intercross markers (3:1)</td>
</tr>
<tr>
<td>Number of co-dominant markers</td>
</tr>
<tr>
<td>Total number of markers produced</td>
</tr>
</tbody>
</table>
Figure 1. The linkage maps of a 'Qinghai' male (A) and a 'Wangqing' female (B) parents. Marker names are shown on the right of each linkage group with map distances (in cm) on the left. Locus nomenclature loci are shown in the form e.g. 74R-800 where 74R refer to the primer codes and 800 polymorphic bands molecular weight. Markers showing significant levels of segregation distortion are indicated by underlining the marker name and the '3:1' loci are indicated by asterisks. QTL bars in the linkage group positions are indicated by SH (stem height) and SC (stem circumference).
female map and two on the female map (Table 3). Two QTLs affecting seedling stem circumference were located in different linkage groups (LG3-M between AFLP markers E7M6-900 and E8M7-240 and LG9-F between AFLP marker E1M3-1860 and SSR marker A23-200). One QTLs affecting seedling height was located in the same linkage group of the female map (LG9-F) but on a different position.
Discussion

Marker identification

In this study, AFLP, RAPD, SSR and ISSR technologies allowed us to generate a large number of molecular markers and a sufficient coverage of the genome for silver birch. Previously studies have shown that RAPD and ISSR are simple and efficient methods to generate a large number of molecular markers suitable for preliminary genetic mapping in a short time (Arcade et al., 2000; Casasoli et al., 2001; Rosa et al., 2003; Jiang et al., 2007). AFLPs have mainly been used as markers for map construction, since they can quickly generate a large number of polymorphisms, produce a great amount of scorable loci in a single assay and generally give a significant coverage of the genome (Hanley et al., 2002; Beedanagari et al., 2005; Ukrainetz et al., 2008).

In our study, the clustering of AFLP markers in some portions of the map is similar to what has been obtained from previous studies (Wu et al., 2000; Cavalcanti et al., 2007). There are several possible explanations for this phenomenon. For instance, the restriction sites of EcoRI-based AFLP markers may not distribute uniformly throughout the genome, especially in regions around centromeres and telomeres (Bovin et al., 1999). Such regions have a markedly reduced frequency of recombination during meiosis, so this feature may account for the tendency of markers to cluster (Luckaszewski and Curtis, 1993). Moreover, although AFLP amplification sites are almost randomly located across the genome, their distribution along most linkage groups is typically uneven. SSR markers are more suitable for the amplification of centromeres and telomere. Furthermore, EST-SSR markers we used may be less vulnerable than genomic DNA-derived SSR markers regarding the competition effects and high levels of mutation around SSR loci, and hence more suitable for homologues detection (Jones et al., 2003).

Marker segregation

Approximately 17.8% of the molecular markers showed a segregation distortion, 124 loci deviated only slightly from Mendelian segregation, while 46 loci showed higher segregation distortion and were discarded prior to the linkage mapping. The percentage of significant distorted markers detected here (13%) was similar to that observed in linkage analyses of pine (14–15%; Kubiak et al., 1995), eucalyptus (15%; Ques et al., 1998), willow (18%; Hanley et al., 2002) and artichoke (14%; Lanteri et al., 2006). Segregation distortion can be due to various reasons, such as statistical bias, genotyping and scoring errors (Plomion et al., 1995). Biological features, including chromosome loss, viability or lethal genes, genetic isolating mechanisms and genetic load, also contribute to segregation distortion (Bradshaw and Stettler, 1994). In the previous work, 35 significant distorted markers were used in the two maps to prevent the exclusion of a significant part of a linkage group (Cervera et al., 2001), while Venkateswarlu et al. (2006) did not consider the distorted markers in the linkage map of mulberry.

Map construction

All the 954 markers were classified into three different groups on the basis of the segregation model (Table 1): (1) Testcross markers segregating 1:1 i.e. heterozygous in one parent and homozygous null in the other. A total of 861 (90.3%) testcross markers were identified (433 for the female parent and 428 for the male parent).
(2) Intercross markers that were heterozygous in both parents and segregated in a 3:1 ratio. Eighty five (8.9%) intercross markers were identified. (3) Eight (0.84%) co-dominant markers were generated by four primer combinations (P61, P79, P96, P114).

Intercross markers are less informative than testcross markers because the dominant phenotype is mixed with three indistinguishable genotypes (+ + , + – , – – ) (Crespel et al., 2002). Thus, the parent of an allele can be determined unambiguously only in the homozygous recessive progeny. These acting accessory markers have been only used to identify homology between parental maps (Casasoli et al., 2001; Scalfi et al., 2004; Sudheer et al., 2005). If two maps share enough common markers, they can then be combined into a single integrated linkage map (Connor et al., 1997; Peace et al., 2003). The co-dominant SSRs have also been used to combine two maps into a single integrated linkage map (Powell et al., 1996). In the present study, although several intercross markers were mapped to linkage groups in both parental maps, they are not enough for the maps to be merged. Integrating the paternal and maternal maps is difficult due to the low percentage of the intercross (8.9%) and co-dominant (0.84%) markers that could be mapped in comparison with testcross markers. We have previously identified a total of 42 anomalous sites in RAPD, AFLP, SSR markers that can be used as the ‘bridges’ to align homologous linkage groups between the two maps (Zhang et al., 2000).

In the paternal linkage map, three LGs (LG11, LG13, LG14) are composed solely of RAPD loci. There is some clustering of the AFLP markers in LG1 and LG5. In the maternal linkage map, only LG15 is composed entirely of RAPD loci. There is also clustering of the AFLP markers in LG1 and LG2. Clustering of AFLP and RAPD markers sporadically observed in both maps is similar to the results of previous studies (Crespel et al., 2002; Bednarek et al., 2003; Sudheer et al., 2005; Cavalcanti et al., 2007).

**QTL analysis**

For QTL mapping in forest species, the problem is the unknown linkage phase. Grattapaglia and Sederoff (1994) came up with the pseudo-testcross strategy, whereby we duplicate the markers to test linkages with all possible phases. In our study we found three QTLs and two of them were clustered on linkage group 9 on the female map. Clusters of QTLs were observed also for different species of the Pinus (Sewell et al., 2000), Fagus (Scalfi et al., 2004), Gossypium (Lin et al., 2005), and Populus (Anne et al., 2008). The clustered QTLs could indicate the possible presence of only one QTL having pleiotropic effect on different traits.

On the male map, we did not find QTL for seedling height. In hybrid pine and beech Kaya et al. (1999) and Scalfi et al. (2004) did not find the QTL in consecutive growing seasons for height. The absence of QTL can be due to the lack of variability for this trait in the parents of the controlled cross progeny or simply due to the small size (80) of our sample. The small size of our sample could also lead to an underestimation of the number of QTLs per trait and to an overestimation of the phenotypic effect associated with each QTL (Beavis, 1995). For accurate QTL analysis for growing traits, LOD was raised to 3.3 (set by the permutation test), even though Grattapaglia et al. (1995) suggested QTL mapping can be done under the likelihood ± 1.

The first QTL study for silver birch is not so accurate because of the small sample size. Hence, with further work more QTLs will be located and become a valuable reference tool for the birch breeding.

**Acknowledgments**

This research was supported by the Specific Basic Science-technology Project (No. 2007FY110400-3), Provincial Key Technologies of Heilongjiang (No. GA06B301-4) and The Graduate Scientific and Technological Innovation Project of Northeast Forestry University (No. gram09). Special thanks to the Key Laboratory of Forest Tree Genetic Improvement and Biotechnology (Northeast Forestry University), Ministry of Education in China for hosting the research group and their critical comments in support of the research. Many thanks also go to Prof. Jiang Jing for her help and support in the realization of the study.

**References**


