Isolation and characterization of microsatellite loci from *Pinus massoniana* (Pinaceae)

Kuo-Hsiang HUNG1,6, Chi-Yung LIN2,6, Chi-Chun HUANG2,6, Chi-Chuan HWANG3,6, Tsai-Wen HSU4, Yau-Lun KUO3, Wei-Kuang WANG2, Cheng-Yu HUNG2, and Tzen-Yuh CHIANG2,*

1Graduate Institute of Bioresources, National Pingtung University of Science and Technology, Pingtung 912, Taiwan
2Department of Life Sciences, Cheng-Kung University, Tainan 701, Taiwan
3Department of Engineering Science, National Cheng-Kung University, Tainan 701, Taiwan
4Taiwan Endemic Species Research Institute, Nantou 552, Taiwan
5Department of Forestry, National Pingtung University of Science and Technology, Pingtung 912, Taiwan

(Received August 19, 2011; Accepted January 6, 2012)

**ABSTRACT.** *Pinus massoniana* is widespread in the central and eastern mainland China, while distributed in Taiwan with a single population. This island population has been declining due to habitat destruction and pine wilt disease. In the study we isolated microsatellite loci from *P. massoniana* for investigating the population structure. Eleven, novel microsatellite markers were developed from *P. massoniana* by using a modified PCR-based isolation of microsatellite arrays (PIMA) method. The number of alleles, observed and expected heterozygosities across loci varied with a range of 2-9, 0.00-0.82 and 0.40-0.83, respectively. The application of these microsatellite markers of *P. massoniana* provides a tool for understanding demography and population structure in Taiwan and mainland China.

**Keywords:** Heterozygosity; Microsatellite; Population structure; Pinaceae; *Pinus massoniana*.

**INTRODUCTION**

*Pinus massoniana* Lamb., a species of sect. *Pinus*, is widespread across the central and eastern mainland China (Richardson and Rundel, 1998). It is frequently used in hedges, windbreak, mine reclamation and timber plantations because of its fast growth rate and delicate wood grain (Richardson and Rundel, 1998). Ecologically, as largely restricted to habitats below 1,500 m in elevations, the species grows in association with *Quercus*, *Cunninghamia*, and *Cryptomeria*, or sometimes in pure stands (Richardson and Rundel, 1998). In contrast to the ecological dominance as an ecological pioneer, colonizing mesic, harsh habitats and competing little with other woody plants in mainland China (Richardson, 1998; Lusk, 2008), on the island of Taiwan, the distribution of a single native population is restricted to the Huoyanshan Nature Reserve in Miaoli County, where the habitat is being degraded. Although the regeneration of *P. massoniana* can occur on the eroded, open habitat, the limited areas seriously constrains its expansion (Chiang, 2008). Recently, the fatal pine wilt disease, carried by the pinewood nematode (*Bursaphelenchus xylophilus*), has also diminished the population size (Huang, 2009), likely leading to the loss of genetic diversity, especially in Taiwan. Ecological studies revealed that *P. massoniana* resources have been decreasing dramatically in mainland China and Taiwan (Guan et al., 2011). For examining the genetic variation in *P. massoniana*, population genetics studies have been conducted by using allozymes (Huang and Zhang, 2000), RAPD fingerprinting (Peng et al., 2003), and organelle DNAs (Zhou et al., 2010). Nevertheless, given characters of being highly polymorphic and abundant, co-dominant inheritance, and analytical simplicity. Microsatellites as DNA markers are more advantageous than the above markers (Morgante and Olivieri, 1993).

As this Taiwanese population decreases in size quickly, strategies to preserve *P. massoniana* are required. For obtaining the necessary information for developing these strategies, microsatellite fingerprinting, with high genetic variability, is an ideal tool to elucidate the population structure and genetic diversity (Freville et al., 2001). The purpose of this study, therefore, is to develop a set of microsatellite markers for *P. massoniana*. While Guan et al. (2011) reported nine microsatellite markers with high variability for *P. massoniana*, SSR markers developed in this study will provide additional resolution to the genetic structuring and diversity in populations of *P. massoniana*. The application of these microsatellite markers of *P. massoniana* provides a tool for conservation genetics in Taiwan and mainland China.
MATERIALS AND METHODS

Samples

In total, 13 individuals of Pinus massoniana from the population at the Huoyanshan Nature Reserve (24°22’ N, 120°43’ E) in Taiwan, and 20 individuals from Mt. Huangshan in Anhui Province (30°07’ N, 118°11’ E) in mainland China were collected. Vouchers were deposited in the herbarium of the Endemic Species Research Institute.

Microsatellite cloning

Genomic DNAs were extracted from dry leaves following a CTAB method (Doyle and Doyle, 1987). The isolation of microsatellites followed a modified PCR-based isolation of microsatellite arrays (PIMA) methodology (Lunt et al., 1999). RAPD-PCR amplifications were performed with a thermal cycler (Bio-Rad, Hercules, CA, USA) in a reaction mixture (50 μL) containing 20-100 ng DNA, 0.2 mM of each dNTP, 2 mM MgCl2, 0.5 U Taq polymerase (Promega, Madison, WI, USA), and 5 pmols of one RAPD primer of 10 bp in length selected from Operon Technologies kits and microsatellite primer (M1 or M2 primer). The PCR programs were as follows: initial denaturing 3 min at 94°C for 1 cycle; 40 cycles of 1 min at 94°C, 1 min at 42°C, 2 min at 72°C; and 10 min at 72°C for an additional extension step.

Fifty RAPD primers (nos. 101-150) and microsatellite primer (M1 primer: 5’-TCTCTCTCTCTCTCA CACACACACAC-3’, or M2 primer: 5’-ACACACACACAC CAGAGAGAGAG-3’) were used to amplify fragments from target species’ genome in separate reactions. PCR products were size-selected to obtain small fragments (ca. 300-800 bp), and the desired fragments were excised from the gel and purified. DNA fragments were ligated into a pGEM T-Easy Vector Systems (Promega), and the reactions were transformed into Escherichia coli. About 1000 clones were screened using a microsatellite primer (M1 or M2 primer) and forward and reverse vector primers (T7 primer: 5’-TAA-TAC-GAC-TCA-CTA-TAG-GG-3’, or SP6 primer: 5’-ATT-TAG-GTG- ACA-CTA-TAG-GG-3’) (Lunt et al., 1999). The PCR programs were as follows: initial denaturing 5 min at 94°C for 1 cycle; 35 cycles of 1 min at 94°C, 1 min at 55°C, 2 min at 72°C; and 10 min at 72°C for an additional extension step. In positive clones with the same microsatellite motif as the primer, PCR electrophoresis would display an additional smaller DNA band that contains microsatellite signal, whereas no amplification was found in negative clones. The plasmid DNAs from the eleven positive clones were purified using the High-Speed Plasmid Mini Kit (Geneaid, Taipei, Taiwan), and sequenced in an Applied Biosystems Model 3730 automated sequencer (Applied Biosystems, Carlsbad, CA, USA). Forward primers were designed according to the nucleotide sequences upstream or downstream of the repetitive DNA using Primer3 (Rozen and Skaletsky, 2000), and reverse primers are designed on the microsatellite sequences themselves.

PCR amplification of microsatellites was performed in a 20 μL volume containing 10 ng of genomic DNA, 0.2 mM dNTP, 2 mM MgCl2, and 5 pmols of each primer. The PCR condition was as follows: 3 min at 94°C; 40 cycles of 30s at 94°C, 30s at primer-specific annealing temperature (Ta) (Table 1), 30s at 72°C, and a final extension step at 5min at 72°C. Electrophoresis was performed in denatured 6% polyacrylamide gels using 10-bp ladder molecular size markers (Invitrogen, Carlsbad, CA, USA) to estimate the allele sizes with ethidium bromide staining.

Data analysis

The observed (Ho), and expected (He) heterozygosities were calculated using the Arlequin program version 3.1 (Excoffier et al., 2005). Tests of Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were conducted using the GENEPOP program version 3.4 (Raymond and Rousset, 1995). Micro-Checker version 2.2.3 (van Oosterhout et al., 2004) was used to estimate the frequencies of null alleles in the microsatellites markers.

RESULTS

Eleven primer pairs were successfully amplified with size matching expectations based on the ranges of initially repetitive DNA fragments. To avoid nonspecific fragments, PCR amplifications have been repeated for at least three times. Besides, for distinguishing fragments with similar sizes, alleles were sequenced. Genotypic data for 11 microsatellite loci were obtained for one population each of mainland and Taiwan (Table 1). Results of the allele number, size range, number of bands per locus are listed in Table 1. The number of alleles across loci ranged from 2 to 5 (with an average of 3.09) in the Taiwanese population. As shown in Table 1, the Ho and He ranged from 0.18 (an average of 0.18) and 0.49-0.83 (an average of 0.61), respectively. In the mainland China population, the number of alleles across loci varied from 1 to 6. Ho and He varied with a range of 0-1.00 (an average of 0.27) and 0.00-0.81 (an average of 0.42), respectively (Table 1). When the data was pooled together, the number of alleles, Ho and He across loci varied with a range of 2-9 (an average of 4.64), 0.00-0.82 (an average of 0.27) and 0.40-0.83 (an average of 0.65), respectively. Significant departures from HWE were detected in all microsatellite loci (P < 0.05) based on the sequential Bonferroni corrections (Rice, 1989). Null alleles, detected by Micro-Checker version 2.2.3, may have occurred in ten loci (P < 0.05), all of which displayed excessive homozygotes, except for the Pin07. No significant linkage disequilibrium (LD) was detected between loci, except for four pairs (Pin01 and Pin02, Pin02 and Pin03, Pin03 and Pin05, and Pin04 and Pin06) (all P < 0.05).

DISCUSSION

It has been generally known that the level of genetic diversity at the molecular markers among populations is...
Table 1. Characteristics of 11 microsatellite primers developed in *Pinus massoniana*. The forward (F) and reverse (R) sequences, repeat type, size of the original fragment (bp), annealing temperature ($T_a$) and the GenBank accession number are given. Shown for each primer pair are the number of alleles ($A$), and mean values of observed ($H_o$) and expected ($H_e$) heterozygosities. The sample size is shown in parentheses.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Repeat motif</th>
<th>Size</th>
<th>$T_a$</th>
<th>GenBank accession number</th>
<th>Taiwan (N=13)</th>
<th>China (N=20)</th>
<th>Total (N=33)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin01</td>
<td>F: CGCAACGAAGTGGAGCTAAT (TC)$_6$(AC)$_8$</td>
<td>241-270</td>
<td>50</td>
<td>JF900332</td>
<td>2</td>
<td>0.00</td>
<td>0.52</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>R: TCTCTCTCTCACCACACACAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin02</td>
<td>F: GTCCCTGGGTTGCGACACTAT (AC)$<em>6$(AG)$</em>{20}$</td>
<td>295-325</td>
<td>60</td>
<td>JF900333</td>
<td>3</td>
<td>0.10</td>
<td>0.63</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>R: ACACACACACACAGAGAGAGAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin03</td>
<td>F: CAACCAGATGTCTCTCCCAAG (AC)$_6$(AG)$_8$</td>
<td>179-250</td>
<td>50</td>
<td>JF900334</td>
<td>3</td>
<td>0.00</td>
<td>0.59</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>R: ACACACACACACAGAGAGAGAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin04</td>
<td>F: CTACGAGCCACAATGTACTTA (TC)$_{30}$ (AC)$_6$</td>
<td>116-170</td>
<td>48</td>
<td>JF900335</td>
<td>2</td>
<td>0.00</td>
<td>0.52</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>R: TCTCTCTCTCACCACACACAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin05</td>
<td>F: AGGTGGTTCAGCAAGATGG (AC)$_6$</td>
<td>224-253</td>
<td>50</td>
<td>JF900336</td>
<td>2</td>
<td>0.00</td>
<td>0.51</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>R: ACACACACACACAGAGAGAGAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin06</td>
<td>F: AAAATCGGGTGCGACAACAGA (TC)$_6$(AC)$_6$</td>
<td>147-153</td>
<td>50</td>
<td>JF900337</td>
<td>2</td>
<td>0.00</td>
<td>0.49</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>R: TCTCTCTCTCACCACACACAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin07</td>
<td>F: AAATCGGGTCGAACATCAAGC (TC)$_6$(AC)$_7$</td>
<td>145-210</td>
<td>52</td>
<td>JF900338</td>
<td>5</td>
<td>0.50</td>
<td>0.70</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>R: TCTCTCTCTCACCACACAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin08</td>
<td>F: CACAAGGATCATTTGAGGAGG (AC)$<em>6$(AG)$</em>{12}$</td>
<td>178-193</td>
<td>50</td>
<td>JF900339</td>
<td>4</td>
<td>0.33</td>
<td>0.72</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>R: ACACACACACAGAGAGAGAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin09</td>
<td>F: TCGGGACCCCTAATGACATA (TC)$_6$(AC)$_6$</td>
<td>186-250</td>
<td>58</td>
<td>JF900340</td>
<td>5</td>
<td>0.75</td>
<td>0.83</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>R: TCTCTCTCTCACCACACAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin10</td>
<td>F: CGGGCTGGTATCTCAAGAGT (AC)$<em>6$(AG)$</em>{14}$</td>
<td>146-218</td>
<td>55</td>
<td>JF900341</td>
<td>3</td>
<td>0.25</td>
<td>0.61</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>R: ACACACACACAGAGAGAGAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin11</td>
<td>F: CGGTAGCAACATCGTCATA (AC)$<em>6$(AG)$</em>{13}$</td>
<td>144-180</td>
<td>55</td>
<td>JF900342</td>
<td>3</td>
<td>0.00</td>
<td>0.62</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>R: ACACACACACAGAGAGAGAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.09</td>
<td>0.18</td>
<td>0.61</td>
</tr>
</tbody>
</table>
apparently associated with the limited effective population size of the species. Small populations of narrowly distributed species are expected to exhibit low levels of genetic variation (Hamrick and Godt, 1989). In this study, 11 loci showed significant heterozygote deficiency. The observed heterozygosity values (average \(Ho=0.27\)) for \(P. massoniana\) are much lower than expected heterozygosity values (average \(He=0.65\)) (Table 1). The observed heterozygosity (\(Ho\)) values of 0.00-0.82 with a mean of 0.27 reported here for \(P. massoniana\) are lower than those for other pine species using microsatellite markers. For example, \(Ho\) values for \(P. radiata\) ranged from 0 to 0.850 with a mean of 0.625 (Smith and Devey, 1994), those for \(P. strobos\) ranged from 0.125 to 0.812 with a mean of 0.515 (Echt et al., 1996), and for \(P. pinaster\) ranged from 0.584-0.690 with a mean of 0.645 (Mariette et al., 2001). The \(Ho\) values for the three sister species were also higher than those in \(P. massoniana\) detected in this study (average \(Ho=0.27\)). For example, the average \(Ho\) value is 0.44 in \(P. thunbergii\), 0.41 in \(P. densiflora\), and 0.42 in \(P. luchuensis\) (Guan et al., 2011). The average allele number (4.64 alleles per locus) across the eleven polymorphic loci characterizing \(P. massoniana\) was lower than that in other other pine species, e.g., 6.00 alleles per locus observed in \(P. radiata\) (Smith and Devey, 1994), 5.40 in \(P. strobos\) (Echt et al., 1996), 6.70 in \(P. sylvestris\) (Soranzo et al., 1998), 7.00 in \(P. thunbergii\), 5.00 in \(P. densiflora\), and 5.13 in \(P. luchuensis\) (Guan et al., 2011).

Compared to Guan et al. (2011), in which \(A = 1-20\) alleles, and \(Ho=0.875\) were detected in the Hubei population of \(P. massoniana\), relatively lower allele number and observed heterozygosity in Mt. Huangshan population were detected in our study. The difference in the genetic diversity between two populations may be due to few sample size and sampling bias. Nevertheless, \(Ho\) and \(He\) values mostly approximated to each other in the Hubei population (Guan et al., 2011), while the observed heterozygosities were lower than the expected values in the Mt. Huangshan population (Table 1), reflecting that the latter population may have experienced demographic fluctuations, subsequently leading to the loss of genetic variation. For population genetic analysis across populations over the distribution range in the future, both primer sets will be used to test genetic diversity and population structure of mainland China and Taiwan for comparison, if the species/populations property attribute to similarly lower genetic diversity.

In this study, the low genetic variation in \(P. massoniana\) in Taiwan is likely attributed to the habitat destruction, which inevitably reduced the effective population size. These HWE deviations detected in all microsatellite loci are also likely due to the effects of bottlenecks that were caused by habitat destruction, fatal disease, or the existence of null alleles. At the population level, the genetic variation in the Taiwanese population (average \(Ho = 0.18\)) was much lower than that in mainland China population (average \(Ho = 0.27\)) (Table 1). The possible explanations include serious habitat destruction, pine wilt disease, and founder’s effects in Taiwanese population of \(P. massoniana\) associated with colonization. In Taiwan, some species/populations have close phylogenetic links to their mainland relatives, e.g., \(P. luchuensis\) ssp. hwangshanensis and \(P. luchuensis\) ssp. taiwanensis (Chiang et al., 2006), \(Cycas taitungensis\) and \(C. revoluta\) (Huang et al., 2001; Chiang et al., 2009). Such a phylogeographical pattern indicates colonization from the Asian continent eastward to Taiwan. One would therefore expect that island populations maintain lower genetic variability as a result of smaller population sizes and numbers due to habitat limitation on islands as well as genetic bottlenecks associated with colonization (Chiang and Schaal, 2006).

For \(P. massoniana\) that experienced a dramatic demographic decline, it is urgent to develop molecular markers to assess the genetic diversity and obtain the necessary information for developing conservation and management strategies of \(P. massoniana\). In the future, more samples in Taiwan and mainland China will be collected, and the application of these microsatellite markers including Guan et al.’s study may provide a tool for understanding demography of \(P. massoniana\) in mainland China and Taiwan and assessing founders’ effects in the latter.

**LITERATURE CITED**


馬尾松 (松科) 微衛星基因座的分離及分析

洪國翔¹ 林其永² 黃啟俊² 黃志川³ 許再文⁴
郭耀綉⁵ 王唯匡² 洪承裕² 蔣鎮宇²

¹ 國立屏東科技大學 生物資源研究所
² 國立成功大學 生命科學系
³ 國立成功大學 工程科學系
⁴ 農委會 特有生物研究保育中心
⁵ 國立屏東科技大學 森林系

馬尾松廣泛分布於中國大陸中部以及東部地區，但在台灣因棲地破壞以及松樹萎凋病的影響而僅存單一現生族群。本研究意欲分離馬尾松微衛星基因座以利未來能應用於探討馬尾松族群遺傳結構，結果顯示共分離出 11 組可用的微衛星基因座，進一步分析顯示在馬尾松中此 11 組微衛星基因座其對偶基因數目為 2 到 9 個；異質度觀測值 (Ho) 介於 0.00 到 0.82；然而異質度期望值 (He) 則為 0.40 到 0.83，此些馬尾松微衛星基因座將可提供一工具來探討台灣以及中國大陸馬尾松族群的遺傳結構。

關鍵詞：異型核子；微衛星；族群結構；松科；馬尾松。