Theor Appl Genet (2002) 104:819–827
DOI 10.1007/s00122-001-0794-z

M. Shepherd · M. Cross · T. L. Maguire
M.J. Dieters · C.G. Williams · R.J. Henry

Transpecific microsatellites for hard pines

Received: 8 January 2001 / Accepted: 18 May 2001 / Published online: 6 February 2002

Abstract Microsatellites are difficult to recover from large plant genomes so cross-specific utilisation is an important source of markers. Fifty microsatellites were tested for cross-specific amplification and polymorphism to two New World hard pine species, slash pine (*Pinus elliottii* var. elliottii) and Caribbean pine (*P. caribaea* var. hondurensis). Twenty-nine (58%) markers amplified in both hard pine species, and 23 of these 29 were polymorphic. Soft pine (subgenus *Strobus*) microsatellite markers did amplify, but none were polymorphic. *Pinus elliottii* var. elliottii and *P. caribaea* var. hondurensis showed mutational changes in the flanking regions and the repeat motif that were informative for *Pinus* spp. phylogenetic relationships. Most allele length variation could be attributed to variability in repeat unit number. There was no evidence for ascertainment bias.

Keywords Conifers · Gymnosperms · Trans-specific polymorphism

Communicated by P. Langridge

M. Shepherd (✉) · M. Cross · R.J. Henry
Cooperative Research Centre for Sustainable Production Forestry,
Centre for Plant Conservation Genetics,
Southern Cross University, P.O. Box 157 Lismore,
NSW 2480, Australia
e-mail: mshepher@scu.edu.au
Fax: +61-2-66222080

T.L. Maguire
Department of Botany, The University of Queensland,
Brisbane, Queensland 4072, Australia

M.J. Dieters
Queensland Forestry Research Institute, M.S. 483 Fraser Rd,
Gympie, Queensland 4570, Australia

C.G. Williams
Graduate Genetics Program, TAMU 2135,
Texas A & M University, College Station, Texas,
77843-2125, USA

Introduction

Transferring microsatellites from related conifer species is appealing because it circumvents de novo microsatellite development. Developing pine microsatellites has proven difficult because the size (approx. 28,000 pg/C) and complexity (approx. 75–86% highly repetitive DNA) of the pine genome represent significant barriers (Echt et al. 1999; Smith and Devey 1994; Soranzo et al. 1998). Secondary screening of enriched libraries (Pfeiffer et al. 1997; Scott et al. 1999), eliminating repetitive regions of the genome (Smith and Devey 1994) or enriching for low-copy genomic sequences (Elsik et al. 2000) have also increased the efficacy of pine microsatellite development. The paucity of pine genomic sequences in the public sequence databases limits the value of this approach to finding microsatellites. Similarly, a survey of several thousand clones from a pine expressed sequence tag (EST) library yielded only eight polymorphic microsatellite markers (Echt and Burns 1999).

Microsatellite transfer in pines is uncertain for several reasons. First, microsatellite transfer in flowering plants tends to be restricted to closely-related genera (Peakall et al. 1998). Similarly, transfer to other subgenera or non-Pinus coniferae has been reported to be poor (Echt et al. 1996; Fisher et al. 1998), although optimisation of polymerase chain reaction (PCR) conditions increased transfer rates for some of these loci (Karhu et al. 2000). Kutil and Williams (2001) reported that tri-nucleotide microsatellites from low-copy genomic regions increased transfer rates within Pinaceae. Unlike previous investigators, they noted that both flanking regions and repeat motifs were highly conserved in soft and hard pines. Consequently, the transfer of pine microsatellites should be feasible in some cases.

Identifying polymorphic microsatellites for *Pinus caribaea* Morelet and *P. elliottii* Little and Dorman (section Pinus; subsection Australes) has application worldwide. These species are indigenous to southern USA, Central America or the Caribbean and are cultivated throughout the world for timber and fibre (Lamb 1973;

Table 1 Results of transfer and PCR conditions for 50 *Pinus* spp. microsatellite loci

Locus	Source taxon[a]	Repeat motif	Expected size[b]	$MgCl_2$ [c] (m*M*)	Buffer[d]	Transfer class[e]	Reference[f]
PtTX 2008	PT	$(GAT)_5$	307	4.0	1	1	3
PtTX 2037	PT	$(GTGA)_8$ $(GT)_{14}$	177	4.0	1	1*	3
PtTX 2034	PT	$(TTTG)_9$	217	5.0	1	1*	1
PtTX 2123	PT	$(AGC)_8$	202	4.0	1	1*	3
PtTX 2142	PT	$(CTG)_5$	262	2.0	1	1*	3
PtTX 2146	PT	$(GCT)_{21}$	180	5.0	1	1*	3
PtTX 3011	PT	$(GAA)_8$ $(GAT)_{18}$ $(GAG)_3$	186	4.0	1	1*	3
PtTX 3013	PT	$(GTT)_{10}$	134	4.0	2	1*	3
PtTX 3018	PT	$(GAT)_{13}$	155	4.0	1	1*	1
PtTX 3019	PT	$(GAA)_{11}$	223	4.0	1	1*	3
PtTX 3020	PT	$(CAA)_9$	211	4.0	1	1*	3
PtTX 3025	PT	$(CAA)_{10}$	266	4.0	1	1*	3
PtTX 3029	PT	$(GCT)_5$ …$(GCT)_8$ …$(GCT)_5$	255	4.0	1	1*	1
RPTest 01	PT	$(ATA)_7$	125	3.5	1	1	4
RPTest 05	PT	$(CAA)_6$ $(CAA)_5$ $(AAC)_4$	197	3.5	1	1	4
RPTest 08	PT	$(GCG)_6$	196	3.5	1	1	4
RPTest 09	PT	$(A)_5$ $(TG)_4$ $(GAG)_5$ $(CAG)_{11}$ $(GCA)_7$ $(GCA)_5$	289	3.5	1	1	4
RPTest 11	PT	$(CAT)_7$	213	3.5	1	1*	4
RPTest 20	PT	$(CAG)_5$	259	3.5	1	1*	4
NZPR1	PR	$(AG)_{17}$	139	1.5	2*	1*	2
NZPR2	PR	$(AG)_{23}$	162	1.5	2*	1	2
NZPR5	PR	$(AG)_{29}$	114	1.5	2	1*	2
NZPR7	PR	$(AT)_{22}$ $(AG)_{27}$	149	1.5	2*	1*	2
PtTX 2128	PT	$(GAC)_8$	245	5.0	1	2*	3
PtTX 3001	PT	$(CAA)_3$ …$(CAA)_3$ $CAG(CAA)_4$	313	4.0	1	2	1
PtTX 3002	PT	$(CAA)_6$ …$(GAG)_4$ $AA(GAG)_4$	194	4.0	1	2	8
RPS105	PS	AC	151	1.5	2	2*	5
RPS150	PS	$(GAG)_4$	248	1.5	2	2*	6
RPS61	PS	AC	195	1.5	2	2*	5
PtTX 2164	PT	$(TCG)_{19}$ $(TCA)_{16}$	252	5.0	1	3	3
PtTX 3034	PT	$(GT)_{10}$ $(GA)_{13}$	207	4.0	1	3	3
PtTX 3021	PT	$(GTTTTT)_4$.$(GTT)_5$.$(GTT)_5$.$(GTT)_4$	471	4.0	1	3	1
PtTX 3023	PT	$(CAA)_4$.$(CAA)_4$	168	4.0	1	3	1
PtTX 3027	PT	$(CAT)_{10}$	280	4.0	1	3	1
PtTX 3035	PT	$(AGG)_7$.$(AGG)_2$ $CGG(AGG)_3$.$(AGG)_7$.$(AGG)_3$	325	4.0	1	3	1
PtTX 3037	PT	$(GA)_9$ $(CAA)_{15}$	144	4.0	1	3	3
RPTest 13	PT	$(CTG)_5$	277	3.5	1	3	4
RPTest 15	PT	$(GTG)_6$ $(TGG)_4$ $(TGG)_4$ $(T)_5$	246	3.5	1	3	4
NZPR3	PR	$(AG)_{14}$ $(AGGG)_3$	148	1.5	2	3	2
NZPR6	PR	$(AG)_{25}$	198	1.5	2	3	2
PR-4.6	PR	$(CA)_{21}$ $(TA)_6$ N_{11} $(TAA)_9$	206–222	1.5	2	3	7
RPS3	PS	$(AC)_{19}$	287	1.5	2	3	5
RPS160	PS	$(ACAG)_3$ $AGGC(AGAC)_3$	246	1.5	2	3	6
PtTX 3014	PT	$(GTT)_{11}$	148	4.0	1	4	1
PtTX 3017	PT	$(GAT)_9$	212	4.0	1	4	3
PtTX 3026	PT	$(ACC)_8$ $(ATC)_{22}$	344	4.0	1	4	8
PtTX 3030	PT	$(TA)_8$ $(GGT)_{10}$	287	4.0	1	4	3
PtTX 3032	PT	$(GAT)_{42}$ $(GAC)_{17}$	335	4.0	1	4	3
NZPR4	PR	$(AG)_{20}$	146	1.5	2	4	2
PR-9.3	PR	$(CA)_{14}$	92–108	1.5	2	4	7

[a] Taxa are: *PT P. taeda ·PR P. radiata ·PS P. strobus*

[b] Published expected PCR product size based on the sequence from the clone from the individual used to develop the library

[c] Final $MgCl_2$ concentration in PCR buffer

[d] See Methds and materials for buffer compositions. An asterisk (*) indicates AmpliTaq Gold DNA Polymerase (Perkin Elmer) was used in the PCR

[e] See Methods and materials. An asterisk (*) indicates primer-pair was tested on a population of ten megagametophytes from one individual each of *P. elliottii* var. elliotti and *P. caribaea* var. hondurensis

[f] References: *1* This paper, *2* Fisher et al. 1998, *3* Elsik et al. 2000, *4* Echt and Burns 1999, *5* Echt and May-Marquardt (unpublished) and Echt et al. 1999, *6* Echt et al. 1996, *7* Smith and Devey 1994, *8* Kutil and Willims (in press)

Nelson et al. 1993). The option of creating de novo microsatellites for every pine species is daunting as pines constitute the largest genus in the Pinaceae family, with 100 species in the northern hemisphere (Little and Critchfield 1969). Transfer may be inefficient because the Pinaceae family is ancient, extending back to the late Triassic (more than 190 million years ago) (Miller 1977). Similarly, the phylogenetic distance between hard and soft pines is great; the fossil record indicates that the subgenera *Pinus* (hard pines) and *Strobus* (soft pines)

were formed by the late Cretaceous (approx. 130 million years ago) (Miller 1977).

The study reported here tested microsatellite transfer to *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis from two other hard pines and a soft pine. Criteria for successful transfer included clear amplification product and polymorphism. Focal species were defined as the source of de novo microsatellite loci and nonfocal taxa referred to the target taxa for transfer, *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis. A nonfocal population was defined as the *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis combined.

Methods and materials

Taxonomic classification and sources of DNA samples

The study encompassed material from four hard pine (subgenus *Pinus*) taxa, *P. radiata*, *P. taeda*, *P. caribaea* var. hondurensis and *P. elliottii* var. elliottii, and a soft pine (subgenus *Strobus*), *P. strobus*. *Pinus radiata* belongs to the subsection Oocarpae (section Pinus), whereas the other three hard pines belong to the subsection Australes (section Pinus) (Little and Critchfield 1969). *Pinus strobus* is a member of the subsection Strobi (section Strobus).

Four *P. elliottii* var. elliottii and five *P. caribaea* var. hondurensis individuals were randomly selected from the Queensland Forest Research Institute breeding populations as the nonfocal samples. Foliage material or DNA from a single individual from each of the focal species was obtained as controls. DNA was extracted from foliage tissue according to Graham et al. (1996). Additionally, ten seeds from an open-pollination of a *P. elliottii* var. elliottii tree (2PEE1-102) and a *P. caribaea* var. hondurensis tree (1PCH1-063) provided megagametophyte tissue. Megagametophyte tissue was freeze-dried and ground prior to extraction using a DNeasy Kit (QIAGEN Valencia, Calif).

Microsatellite transfer and optimisation

Fifty microsatellite markers (Table 1) were examined for transfer: 28 markers from the *Pinus taeda* (PtTX) series were developed from total genomic or low-copy libraries (Elsik et al. 2000; Kutil and Williams 2001; Table 2); eight *P. taeda* (RPTest) markers were derived from ESTs (Echt and Burns 1999); two *P. radiata* (PR) markers were derived from a total genomic library of *P. radiata* (Smith and Devey 1994); seven *P. radiata* (NZPR) markers were derived from total genomic libraries (Fisher et al. 1998); five *P. strobus* (RPS) markers were derived from total genomic libraries (Echt et al. 1996).

Microsatellite transfer was a stepwise process. First, microsatellite loci were evaluated for amplification in the nonfocal population and the focal taxa control using published conditions. PCR amplification was assessed on 3.5% agarose gels. If positive, then amplification products were tested for polymorphism using an ABI 310 Genetic Analyser (Perkin Elmer, Foster City, Calif.) and detected by fluorescence from dye-labelled dUTPs (R110, Perkin Elmer) or from a labelled primer (FAM, HEX or TET). In cases where primer-pairs gave low yield or complex banding patterns, PCR optimisation was attempted by testing a range of $MgCl_2$ concentrations on the control focal species sample.

Hot start PCR, using an antibody-inactivated *Taq* (AmpliTaq Gold, Perkin Elmer), was tested on nine recalcitrant microsatellites (Table 1). Amplifications were carried out using GeneAmp 9700 or a 9600 thermocycler (Perkin Elmer), using 25-u1 or 12.5-u1 reactions with 20 ng of DNA template per reaction and final concentrations of 100 u*M* each dNTP and 200 n*M* each primer. One of two PCR buffers was used. Buffer 1 consisted of 10 × PCR reaction buffer, 100 m*M* Tris-HC1, l5 m*M* $MgCl_2$, 500 m*M* KC1, pH 8.3 (Roche, Indianapolis, Ind buffer 2 of 10×stock – 500 m*M* Tris-HCl, 200 m*M* ammonium sulfate, 120% sucrose, l mg /ml gelatine. The final concentration of magnesium chloride varied depending on the primer-pair (Table 2).

Table 2 Primer-pairs for microsatellite loci obtained from *P. taedu* low-copy libraries

Marker	Primer sequences (5′ -3′)
PtTX2034[a]	TCTGAGGAGGAACATGTCATTTACT - F
	GCATGTCTGAATTATTGTGTTCTAT - R
PtTX3001	ATAAAGGCAGAGGATGAACA - F
	CCCAATTGTTATTTCTGATT - R
PtTX3002	TTGTTGTGCTCATAATTACTAGTGT - F
	CTCCTAAGCTTGCTCATGTG - R
PtTX3014	CCATTGACGCTCCCGTTACGTTACT - F
	CGAGAGACGTGCGGATACAAGACCT - R
PtTX3018	CCATTTATGAACCAGAGA - F
	ATTAAAACCATGAGACCTT - R
PtTX3021	TTCATCCTAGCTGCTTGCTTT - F
	CTCAGCGTCTACCCCATCAA - R
PtTX3023	CATCTAGTTACCAAAGTTAT - F
	ATTTATGAAAATGGTAAGT - R
PtTX3027	TCCATTTGAGAACTTTTT - F
	AGGAGCCACAACATAATA - R
PtTX3029	CTTGTTGCTGCTTCTGC - F
	AACAAAATAATATAAATGCTCTGC - R
PtTX3035	AGGAGGAGGAGTTGGAGTT - F
	ATCGCCCTAGCTGGTTTAT - R

[a] See Elsik et al. (2000) for methods

Complex banding patterns for a microsatellite locus were viewed as either multiple alleles or multiple loci. These markers were tested on ten haploid megagametophytes from a *P. elliottii* var. elliottii individual (2PEE1-102) and a *P. caribaea* var. hondurensis individual (1PCH1-063) and a population of 93 controlled-pollinated hybrid offspring. Allele segregation patterns also distinguished homozygotes from heterozygotes with null alleles.

Successful microsatellite transfer to *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis was based on a clear amplification product of the expected size and polymorphism. The four classes for microsatellite marker transfer criteria were as fo1lows: (1) polymorphic, (2) monomorphic, (3) poor amplification resulting in low product yield or non-specific amplification or (4) no amplification.

Microsatellite sequencing

Six *P. taeda* (PtTX) microsatellite loci were sequenced in *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis to compare repeat structure and sequence using haploid megagametophyte DNA. The sequencing of a PCR-amplified product was carried out using Big Dye Sequencing Kit (Perkin Elmer). Gel separation of the sequencing reactions was carried out at the Australian Genome Research Facility, Brisbane. Sequences were aligned using CLUSTAL-W multiple (accurate) (Thompson et al. 1994) using a gap opening penalty of 10 and manual alignment adjustments.

Statistical analysis

Heterozygosity (H) was calculated according to Crow (1986) using the following formula.

$$\mathrm{H} = 1 - \sum_{i=1}^{\mathrm{n}} f_i^2 \qquad (1)$$

where f_i is the frequency of the *i*th allele and n is the allele number.

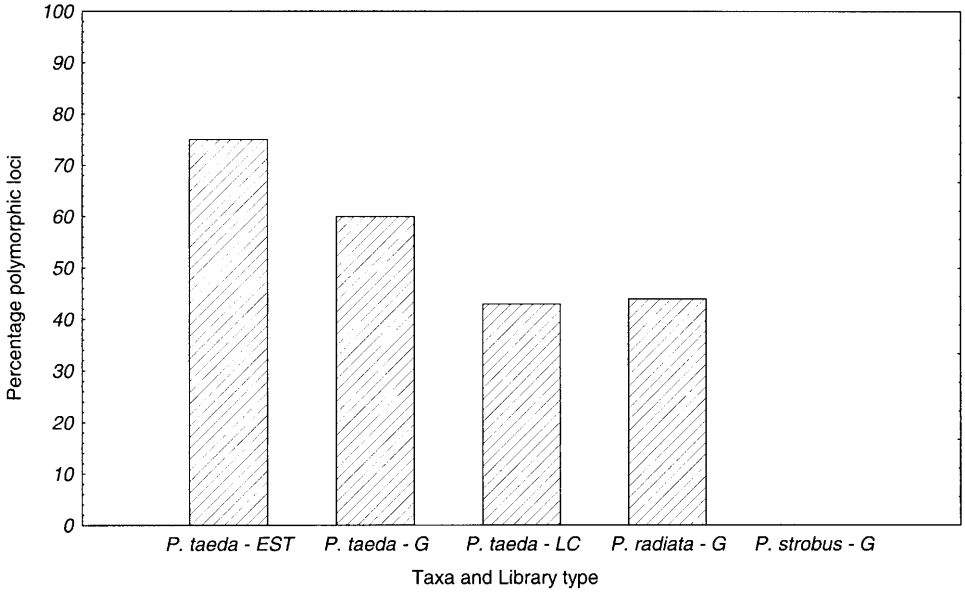

Fig. 1 Transfer of polymorphic *Pinus* spp. microsatellite loci categorised by source. Microsatellite sources are identified by the focal taxon and library type. Library types: *G* genomic, *LC* low-copy, *EST* expressed sequence tag

Table 3 Transfer of *Pinus* spp. microsatellite loci to nonfocal population species as categorised by the focal taxa and library type

Locus source (taxon and library type)	Number of loci tested	Number of loci amplified	Number of loci polymorphic
P. strobus – G[a]	5	3	0
P. radiata – G	9	4	4
P. taeda – G	5	4	3
P. taeda – LC	23	12	10
P. taeda – EST	8	6	6
Totals	50	29	23

[a] Library types: *G* Total genomic library including enriched · *LC* low-copy library · *EST* EST library
[b] Number of amplified loci refers to number of loci which were polymorphic or monomorphic

Linear regression and *t*-tests were calculated using a spreadsheet. One-way analysis of variance (ANOVA), least significant difference test (LSD) and Fisher's Exact tests were calculated using Statistica v 4 (Statsoft, Tulsa, Okla.)

Results

Microsatellite marker transfer to *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis was successful for 23 of the 50 loci (Table 1). Twenty-nine microsatellite loci amplified in the two nonfocal taxa, and 23 of these 29 were polymorphic (Tables 2 and 3). If a microsatellite was polymorphic within a hard pine focal species, then it was also polymorphic within the nonfocal taxa. The two exceptions were PtTX 3011 for which *P. caribaea* var. hondurensis was monomorphic and NZPR2 where *P. elliottii* var. elliottii was monomorphic. Six of the 50 microsatellite loci were monomorphic in both focal and nonfocal species, and 21 microsatellite loci had either poor or no amplification.

Five sources of microsatellites markers were recognised based on the focal taxa and the library type from which they were derived. The three library types were EST, total genomic or low-copy. The *P. strubus* and *P radiata* markers were coded as PS – G and PR – G, respectively. The *P. taeda* markers were coded by library type: EST (PT – EST), total genomic (PT – G) or low-copy (PT – LC). With respect to transfer by amplification, the PT – G and PT – EST classes had high transfer rates, 80% and 75%, respectively. The PS – G and PT – LC classes demonstrated moderate rates of transfer (60% and 52%, respectively), whereas PR – G had the lowest transfer (44%). When transfer was assessed, by polymorphism, PT – EST had the highest proportion (75%) of polymorphic markers, followed by PT – G (60%), PR – G (44%) and PT – LC (43%) were lower and similar (Fig. 1). A Fisher's exact test was used to compare the transfer rate of polymorphic markers at a taxon level. As many markers from *P. radiata* transferred and were polymorphic as *P. taeda*, despite the greater phylogenetic distance for *P. radiata* (Fisher's Exact test; two-tailed *P* value = 0.72).

Heterozgyosity was estimated for 19 of the 23 markers where at least four individuals amplified in each taxon (Table 4). Heterozygosity values ranged from zero to 0.86 for individual loci within the two taxa with an average of 0.5 ± 1.9 for *P. elliottii* var. elliottii and 0.49 ± 0.23 for *P. caribaea* var. hondurensis. The mean H value and the mean allele number for all loci were the same in both *P. elliottii* var. elliottii and *P. caribaea* var. hondu-

Table 4 Heterozygosity (H) and number of alleles for 19 *Pinus* spp. microsatellite markers based on a sample of four *P. elliottii* var. elliottii and five *P. caribaea* var. hondurensis individuals. Alleles were sized by capillary electrophoresis (see Methods and materials). Where a primer-pair was multicopy, only alleles at one locus were examined

Locus[a]	Focal sp.	*P. caribaea* var. hondurensis			*P. elliottii* var. elliottii			Nonfocal population	
	Allele sizes (bp)	Number of alleles	Allele size range (bp)	H	Number of alleles	Allele size range (bp)	H	Total no. non-focal alleles	Percentage of alleles unique to each taxa
PtTX 3013	129,132	4	128–141	0.66	3	138–145	0.59	6	67
PtTX 2008	305,305	2	305,316	0.38	2	305,316	0.38	2	0
PtTX 2123	204,204	2	198,201	0.50	2	198,201	0.47	2	0
PtTX 3034a	199,211	2	205,213	0.38	7	191–209	0.84	8	88
PtTX 3025	256,266	3	254–268	0.41	2	252,268	0.22	4	75
PtTX 3019	219,223	4	210–221	0.72	2	205,208	0.50	6	100
PtTX 3020	171,183	3	147–163	0.53	4	151–159	0.75	7	100
PtTX 2128	244,270	2	231,238	0.50	2	231,238	0.50	2	0
RPTest 01	124,124	3	112–130	0.53	2	112,124	0.38	3	33
RPTest 05	199,199	1	195	0.00	2	192,195	0.22	2	50
RPTest 08	194,194	2	192,195	0.22	2	192,195	0.38	2	0
RPTest 09a	264,279	4	269–282	0.69	2	252,258	0.22	6	33
RPTest 11	214,214	2	209,214	0.22	3	209–217	0.41	3	33
RPTest 13	267,276	3	267,276	0.53	2	267,276	0.50	3	33
RPTest 20	254,254	1	254	0.00	2	232,254	0.47	2	50
NZPR1	133,133	8	121–145	0.86	5	121–149	0.76	11	82
NZPR5a	112,114	3	80–84	0.64	3	78–82	0.41	4	50
NZPR6	184,184	5	172–190	0.74	5	184–196	0.78	8	75
NZPR7	113,139	5	77–101	0.80	5	93–113	0.83	9	89
Mean±SD		3.00±1.57		0.49±0.23	2.86±1.42		0.50±0.19	4.5±2.7	49±33

[a] Multicopy loci are indicated by an "a" suffix on the locus label. Null alleles are known in NZPR7 but this was scored as a homozygote to be consistent with other loci where it was unknown whether there were null alleles

rensis (two-tailed *t*-tests, *P* values > 0.5). The number of alleles for each locus ranged from one to eight in *P. caribaea* var. hondurensis and from two to seven in *P. elliottii* var. elliottii (Table 4). Across both taxa, the number of alleles ranged from two to eleven per locus, with an average of 4.5 ± 2.7 and with 49% ± 33% of alleles unique to each taxa.

The H values for eight PtTX markers tested in this study were compared to H values reported for the same markers in *P. taeda* (Elsik et al. 2000). The loci tested were PtTX 3025, PtTX2008, PtTX2123, PtTX2128, PtTX3019, PtTX3034, PtTX3013 and PtTX3020. There was no difference in the average H for the markers when they were transferred to the nonfocal population of *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis (mean ± SD of H focal = 0.59±0.24, nonfocal = 0.55±0.10; two-tailed paired *t*-test, *t* value of 0.68, *P* = 0.51).

To test whether different sources of markers yielded different levels of variability in the nonfocal population, we compared the H values between sources of polymorphic transpecific markers. The PS – G source was excluded because all of these soft pine microsatellites were monomorphic. A one-way ANOVA indicated a significant difference in average H between the sources (one-way ANOVA *F*-value = 6.9, *P* value = 0.004; Table 5). The PR – G source had the highest average H (0.73±0.14), although it was not significantly higher than the PT – LC or PT – G sources (0.53±0.14 and 0.49±0.01, respectively). The PT – EST source had the lowest H (0.34±0.14) and was significantly lower than all other sources except PT – G.

Table 5 A comparison of microsatellite variability (H) in the nonfocal population by focal taxa and library type[a]. Least significant difference (LSD) test was applied at $P<0.05$ following a one-way ANOVA test; *P* value = 0.004

Locus source	Number of loci (taxa and library type)	H (mean ± SD)	LSD[b]
P. radiata – G	4	0.73±0.14	a
P. taeda – LC	6	0.53±0.14	a,b
P. taeda – G	2	0.49±0.01	a,b,c
P. taeda – EST	7	0.34±0.14	c

[a] See Table 3

[b] Different letters indicate that H was significantly different (See Results)

The variability at microsatellite loci in different repeat type classes was compared using an average of the heterozygosity for each of the two nonfocal taxa. Di-nucleotide repeat microsatellite loci were significantly more variable than the tri-nucleotide repeat loci derived from either genomic or EST libraries (Table 6).

Markers NZPR7 and PtTX2037 had null alleles. No other null alleles were confirmed in a sample of 21 polymorphic or monomorphic markers (see Table 1). As previously reported for hard pines, microsatellite loci RPTest 09 and NZPR5 were multilocus in *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis (Echt and Burns 1999;

Table 6 A comparison of microsatellite variability (H) for microsatellite repeat type and library type in the nonfocal population of *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis. Heterozygosity is an average of values for *P. elliottii* var. elliotttt and *P. caribaea* var. hondurensis. See Table 3 for library type categories. Only loci which were perfect or compound as defined by Weber (1990) and modified by Elsik (2000) were included. Loci which had a combination of di- and trimers were excluded from the analysis. Least significance difference test applied following one-way ANOVA on the three group: $F = 9.5$, P-value=0.001

Microsatellite repeat type[a] and source class	*n*	H (mean ± SD)	LSD[b]
Di – G	8	0.64±0.13	a
Tri – G	6	0.49±0.13	b
Tri – EST	6	0.32±0.15	b

[a] *Di* Di-nucleotide repeat motif · *Tri* tri-nucleotide repeat motif
[b] Different letters indicate a significant difference in H (see Results)

Fisher et al. 1998). Microsatellites NZPR3, PtTX3029 and PtTX2034 also appeared to be multicopy in *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis.

Transpecific polymorphism: a sequence-based comparison

One or two alleles from each of six PtTX loci that amplified from megagametophytes from both a *P. elliottii* var. elliottii and a *P. caribaea* var. hondurensis individual were sequenced and aligned with the corresponding loci from *P. taeda*. The repeat unit number accounted for most of the differences between the focal and the nonfocal species (Table 7). The repeat motif was conserved in all transfer microsatellites except for locus PtTX 2037. At this locus, a dinucleotide motif (GA) was present in both nonfocal species that was absent in *P. taeda*. The focal and nonfocal species were found to also differ by a total of three single-base pair insertions and 13 single-base pair substitutions within the flanking regions of the six loci. Linear regression was carried out on the PCR product length and repeat length for four loci where there was sufficient data and variability. The relationship was highly significant for PtTX 3013 ($R^2 = 0.99$, P value << 0.001) and PtTX 3034 ($R^2 = 1$, P value = 0), significant for PtTX 2037 ($R^2 = 0.99$, P value = 0.03) but not highly significant for PtTX 2123 ($R^2 = 0.89$, P value = 0.06).

Ascertainment bias was not evident between *P. taeda* and the two nonfocal species. Within the nonfocal group, only two single base pair substitutions were found between loci in *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis. The sequence length of alleles in the nonfocal group were generally shorter for 12 of the 15 alleles, but there was not a significant reduction in sequence or repeat length between the focal and nonfocal population (two-tailed paired t-tests; $t = 1.49$, $P = 0.19$ and $t = 1.628$ $P = 0.17$ for sequence and repeat length, respectively).

Table 7 Repeat region structural and flanking sequence mutations at six microsatellite loci from an individual from each of the taxons *P. elliottii* var. elliottii, *P. caribaea* var. hondurensis, and a *P.taeda*. Mutations are relative to sequence position in Genbank accessions for *P. taeda.* Substitutions noted as, for example, G-C represents a change from a G in *P. taeda* to a C in *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis. Insertions noted as, for example, -T is an insertion of a T into *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis. AC to T mutation at locus PtTX3002 only occurs in *P. caribaea* var. hondurensis

Locus	Taxon[a]	GenBank acc. no.	Allele	Repeat motif	Repeat length (bp)	Sequence length (bp)	Mutuations in flanking sequence
PtTX2037	PT	AF143959	1	$(GTGA)_8\,(GT)_{14}$	60	176	A-C (34); 2 ×-T (139, 154)
	PE		1	$(GTGA)_1\,(GA)_2\,(GT)_{12}$	32	150	
	PC		1	$(GTGA)_1\,(GA)_4\,(GT)_{17}$	46	164	
PtTX2123	PT	AF143960	1	$(AGC)_8$	24	202	-G (301); 3 × T-A (318, 325, 327); 3 × A-G (321, 323, 326)
	PE		1	$(AGC)_8$	24	203	
	PC		1	$(AGC)_8$	24	203	
	PC		2	$(AGC)_7$	21	200	
PtTX3002	PT	AF277846	1	$(AGG)_6$	18	194	C-T (75); T-C (93)
	PE		1	$(AGG)_6$	18	194	
	PC		1	$(AGG)_6$	18	194	
PtTX3013	PT	AF143966	1	$(GTT)_{10}$	30	134	G-T (95); -T (175)
	PE		1	$(GTT)_9$	27	132	
	PE		2	$(GTT)_{12}$	36	141	
	PC		1	$(GTT)_9$	27	132	
	PC		2	$(GTT)_{12}$	36	141	
PtTX3020	PT	AF143969	1	$(CAA)_8$	24	211	G-A (152)
	PC		1	$(CAA)_1$	3	190	
PtTX3034	PT	AF143974	1	$(GT)_{10}\,(GA)_{13}$	46	207	G-C (101); T-A (172)
	PE		1	$(GT)_9\,(GA)_{10}$	38	199	
	PC		1	$(GT)_{14}\,(GA)_{12}$	52	213	
	PC		2	$(GT)_9\,(GA)_{13}$	44	205	

[a] *PT Pinus taeda* · *PE P. elliottii* var. elliottii · *PC P. caribaea* var. hondurensis

Discussion

Microsatellite loci transferred to *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis equally from *P. radiata* and *P. taeda*. The transfer rates to both of the nonfocal taxa was similar, and no ascertainment bias was detected. The length of the repeat sequence in microsatellites transferred to the nonfocal population closely corresponded with PCR product size. Changes in the repeat structure or flanking sequence were minor and consistent with current taxonomic and phylogenetic relationships. Homologous microsatellite loci were as polymorphic in the two nonfocal species as in the focal species.

Efficient microsatellite transfer

The success of the transfer of polymorphic markers from *P. radiata* (subsection Oocarpae) was the same as for markers from *P. taeda* (subsection Australes). Also, despite a higher proportion (36/50) of loci from a focal species within the same subsection than nonfocal species, the overall transfer of polymorphic loci was similar (23/50) to the rate across a subsection. The rate of transfer for polymorphic loci across a subsection (4/9) was similar to that found previously for transfer of loci from *P. radiata* to *P. taeda* (9/20) (Devey et al. 1999). These subsections are believed to have evolved separately for at least 10 million years but perhaps as long as 35 million years, (Miller 1993; Krupkin et al. 1996). Microsatellite markers originating from focal species external to a nonfocal species subsection may be as useful as markers transferred from within the same subsection in hard pines, at least where they they have only diverged within the past 10–35 million years.

Microsatellite loci from *P. strobus* (section Strobus), while transpecific, were not polymorphic between the taxa. This was consistent with previously reported low levels of variation for these markers amongst species of hard pines (Echt et al. 1999). The hard and soft pine sections have been separated for around 130 million years (Mirov 1967). Nonetheless, the transfer of microsatellite markers over these evolutionary – and even greater (140–195 million years) – distances in the pines and of those which exhibit variation amongst hard pine taxa has been possible (Karhu et al. 2000; Kutil and Williams 2001). Conserved microsatellite loci in pines were those with perfect triplet-repeat motifs (Kutil and Williams 2001).

The type of library from which a microsatellite originated affected its H in the nonfocal group. EST-derived microsatellites were less polymorphic than loci identified from other library types in *P. taeda*. In rice, microsatellites in EST were found to be less polymorphic than those derived from genomic libraries (Cho et al. 2000). Our data suggests this may also be the case in *Pinus*. However, caution is required as the effect may be confounded with differences due to repeat types, as all EST-derived loci we tested were perfect trinucleotide or compound trinucleotide repeat loci. Loci with trinucleotide repeats as a group were less variable than dinucleotide repeat sequences in our study, and a similar effect has also been found in both *Drosophila* and humans (Chakraborty et al. 1997; Schug et al. 1998). A comparison of EST-derived microsatellite loci with microsatellite loci derived from genomic libraries unconfounded by differences in repeat types will be required to confirm whether EST-derived loci have lower polymorphism.

Although EST-derived microsatellite markers with trinucleotide repeats had a lower average H than dinucleotide repeats, they had similar H to other microsatellite loci containing trinucleotide repeats. EST-derived microsatellite markers also had a higher transfer rate than the average for all markers. This high transfer rate of EST-derived microsatellites compared with markers from other sources probably resulted from a higher sequence conservation in primer binding sites. Furthermore, EST-derived markers were proportionately well represented amongst loci that were mapped in a segregating population (data not shown), indicating that despite a lower polymorphism EST are a valuable source of markers.

Repeat structure and flanking mutations are phylogenetically informative

Differences in repeat structure and flanking sequence mutations were found to align with the taxonomic assignment of *P. elliottii* var. elliottii, *P. caribaea* var. hondurensis and *P. taeda* and were in agreement with phylogenetic relationships based on other molecular studies (Dvorak et al. 2000; Little and Critchfield 1969; Nelson et al. 1994), Dvorak et al. (2000) proposed evolutionary scenarios for the Australes and Oocarpae subsections based on a random amplified polymorphic DNA (RAPD) phylogeny, other marker data and information from crossing experiments and observations of natural hybridisation and distributions. They proposed a common ancestral origin for the Mesoamerican Oocarpae and Australes subsections. Because *P. caribaea* var. hondurensis has a genetic constitution of both Oocarpae and Australes, it shared a common ancestor with Mesoamerican Oocarpae and Australes but diverged before this split. A divergence between an ancestral. *P. caribaea* var. hondurensis and the ancestors that formed two clades within Australes – one consisting of the remaining *P. caribaea* varieties, caribaea and bahamensis, and the two *P. elliottii* varieties, *elliottii* and densa, and the second clade containing *P. taeda*, *P. palustris* and *P. echinata* – was not resolved. On the basis of this genetic distance data, *P. caribaea* var. hondurensis should be equidistant to *P. elliottii* var. elliottii, and *P. taeda*. Nevertheless, morphological and biochemical data supports a close relationship between *P. caribaea* var. hondurensis and *P. elliottii* var. elliottii, and they were once classified as a single species (Little and Dorman 1954; Nikles 1966). The affinity of *P. elliottii* var. elliottii to *P. caribaea* var. hondurensis compared with *P. taeda* was also evident in studies

of chloroplast DNA variation (Nelson et al. 1994). *Pinus elliottii* var. elliottii shared its second most common haplotype with *Pinus caribaea* var. hondurensis, but neither taxa shared haplotypes with *P. taeda*. Hence, in our study, it was expected that *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis would show a greater similarity in DNA sequence to each other than to *P. taeda*.

Most of the variation in allele length was attributed to changes in the number of repeat units in our study. Where the repeat structure differed in the nonfocal samples from focal taxa, the change was consistent between the two nonfocal taxa. Similarly, the majority of flanking sequence single-base pair insertions or deletions and substitutions were constant between *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis. Our data demonstrated a strong affinity between *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis compared to *P. taeda*, consistent with previous phylogenic analysis. Microsatellite repeat structure and flanking sequence mutations should be useful for phylogenetic analysis within these groups. Repeat structure in microsatellites and flanking sequence mutations has been found to be phylogenetically informative across a subfamily of wasps (Zhu et al. 2000).

No evidence of ascertainment bias for *P. taeda* microsatellite loci

Ascertainment bias is the reduction in allele length or variability in the nonfocal species (Ellegren et al. 1995). It is attributed to an artifact of the microsatellite marker development process such that loci with large numbers of repeats tend to be preferentially isolated and cloned. Evidence for ascertainment bias from reciprocal studies has been found for *Drosophila* spp. (Hutter et al. 1998) and swallows (Kirchman et al. 2000) but did not fully explain the reduction in allele sizes of human microsatellites transferred to chimpanzees (Cooper et al. 1998) or may only be a minor affect in other animals (Crawford et al. 1998). We found no evidence for ascertainment bias for markers developed from total genomic or low-copy libraries in *P. taeda* as there was no reduction in the mean H values in the nonfocal population. This lack of significant loss in variability may reflect the relative closeness of the taxa in this study, as there should be a negative relationship between genetic distance and the degree of polymorphism with ascertainment bias (Ellegren et al. 1995).

Microsatellite markers from other hard pines transferred equally to *P. elliottii* var. *elliottii* and *P. caribaea* var. hondurensis. Variation in the structure and mutations at these loci are consistent with current phylogenetic relationships, and the lack of ascertainment bias for markers is promising for exchanging markers in population and mapping studies in an important group of tropical pines.

Microsatellite marker transfer was an efficient strategy for obtaining polymorphic markers among closely related taxa. Microsatellite markers from *P. contorta* (section Pinus, subsection Contortae) (Hicks et al. 1998) and *P. sylvestris* (section Pinus, subsection Sylvestres) (Kostia et al. 1995; Soranzo et al. 1998) may also be useful in *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis, although more distant than *P. radiata* (Krupkin et al. 1996).

Acknowledgements The authors thank P. Toon and L. Perkins for assistance in collecting *P. elliottii* var. elliottii and *P. caribaea* var. hondurensis samples, A. Cameron and M. Devey for tissue or DNA of *P. radiata* and *P. strobus*, respectively, and B. Potts, M. Rossetto, R. Stokoe and K. Shepherd for comments on early versions of the manuscript. This work was supported by the Cooperative Research Centre for Sustainable Production Forestry. Experiments complied with the current laws in Australia.

References

Chakraborty R, Kimme1 M, Stivers DN, Davison LJ, Deka R (1997) Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proc Natl Acad Sci USA 94:1041–1046

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:713–722

Cooper G, Rubinsztein DC, Amos W (1998) Ascertainment bias cannot entirely account for human microsatellites being longer than their Chimpanzee homologues. Hum Mol Genet 7:1425–1429

Crawford AM, Kappes SM, Paterson KA, Degotari MJ, Dodds KG, Freking BA, Stone RT, Beattie CW (1998) Microsatellite evolution – testing the ascertainment bias hypothesis. J Mol Evol 46:256–260

Crow JF (1986) Basic concepts in population, quantitative and evolutionary genetics. Freedman, New York

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

Dvorak W, Jordan A, Hodge G, Romero J (2000) Assessing evolutionary relationships of pines in the Oocarpae and Australes subsections using RAPD markers. New For 20:163–192

Echt C, Burns R (1999) SSR derived from *Pinus taeda* ESTs. http://dendromeucdavisedu/Data/primerhtml

Echt CS, Maymarquardt P, Hseih M, Zahorchak R (1996) Characterisation of microsatellite markers in eastern white pine. Genome 39:1102–1108

Echt CS, Vendramin GG, Nelson CD, Marquardt P (1999) Microsatellite DNA as shared genetic markers among conifer species. Can J For Res 29:365–371

Ellegren H, Primmer C, Sheldon B (1995) Microsatellite 'evolution': directionality or bias. Nat Genet 11:360–362

Elsik CG, Minihan VT, Hall SE, Scarpa AM, Williams CG (2000) Low-copy microsatellite markers for *Pinus taeda* L. Genome 43:550–555

Fisher PJ, Richardson TE, Gardner RC (1998) Characteristics of single- and multi-copy microsatellites from *Pinus radiata*. Theor Appl Genet 96:969–979

Graham GC, Henry RJ, Godwin ID, Nikles DG (1996) Phylogenetic position of Hoop Pine. Aust Syst Bot 9:893–902

Hicks M, Adams D, O'Keefe S, Macdonald E, Hodgetts R (1998) The development of RAPD and microsatellite markers in lodgepole pine *(Pinus contona* var. latifolia). Genome 41:797–805

Hutter CM, Schug MD, Aquadro CF (1998) Microsatellite variation in *Drosophila melanogaster* and *Drosophila simulans*: a reciprocal test of the ascertainment bias hypothesis. Mol Biol Evol 15:1620–1636

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

Kirchman JJ, Whittingham LA, Sheldon FH (2000) Relationships among cave swallow populations (*Petrochelidonfulva*) determined by comparisons of microsatellite and cytochrome *b* data. Mol Phylo Evol 14:107–121

Kostia S, Varvio SL, Vakkari P, Pulkkinen P (1995) Microsatellite sequences in a conifer, *Pinus sylvestris*. Genome 38:1244–1248

Krupkin AB, Liston A, Strauss SH (1996) Phylogenetic analysis of the hard pines (*pinus* subgenus *Pinus*, Pinaceae) from chloroplast DNA restriction site analysis. Am J Bot 83:489–498

Kutil B, Williams CG (2001) Triplet-repeat microsatellites shared among hard and soft pines. J Hered (in press)

Lamb A (1973) *Pinus caribaea*: volume 1: fast-growing timber trees of the lowland tropics. Trop For Pap no. 6, University of Oxford, Oxford

Little L, Critchfield W (1969) Subdivisions of the Genus *Pinus*. Report no. 1144, US For Serv Misc Publ

Little E, Dorman K (1954) Slash pine (*Pinus elliottii*), including south Florida slash pine: nomenclature and description (SE-36). USDA For Serv Res Pap

Miller C (1977) Mesozoic conifers. Bot Rev 43:217–280

Miller C (1993) Impact of the Ecocene on the evolution of *Pinus* L. Ann Mo Bot Gard 80:471–498

Mirov NT (1967) The Genus *Pinus*. Ronald Press, New York

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

Nelson CD, Nance WL, Wagner DB (1994) Chloroplast DNA variation among and within taxonomic varieties of *Pinus caribaea* and *Pinus elliottii*. Can J For Res 24:424–426

Nikles D (1966) Comparative variability and relationships of Caribbean pine (*Pinus caribaea* Mor.) and slash pine *(Pinus elliottii* Engelm.) PhD thesis, North Carolina State University, Raleigh, N. C.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Mol Bio1 Evol 15:1275–1287

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

Schug MD, Hutter CM, Wetterstrand KA, Gaudette MS, Mackay TFC, Aquadro CF (1998) The mutation rates of di-, tri- and tetranucleotide repeats in *Drosophila melanogaster*. Mol Bio1 Evol 15:1751–1760

Scott LJ, Cross M, Shepherd M, Maguire T, Henry W (1999) Increasing the efficiency of microsatellite discovery from poorly enriched libraries in coniferous forest species. Plant Mol Bio1 Rep l7:351–354

Smith DN, Devey ME (1994) Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37:977–983

Soranzo N, Provan J, Powell W (1998) Characterisation of microsatellite loci in *Pinus sylvestris* L. Mol Ecol 7:1247–1263

Thompson J, Higgins D, Gibson T (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Weber J (1990) Informativeness of human (dC-dA)n - (dG-dT)n polymorphisms. Genomics 7:524–530

Zhu Y, Queller DC, Strassmann JE (2000) A phylogenetic perspective on sequence evolution in microsatellite loci. J Mol Evol 50:324–338