Theor Appl Genet (2001) 103:1189–1195 © Springer-Verlag 2001

C.G. Elsik · C.G. Williams

Low-copy microsatellite recovery from a conifer genome

Received: 19 April 2000 / Accepted: 27 February 2001

Abstract Microsatellite development has been stymied by highly repetitive DNA in the large, highly duplicated conifer genome and by so few genomic conifer sequences in public databases. Recovery of microsatellites from the low-copy component was tested as an efficient approach to marker development. Microsatellites were isolated from *Pinus taeda* L. via low-copy enrichment and filter-hybridization of tri- and tetra-nucleotide repeat motifs. Efficiency at three phases of marker development was compared for low-copy and total-genome control libraries. In the first phase, enrichment for microsatellites was slightly lower in the low-copy libraries. In the second phase, redundancy was higher in the low-copy libraries. In the third phase, low-copy libraries provided more polymorphic markers than total-genome libraries. Of 418 sequenced low-copy clones, 102 were unique sequences with repeat motifs. Of these unique sequences, twice as many were useful for marker development compared to the total-genome control. Difficulty in microsatellite marker development due to highly repetitive DNA can be abated by low-copy enrichment or circumvented by selecting for specific CG-rich trinucleotide repeat motifs. Sixteen new low-copy and genomic *P. taeda* microsatellites were given as an example.

Keywords Low-copy kinetic component · *Pinus taeda* · Reassociation kinetics · Triplet repeat sequences · Gymnosperms

Introduction

Highly repetitive DNA has been reported as the major obstacle to recovering microsatellites from genomic libraries for several conifer species (Smith and Devey 1994; Kostia et al. 1995; Pfeiffer et al. 1997; Fisher et al. 1998). The conifer genome requires special tools for microsatellite development. Approaches for circumventing this problem include probing cDNA libraries for repeat motifs, which has resulted in very low yields (Scotti et al. 2000). Using large-insert BAC libraries as proposed by Cardle et al. (2000) is problematic for conifers due to their large genome sizes. Due to genome size, only about 5% of the pine genome has been covered by a BAC library.

Of the conifers, pines have the largest genomes with 19 to 31 pg per haploid nucleus (Murray 1998). The pine genome is composed of 75–86% highly repetitive DNA (Kriebel 1985; Elsik and Williams 2000). Using *Pinus taeda* L., our approach was to recover sequences from the low-copy kinetic component and then enrich them for specific repeat motifs. The aim of this study is to test low-copy enrichment and subsequent recovery of low-copy microsatellites against genomic library construction. Motif enrichment, marker recovery and marker quality were compared for three pairs of low-copy and total-genome control libraries. Low-copy enrichment is a novel protocol with potential application to other large plant genomes.

Communicated by A. Kahler

C.G. Elsik, C.G. Williams (✉)
Texas A&M University, 305 Horticulture Building,
College Station, Texas 77843-2135, USA
e-mail: claire-williams@tamu.edu
Tel.: +1-979-862-3745, Fax: +1-979-845-6049

Materials and methods

A general overview of low-copy enrichment

Low-copy enrichment began with dividing nuclear DNA into two pools (Fig. 1). The first pool was prepared for cloning by fragmenting, then ligating linkers to nuclear DNA. The second pool was composed of low-copy DNA fragments, prepared by removing repetitive DNA from nuclear DNA. Removal of repetitive DNA was done by partial self-annealing to C_0T 1000 followed by hydroxyapatite (HAP) chromatography. These low-copy fragments were then biotinylated and used as a “fish-hook” (*sensu* Lovett 1994) to capture low-copy sequences from the first pool. Streptavidin-coated magnetic beads were then used to sequester the biotinylated hybrid fragments (e.g., Morgan et al. 1992). Low-copy linker-ligated DNA was eluted from the hybrid molecules, and then amplified by PCR using lone-linker PCR (e.g., Ko et al. 1990). A more-detailed description is given in Elsik (2000).

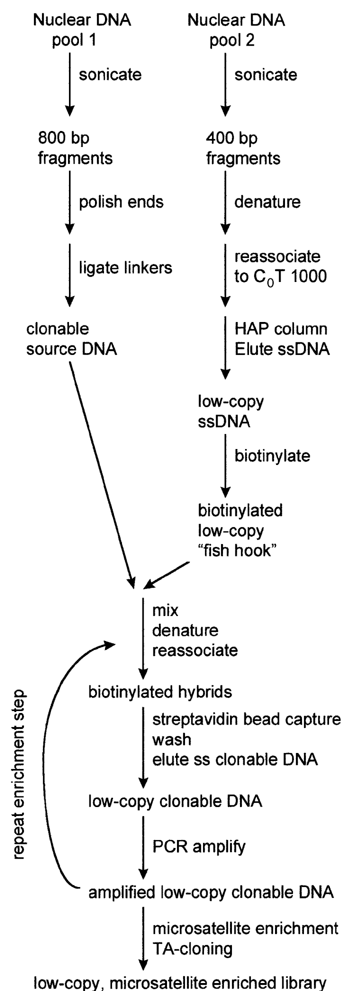

Fig. 1 A schematic diagram of the two-step low-copy enrichment method is shown here. The low-copy enrichment begins with the preparation of the fish-hook. The low-copy fish-hook is then mixed with genomic DNA to capture biotinylated hybrid DNA

Plant material and DNA extraction

Nuclear DNA was isolated from fresh needle tissue of *P. taeda* parent 11–1060 after isolating nuclei and organelles using a modification of Wagner et al. (1987). Pellets were incubated for 10 min in wash buffer (Britten et al. 1974) containing 0.5% Triton-X-100 to lyse chloroplasts and mitochondria. The resulting nuclear pellets were washed three times in wash buffer, prior to continuing with the modified CTAB procedure.

Low-copy enrichment: preparation of clonable DNA

Nuclear DNA was sonicated to fragments smaller than 1,200 bp and purified of metal ions (Werman et al. 1996). Fragments smaller than 400 bp were removed using glassmilk (Geneclean, Bio 101) resulting in an average size of 800 bp. Fragment ends were polished and then ligated to phosphorylated linkers (5´-pTA-GTCCACGCGTAAGCAAGAGCACA-3´/3´-ATCAGGTGCGCA-TTCGTTCTCG-5´) (Edwards et al. 1996). This was the pool of clonable DNA used as the source in the low-copy enrichment and the DNA used to construct non-enriched genomic libraries.

Low-copy enrichment: fish-hook preparation

Reassociation conditions were chosen based on the *Pinus strobus* C_0T curve (Kriebel 1985). Nuclear DNA was sonicated to an average fragment size of 400 bp and purified of metal ions (Werman et al. 1996). After removing fragments smaller than 400 bp with Glassmilk, fragments were re-suspended in 0.48 M sodium phosphate buffer, pH 6.8 (PB). A self-annealing reaction containing 12.8 of µg DNA in 40 µl of 0.48 M PB was boiled for 5 min, then incubated at 60°C for 25 h to C_0T 1,000.

Hydroxyapatite (Bio-Gel-HTP, Bio-Rad) was prepared in 0.14 M PB pH 6.8 (Britten et al. 1974). The hydroxyapatite column was equilibrated to 50°C. The annealing reaction was adjusted to 0.14 M PB, equilibrated to 50°C, loaded onto the HAP column and incubated at 50°C for 5 min. Single-stranded DNA was eluted using 0.14 M PB. The eluate was concentrated to a volume of 100 µl by butanol extraction and then de-salted using a Sephadex-G50 column (Sambrook et al. 1989). The resulting pool of low-copy DNA was ethanol-precipitated and re-suspended in 100 µl of de-ionized water. This low-copy single-stranded DNA was biotinylated by random prime labeling using the BioPrime DNA Labeling System (Life Technologies) which both labels, and amplifies the DNA. Non-incorporated nucleotides were removed using a Sephadex-G50 column.

To verify the removal of repetitive sequences from the fish-hook, a dot blot was used to compare single-stranded fish-hook DNA eluted from the HAP column with a control, total genomic DNA. Both dots were probed with nick-translated ^{32}P-labelled genomic DNA. The presence of repetitive sequences increased the signal intensity. DNA samples (400 ng) in 10×SSC were blotted onto a Hybond-N+ nylon membrane (Amersham Pharmacia Biotech) using a Bio-Dot microfiltration unit (Bio-Rad Laboratories). Filters were prehybridized, hybridized and washed using tetramethylammonium chloride (TMAC), eliminating the T_m difference between AT-rich and CG-rich sequences (Wood et al. 1985). The blot was pre-hybridized and hybridized at 75°C with the addition of herring sperm DNA, and then washed at 85°C (Wood et al. 1985). The ^{32}P-labeled nick translated genomic DNA probe was prepared using the Promega Nick Translation System (Promega). Blots were exposed to Biomax MS autoradiography film (Eastman Kodak) for 24 h at –80°C using a Biomax MS intensifying screen.

Low-copy enrichment: capture of clonable DNA using the fish-hook

A hybridization reaction was prepared containing a mixture of 250 µg of biotinylated low-copy DNA (fish-hook), 2.5 µg of clonable source DNA, 10 µg of the 22-mer linker oligo, 6×SSC, 1% SDS in a total volume of 500 µl. Low-copy fish-hook DNA was used in 100-fold excess over source DNA to prevent the repetitive sequences in the source DNA from driving the reaction. The solution was incubated at 95°C for 5 minutes and then incubated at 60°C for 24 h.

Streptavidin-coated magnetic beads (Promega) (0.6 ml) were prepared in 6×SSC. The hybridization mix was added to the beads and mixed for 15 min. The beads were washed, using a magnetic bead stand, twice for 15 min at room temperature in 6×SSC, 0.1% SDS, three times for 15 min at 60°C in 6×SSC, 0.1% SDS, and then rinsed twice at room temperature in 6×SSC.

Single-stranded clonable low-copy DNA was eluted from the beads in 100 µl of 0.1 N sodium hydroxide at 60°C for 5 min, followed by the addition of 100 µl of 1 M Tris, pH 7.5. The resulting solution of low-copy clonable DNA was de-salted and concentrated to 20 µl in de-ionized water in an Ultrafree-MC filter (Millipore).

Low-copy clonable DNA was amplified in 50 µl, containing 4 µl of magnetic bead eluate, 400 nM of linker 22-mer oligo, 200 µM of each dNTP, 1 unit of *Taq* polymerase (Life Technologies), 1×PCR buffer (Life Technologies), 1.5 mM $MgCl_2$, 0.1 mg/ml of gelatin. PCR took place in a Robocycler thermal cycler with the following protocol: 3-min denaturation 95°C; 1 min at 95°C,

1 min at 60°C, 2 min at 72°C (30 cycles); 5 min extension at 72°C.

A portion of the PCR product (36 ul) was used in a second biotin-streptavidin enrichment cycle. After eluting, de-salting and concentrating, the twice-enriched low-copy clonable DNA fragments were PCR amplified as before. The resulting fragments were then used in microsatellite enrichment.

Microsatellite enrichment: filter hybridization

Microsatellite enrichment used the modifications of Karagyozov et al. (1993) and Edwards et al. (1996). Repeat motif oligos were grouped according to T_m to allow wash stringencies to be optimized. Filters were prepared for the enrichment of three microsatellite motif sets. Although $(GTTT)_n$ is a tetranucleotide motif, it was grouped with the AT-rich trinucleotide motifs because of similarity in T_m.

For each oligo group, there were two libraries: one using low-copy clonable DNA (the low-copy libraries) and the other using total genomic clonable DNA (the control libraries). The CG-rich low-copy microsatellite and control microsatellite libraries used an oligo group composed of $(CCT)_{10}$, $(CGT)_{10}$, $(GCT)_{10}$, $(GGT)_{10}$. The AT-rich low-copy microsatellite and control microsatellite libraries used an oligo group composed of $(AGT)_{10}$, $(CTT)_{10}$, $(GAT)_{10}$, $(GTT)_{10}$, $(GTTT)_{10}$. The tetranucleotide libraries used an oligo group composed $(CGTT)_{10}$, $(CTGT)_{10}$, $(AGGT)_{10}$, $(GAGT)_{10}$, $(GGAT)_{10}$. Low-copy microsatellite libraries were constructed with low-copy clonable target DNA. Control microsatellite libraries were constructed with PCR-amplified clonable source DNA which had not been enriched for low-copy sequences.

Two 0.5 cm^2 pieces Hybond-N+ (Amersham) filters were prepared for each library (four pieces for each group of oligos). For each filter, 325 ng of each oligo were pooled in 80 µl 10×SSC and spotted onto a filter. Filters were air-dried for 1 h, then baked at 80°C for 2 h and UV crosslinked. The filters were washed twice in 10 ml of hybridization buffer (50% formamide, 5×SSC, 50 mM sodium phosphate pH 7.0, 1% SDS) at 37°C for 48 h. Each filter was then boiled in 500 µl 1% SDS for 5 min, removed from solution and stored in a microfuge tube at –20°C until use.

Four 300-µl-hybridization reactions (one per library) contained 50% formamide, 5×SSC, 0.5% SDS, 50 mM sodium phosphate, 2 µg of the 21-mer linker oligo, 43 µl of the appropriate PCR product (either low-copy or total genomic), which had been denatured by boiling for 5 minutes. Hybridization took place at 37°C for 2 days. Following hybridization, filters were washed in 500-µl volumes at varying wash stringencies for 5 min in each wash, starting at low stringency (5×SSC, 0.1% SDS, room temperature) and ending with high stringency [0.1×SSC, 0.1% SDS, 65°C (CG-rich and tetranucleotide libraries) or 45°C (AT-rich libraries)]. Filters were boiled in 500 µl of 0.1% SDS for 10 min to elute microsatellite-enriched DNA. SDS was removed from eluates using Ultrafree-MC filters (Millipore), concentrating the solutions to 20 µl in de-ionized water. Microsatellite-enriched DNA was PCR amplified in a 50 µl volume using 4 µl of concentrated eluate as the template and the same reaction mix and conditions as used in the previous PCR amplification. The resulting PCR product was used in a second round of microsatellite enrichment.

Library construction

Libraries were constructed using PCR fragments. Six libraries, two low-copy and two control libraries, were constructed after microsatellite enrichment for CG-rich, AT-rich or tetranucleotide repeats. PCR fragments were ligated into the plasmid pCR2.1 (Invitrogen TA Cloning Kit). Plasmids were used to transform TOP10F´ cells (Invitrogen). Colonies were picked for long-term storage in 384-well plates. To compare low-copy and control libraries, microsatellite development was divided into three phases: (1) microsatellite enrichment, (2) sequence redundancy, and (3) marker quality.

Phase 1: microsatellite enrichment

A Beckmann robot was used to replicate cultures from 384-well plates to filters for library screening. Filters were prepared, hybridized with the appropriate ^{32}P-labelled microsatellite oligos and exposed to autoradiography film (Sambrook et al. 1989). Microsatellite enrichment was scored as the number of clones positive for microsatellite motifs in each library.

Phase 2: redundancy and recovery of sequences suitable for primer design

A sample of positive clones from each library were sequenced. Sequence redundancy was identified through contig analysis using a minimum match set at 95%. Duplicate and overlapping sequences from each library were identified using the Seqman II contig assembly program (DNAStar, Madison, Wis.). Three factors accounted for the contig grouping: (1) exact replicates, (2) overlapping fragments along the chromosomes or (3) similar repetitive sequences.

Microsatellite recovery was tabulated for all six libraries, low-copy and control alike. The sequences were subjected to contig analysis and grouped according to unique contigs (see Table 1). The number of contigs containing a sequence with sufficient flanking region for primer design was recorded. Unsuitable flanking sequences were defined as flanking sequences containing direct or inverted repeats that did not provide unique sites for primer annealing.

Phase 3: marker quality

Only one sequence per contig was selected for primer design if flanking regions were sufficiently long. Contigs without a single sequence with a sufficient flanking region for primer design were not considered for marker development (Table 1). Primer design was done with the PrimerSelect program (DNAStar, Madison Wis.). After testing for amplification and temperature screening using *P. taeda* 11–1060 genomic DNA, primers that resulted in amplification were tested for polymorphism and Mendelian inheritance using the grandparents, parents and ten progeny of a three-generation outbred loblolly pine pedigree which included *P. taeda* 11–1060 (Devey et al. 1991). Markers were verified to be single-locus using haploid megagametophytes of *P. taeda* parent 11–1060. Amplification conditions for low-copy *P. taeda* microsatellite markers are given in Elsik et al. (2000). PCR products were visualized on 6% denaturing polyacrylamide gels by silver staining using modifications of Elsik et al (2000).

Microsatellite quality was categorized as follows. Class 1 markers had polymorphic Mendelian inheritance with a single-locus product or two clearly discernible loci. Class 2 markers had a monomorphic banding pattern where the same PCR product was amplified among all members. Class 3 markers had polymorphic non-Mendelian inheritance where the PCR product amplified clearly discernible codominant loci with at least one aberrant allele among the parents or progeny. Class 4 markers had multiple amplification products with patterns of inheritance between multiple loci which were not discernible. Class 5 markers had poor amplification which was defined as a high degree of non-specific amplification, inconsistent amplification between samples or PCR products outside the expected size. Marker informativeness was based on Williams (1998).

Results

Recovery using the low-copy libraries yielded more polymorphic markers with Mendelian inheritance than the control genomic libraries. This occurred despite lower efficiency in two out of three phases of microsatellite development. In the first phase, the capture step did

use low-copy fragments which showed removal of repetitive DNA in the pooled dot blot. Subsequent microsatellite enrichment was slightly lower in the low-copy libraries. Also, low-copy libraries had clone-specific redundancy. In the second phase, the proportion of useful sequences was higher for the low-copy libraries, especially the AT-rich library. In the third phase, the low-copy library method yielded two to three times more highly informative class-1 markers compared to the control libraries (Tables 1–3).

Low-copy capture, microsatellite enrichment and clone-specific redundancy

The fish-hook DNA pool had a low-intensity signal when probed with total genomic DNA, showing removal of repetitive DNA (Fig. 2). The pooled dot blot confirmed that the aggregate pool of fish-hook fragments used in the capture step were indeed low-copy. The filter hybridization method proved effective for microsatellite enrichment with 32% microsatellite-positive clone frequencies for both CG-rich trinucleotide repeats and tetranucleotide repeats. Enrichment frequency increased at least 267- and 533-fold, respectively, for CG-rich repeats. The filter enrichment resulted in a 34-fold increase in microsatellite-rich clone frequency for AT-rich trinucleotide motifs. Enrichment was slightly higher for the CG-rich total-genome library (32%) and the tetranucleotide library (38.3%) (Table 1).

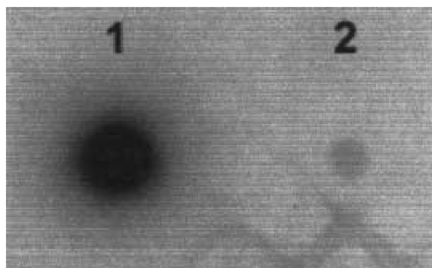

Fig. 2 Dot blot of pooled fragments. *Lane 1* is the total genomic DNA and *lane 2* is the fish-hook DNA. Both pools were probed with nick-translated total genomic DNA

Redundancy in the low-copy CG-rich microsatellite library was nearly twice as high as the control (Table 1). Most sequences were exact replicates, and thus redundant clones. Redundancy was clone-specific: three sequences were found in 85, 33 and 23 copies in a total of 206 clones. The other 65 sequences occurred on an average of 1.6 copies. Similarly, the other two low-copy libraries also had clone-specific redundancy but the low-copy AT-rich microsatellite library had only a single redundant sequence with 15 copies. The low-copy tetranucleotide microsatellite library contained two sequences with 48 and 28 copies, respectively, out of a total of 168 clones. Clone-specific redundancy did not occur in any of the total genome control libraries.

Recovery of sequences suitable for primer design

Recovery of useful sequences was higher for the combined control libraries. The control libraries had 71 unique contigs of which 37 could be used for primer design (Table 1). Only 29 of the 102 unique contigs in the low-copy library could be used for primer design due to short flanking sequences. The absence of suitable primer binding sites also reduced the number of useful sequences but this factor affected both types of libraries equally. Direct repeats reduced the number of suitable primer binding sites. Nine CG-rich microsatellite sequences had direct repeats ranging in length from 12 to 144 bp which flanked the repeat region. Five of these sequences contained direct repeats which precluded primer design.

Marker quality

In the third phase, the low-copy libraries yielded more polymorphic markers with Mendelian inheritance than the control libraries (Table 2). Primer sets for 25 out of 29 low-copy sequences amplified via PCR and 15 of 25 were polymorphic with Mendelian inheritance (Table 2). The control libraries had only 17 out of 37 sequences which amplified via PCR and, of these, only eight could be developed into markers. The proportion of microsatel-

Table 1 Recovery of sequences suitable for primer design. Redundancy decreased the proportion of sequences suitable for primer design in the low-copy libraries

STAGE	Low-copy library			Total-genome library		
	CG-rich	AT-rich	Tetra	CG-rich	AT-rich	Tetra
Total clone count	2169	511	548	384	384	384
Microsatellite-positive	521	46	181	123	17	147
% Positive	24.0	9.0	33.0	32.0	4.4	38.3
Sequence sample	206	44	168	59	15	50
Unique sequences	43	25	61	38	15	27
Unique contigs (UC)	39	22	50	36	15	20
UC with microsatellites	36	17	49	36	14	19
Sequences with primers	13	8	8	24	6	7
Useful sequences	13/206 (6%)	8/44 (18%)	8/168 (5%)	24/59 (41%)	6/15 (40%)	7/50 (14%)

Table 2 Low-copy enrichment more than doubled recovery for each microsatellite library compared to its total-genome control. Class 1 represented a high-quality marker: polymorphic Mendelian inheritance with a single-locus product or two clearly discernible loci. Class 2 had monomorphism. Class 3 had polymorphic non-Mendelian inheritance. Class 4 had multiple amplification products. Class 5 had poor amplification

Type	Low-copy *n*=29			Genomic *n*=37		
	CG-rich library *n*=13	AT-rich library *n*=8	Tetra library *n*=8	CG-rich library *n*=24	AT-rich library *n*=6	Tetra library *n*=7
Class 1	6	6	3	5	2	1
Class 2	3	0	1	2	1	0
Class 3	0	0	1	3	1	0
Class 4	2	1	1	1	0	1
Class 5	2	1	2	13	2	5
Class 1 markers (%)	46%	75%	38%	21%	33%	14%

lite markers was higher in the low-copy compared to the control, an observation which is consistent with enrichment for lower-copy number sequences. As an example, 16 polymorphic low-copy and genomic *P. taeda* microsatellites are given in Appendix 1. These microsatellites were developed from several *P. taeda* low-copy and genomic libraries including the exact paired libraries described for phases 1–3.

Discussion

Developing microsatellites from a low-copy library was more efficient than using total-genome libraries despite higher sequence redundancy and lower recovery. Low-copy enrichment increased marker recovery. By adding steps to decrease sequence redundancy, low-copy enrichment could increase marker recovery per primer set tested.

Low-copy enrichment gave the greatest benefit to the development of AT-rich trinucleotide and tetranucleotide microsatellite markers. Lower enrichment in the AT-rich libraries was attributed to the lower T_m of these motifs. Lower wash temperatures after filter hybridization caused increased non-specific binding to filters which in turn diluted the capture of sequences with AT-rich repeat motifs.

The value of low-copy enrichment was lower for CG-rich motifs than for the other two libraries. Possibly CG-rich trinucleotide repeats are sequestered in the low-copy kinetic component or near genic regions. If so, enriching for sequences with CG-rich motifs may be an efficient alternative to using the low-copy protocol and screening against sequence redundancy. Either low-copy enrichment with microsatellite selection, or even selecting for microsatellite motifs preferential to low-copy DNA, should improve the recovery of useful microsatellite markers for large plant genomes.

High sequence redundancy in the low-copy library method was attributed to the higher number of PCR steps compared to the control library. Although there are families of clustered microsatellites, members of these families have degenerate flanking regions as well as variable motif composition and length (Elsik and Williams 2001). The low-copy libraries had four PCR amplification steps compared to two in the control libraries. Redundancy could be reduced by decreasing the number of cycles per PCR step or by decreasing the number of enriching rounds. Fewer redundant clones would be sequenced if only a sample of clones was initially sequenced. The library could then be screened with the highly redundant clones prior to large-scale sequencing.

Marker recovery was higher for the low-copy method compared to microsatellite-enriched *Pinus sylvestris* libraries (Soranzo et al. 1998). Low-copy microsatellite development was more efficient than marker development from *P. taeda* RFLP probes (Harry et al. 1998). To-date, markers developed from RFLP sampled from conserved regions of the conifer genomes show low levels of polymorphism and few alleles per locus. For example, eight PCR-based *P. taeda* were developed from RFLP probes and all markers had the least-informative backcross (BC) mating-type configuration in the public RFLP pedigree (Harry et al. 1998).

The low-copy method recovered more markers with higher levels of informativeness compared to probing for repeats in *P. sylvestris* cDNA libraries (Scotti et al. 2000). Developing microsatellite markers from 50,000 *P. sylvestris* cDNA clones yielded 23 positive clones of which only four (17.4%) markers were polymorphic. These markers had 3 to 11 alleles per locus. By comparison, there were 102 low-copy clones where unique sequences with repeat motifs were found among a total of 418 low-copy clones. Of these 102 clones, 29 (28.4%) proved useful for marker development. Of these, the resulting low-copy microsatellite had 3 to 29 alleles per locus and displayed the more-highly informative intercross mating-type configurations (Elsik et al. 2000). Difficulty in microsatellite marker development due to highly repetitive DNA can be abated by low-copy enrichment prior to microsatellite selection.

Acknowledgements This work was sponsored by the Texas Higher Education Coordinating Board grants 999902–090 and 999902–201, PE Celera AgGen, The Timber Company and the USDA-Forest Service Challenge Grant Program. Germplasm was supplied by Westvaco Corporation and Weyerhaeuser Company. Special thanks to Dr. Craig Echt, Forest Research Institute (NZ) and to Dr. Keith Edwards, University of Bristol (UK), for their assistance. All experiments were performed according to U.S. laws.

Appendix 1 Sixteen di-, tri- and tetra-nucleotide microsatellites are given as an example of polymorphic markers recovered from several low-copy and genomic *P. taeda* libraries. Year of library is denoted by the first digit (2=Year 1; 3=Year 2). Type of library is denoted by the second digit (0=low-copy, 1=total genomic).

Locus	Primer sequence (5´-3´)	Size (bp)	Tm (°C)	GenBank accession #
PtTX2033	F: CATTCCTACAAAACTTCTAAATTAA R: CCATATTTGATGCGTTGATT	258	55	AF333774
PtTX2090	F: CCCGCCTATTCCACCTA R: CTACACATTTCACCCATAAGTCC	283	59	AF333775
PtTX2091	F: ACCAAATCTCCCCACAT R: AATCATACCCGTTTCAGT	267	59	AF333776
PtTX2159	F: TCAGCTCCCCAATGCCCTCCACTCA R: CAGCTGCTCTTTCCAACCTGTCACC	167	61	AF333777
PtTX2183	F: TTAGTTGCAAAGAATATTTAAGGT R: CCTGCACTAGCTTTATATTTCATA	205	55	AF333778
PtTX3005	F: TGTTGATGATGAGGATGACGA R: CATTAATTTAGTGTGGCTTTTT	80	61	AF333779
PtTX3022	F: CTCGCGGTAGTAATCTT R: CGAGTAGTAGGCGTATCT	292	57	AF333780
PtTX3045	F: CATCGCATATCGCAATCAGG R: ATCGGAGTCAAAACACAAAAGAAA	226	55	AF333781
PtTX3046	F: TATAGCTAGACCCGAAACA R: AGACCCGCTGGCATTAT	162	55	AF333782
PtTX3096	F: TAATTGGTTATCATTTGTCTTT R: CATTGACTTAAAATCCATACAT	307	57	AF333783
PtTX3105	F: TGTCGGTGGAGTTGGCAGTAGACT R: AGGGCCCAGCGTTTCCTG	258	55	AF333784
PtTX3107	F: AAACAAGCCCACATCGTCAATC R: TCCCCTGGATCTGAGGA	217	55	AF333785
PtTX3112	F: AAAAGGGCCTCAAAGAAAAAT R: ATAGGGAGATAAGTTGAAAATA	161	65	AF333786
PtTX3116	F: CCTCCCAAAGCCTAAAGAAT R: CATACAAGGCCTTATCTTACAGAA	191	55	AF333787
PtTX3122	F: AAATCAAAGCAGCTAGAAAGTGT R: AATATGCCTGAGGTTGGTTAC	324	65	AF333788
PtTX3127	F: ACCCTTACTTTCAGAAGAGGATA R: AATTGGGGTTCAACTATTCTATTA	265	55	AF333789

References

Britten RJ, Graham DE, Neufeld BR (1974) Analysis of repeating DNA sequences by reassociation. Methods Enzymol. 29:363–418

Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R (2000) Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156:847–854

Devey ME, Fiddler T, Liu B, Knapp S, Neale DB (1994) An RFLP map for loblolly pine based on a three-generation outbred pedigree. Theor Appl Genet 88:273–278

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques 20:758–760

Elsik CG (2000) The low-copy component of the *Pinus taeda* genome. PhD dissertation, Texas A&M University, College Station Texas

Elsik CG, Williams CG (2001) Families of clustered microsatellites in a conifer genome. Gen Genet 265:535–542

Elsik CG, Williams CG (2000) Retroelements contribute to the excess low-copy number DNA in pine. Mol Gen Genet 264:47–55

Elsik CG, Minihan VT, Scarpa AM, Hall SE, 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

Harry DE, Temesgen B, Neale DB (1998) Codominant PCR-based markers for *Pinus taeda* developed from mapped cDNA clones. Theor Appl Genet 97:327–336

Karagyozov L, Kalcheva ID, Chapman VM (1993) Construction of random insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Res 21:3912–3913

Ko MSH, Ko SBH, Takahashi N, Nishiguchi K, Abe K (1990) Unbiased amplification of a highly complex mixture of DNA fragments by 'lone linker'-tagged PCR. Nucleic Acids Res 18:4293–4294

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

Kriebel HB (1985) DNA sequence components in the *Pinus strobus* nuclear genome. Can J For Res 15:1–4

Lovett M (1994) Fishing for complements: finding genes by direct selection. Trends Genet 10:352–357

Morgan JG, Dolganov GM, Robbins SE, Hinton LM, Lovett M (1992) The selective isolation of novel cDNAs encoded by the regions surrounding the human interleukin 4 and 5 genes. Nucleic Acids Res 20:5173–5179

Murray BG (1998) Nuclear amounts in gymnosperms. Ann Bot 82:3–15

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

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Scotti I, Magni F, Fink R, Powell W, Binnelli G, Hedley PE (2000) Microsatellite repeats are not randomly distributed within Norway spruce (*Picea abies* K.) expressed sequences. Genome 43:41–46

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

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

Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP, Allard RW (1987) Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proc Natl Acad Sci USA 84:2097–2100

Werman SD, Springer MS, Britten RJ (1996) Nucleic acids. I. DNA-DNA hybridization. In: Hillis DM, Moritz C, Mable BK (eds) Molecular systematics, 2nd edn, chapter 6. Sinauer Associates, Sunderland, Massachusetts, pp 169–203

Williams CG (1998) QTL mapping in outbred pedigrees. In: Chapter 5: Molecular dissection of complex traits. Paterson AH (ed) CRC Press, Boca Raton, Florida, pp 81–94

Wood WI, Gitschier J, Lasky LA, Lawn RM (1985) Base composition-independent hybridization in tetramethlyammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. Proc Natl Acad Sci USA 82:1585–1588

Note added in proof: more detailed information on low-copy P. taeda microsatellites may be found at http://forestry.tamu.edu/genetics