Theor Appl Genet (2001) 102:296–306

S.A. Cato · R.C. Gardner · J. Kent · T.E. Richardson

A rapid PCR-based method for genetically mapping ESTs

Received: 4 January 2000 / Accepted: 26 May 2000

Abstract A simple, semi-automatable procedure was developed for converting expressed sequence tags (ESTs) into mappable genetic markers. The polymerase chain reaction is used to amplify regions immediately 5′ or 3′ to the coding regions of genes in order to maximise sequence variability between alleles. Fragment length and nucleotide substitution polymorphisms among amplified alleles can be detected using either ethidium bromide staining or automated laser-based fluorescence. A 6% non-denaturing acrylamide gel, analysed with an ABI 377 DNA sequencer, proved capable of resolving homoduplexes and heteroduplexes formed between amplified alleles containing nucleotide substitutions as well as resolving allelic length differences. With this approach 75% of 60 ESTs from a range of *Pinus* species could be genetically mapped in each of three pedigrees from *P. radiata* and *P. taeda*. Furthermore, three or four alleles were detected in each pedigree for 42% of the EST markers.

Keywords Expressed sequence tag · Polymerase chain reaction · Fragment length polymorphism · Single nucleotide polymorphism · Heteroduplex · *Pinus radiata* · *Pinus taeda*

Introduction

In recent years a number of random and targeted gene discovery programmes have been initiated for a range of species, and this has led to a dramatic increase in the number of expressed sequence tags (ESTs) in public and private databases. EST databases are now becoming the basis for genomic approaches to drug discovery, plant and animal genetic improvement and the study of human genetic diseases (Messing and Llaca 1998; Saier 1998; Picoult-Newberg et al. 1999). These databases are also a potentially valuable source of genetic markers and provide an opportunity to construct syntenic genome linkage maps of expressed genes among related species.

There are a number of advantages to using expressed genes instead of ‘anonymous’ sequences as genetic markers. Firstly, if an EST marker is found to be genetically associated with a trait of interest, it is possible that the mapped gene directly affects the trait. Secondly, ESTs that share homology to candidate genes, or differentially expressed ESTs in a tissue of interest, can be specifically targeted for genetic mapping. Moreover, because ESTs are derived from coding DNA, which generally has a high degree of sequence conservation, EST markers are more likely to be transportable across pedigree and species boundaries than are markers derived from non-expressed sequences, like amplified fragment length polymorphisms (AFLPs), random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) markers. As such, EST markers will be especially useful for aligning genome linkage maps and comparing quantitative trait loci (QTL) information across distantly related species. Likewise, if DNA sequence information is lacking for a target species, ESTs derived from other species could be used as the basis for genetic mapping in the species of interest. Genetic mapping with ESTs would thus enable a more rapid transfer of linkage information between species.

Earlier methods for mapping ESTs involved restriction fragment length polymorphism (RFLP) analysis using ESTs as hybridisation probes. This approach has since been augmented by more efficient polymerase chain reaction (PCR)-based approaches where EST-specific primers are used to amplify coding sequence, followed by either restriction enzyme analysis, heteroduplex analysis or single-stranded conformational polymorphism (SSCP) analysis (Fischer and Lerman 1983; Harry et al.

Communicated by M.A. Saghai Maroof

S.A. Cato (✉) · J. Kent · T.E. Richardson
Forest Research, Private Bag 3020, Rotorua, New Zealand
Fax: +64-7-3435444
e-mail: sheree.cato@forestresearch.co.nz

S.A. Cato · R.C. Gardner
School of Biological Sciences, University of Auckland,
Private Bag 92019, Auckland, New Zealand

1998; Schafer and Hawkins 1998; Orita et al. 1989; Konieczny and Ausubel 1993). If extensive allele sequence data is available, the mapping of characterised single nucleotide polymorphisms (SNPs) is an efficient approach to mapping genes. However, there are currently few documented SNPs available for many plant and animal species. Consequently, we sought to develop an EST mapping method that reveals high levels of non-sequenced polymorphisms and is automatable for high-throughput. Here we report on the development of a PCR-based method for genetically mapping ESTs based on fragment length polymorphisms that can be resolved on automated gel systems. The method detects both insertions and deletions and nucleotide changes in stretches of coding and non-coding DNA adjacent to expressed genes. This paper outlines the development of the technique and evaluates the efficiency and utility of the approach for genetic mapping in two *Pinus* species.

Materials and methods

Isolation of genomic DNA

Pinus radiata needle tissue was collected from two parents, ten megagametophytes and 53 progeny from the control-pollinated full-sibling family 850.055 × 850.096. Genomic DNA was extracted from the tissue using a BIO 101 FastDNA (H) kit and a FastPrep FP120 machine (Savant) following the manufacturer's instructions. Needle tissue was also collected from the parents of the *qtl* and *base P. taeda* full-sibling mapping pedigrees (Devey et al. 1994; Groover et al. 1994) and genomic DNA was extracted from the tissue using a CTAB method (Harry et al. 1998).

Digestion of genomic DNA and ligation of adaptors

One microgram of genomic DNA was cut with 5 U of either *Alu*I, *Dra*I, *Eco*RV or *Ssp*I restriction endonucleases in 40 µl of 1× or 2× OPA+ buffer (Pharmacia) for 1 h at 37°C according to the manufacturer's instructions. Next, 10 µl of a solution containing 100 pmol. of GenomeWalker adaptor (Clontech), 1 U T4 DNA ligase, 5 m*M* ATP in 1× or 2× OPA+ buffer (Pharmacia) was added and incubated at room temperature overnight. The reaction mixture was diluted tenfold with 10 m*M* TRIS-HCl, 0.1 m*M* EDTA (pH 8.0) and stored at 4°C. The GenomeWalker adaptor was prepared by adding equimolar amounts of AdaptorF (5′GTAATACGACTCACTATAGG GCACGCGTGGTCGACGGCCCGGGCAGGT) and AdaptorR (5′PO_4-ACCTGCCC-NH_2), then incubating at 37°C for 10 min.

Primer design

Nested PCR primers were designed to 60 ESTs (Table 1) from a range of *Pinus* species using PRIMER3 software (Rozen and Skaletsky 1998). To assist with primer design, we analysed ESTs for homology to existing sequences in the public databases 'nr' and 'dbest' using the NCBI Blastn and Blastx programmes (Altschul et al. 1990). Where possible, PCR primers were designed within the 5′ or 3′ UTR region (or near to the start or stop codon within coding DNA) so that amplification would occur towards the non-coding DNA, either 5′ or 3′ of the EST (Fig. 1). As a positive control, 12 of the 60 EST primers were designed downstream from a known *Alu*I, *Dra*I, *Eco*RV or *Ssp*I site in either *P. radiata* or *P. taeda* to test for amplification of the expected product size. The nested adaptor primers were AP1 (5′GTAATACGACTCACTATAGGGC) and AP2 (5′ACTATAGGGCACGCGTGGT) (Siebert et al. 1995).

Twelve of the ESTs that were selected (Table 1) had previously been mapped in *P. taeda* and *P. radiata* pedigrees as PtIFG-RFLP probes (Devey et al. 1994, 1996, 1999; Groover et al. 1994). PCR primers had also been designed to 16 of the ESTs (including seven PtIFG clones) (Table 1) in previous studies, and the ESTs screened for polymorphism using either restriction enzyme analysis, denaturing gradient gel electrophoresis (DGGE) or SSCP gels (Harry et al. 1998; B. Temesgen and G. Brown, unpublished data; Plomion et al. 1999).

Amplification of genomic DNA fragments using an EST-specific primer and an adaptor primer

PCRs were performed using an EST-specific primer (ESTSP1) (Table 1) and an adaptor primer (AP1). PCR amplifications were carried out in a total volume of 25 µl using 5 µl of digested and ligated template DNA, 0.4 µ*M* of each primer, 1.25 U *Taq* polymerase (Roche), 1.5 m*M* $MgCl_2$, 0.2 m*M* each of dATP, dCTP, dGTP, dTTP, 50 m*M* KCl, and 10 m*M* TRIS-HCl (pH 8.3). Samples were overlayed with paraffin oil and amplified in 96-well plates on a Techne PHC-3 thermal cycler for 45 cycles with the following cycle profile: 30 s at 94°C, a 30 s annealing step, 60 s at 72°C. The annealing temperature in the first cycle was 65°C; this was reduced by 1°C in each of the next 4 cycles, and was continued at 60°C for the remaining 40 cycles. PCR amplifications were followed by a 15-min incubation at 72°C. Separate PCRs were performed using template DNA from each of the four genomic digests.

After the first amplification, PCR products were diluted 200-fold with distilled water, and a 5-µl aliquot was used as a template for the second amplification reaction. PCRs were performed in a total volume of 25 µl using a nested EST-specific primer (ESTSP2) (Table 1) and a nested adaptor primer AP2 as described above. Amplifications were performed for 30 cycles with the following cycle profile: 30 s at 94°C, a 30 s annealing step, 60 s at 72°C. The annealing temperature in the first cycle was 65°C; this was reduced by 1.0°C for the next four cycles, and was continued at 60°C for 25 cycles. PCR amplifications were followed by a 5 min incubation at 95°C, then slowly cooled to 42°C using a ramp rate of approximately 2°C per minute to promote the formation of heteroduplexes between heterozygous alleles.

Aliquots of 5 µl of PCR product from both the first and nested PCR amplifications were electrophoresed on 2% agarose gels using 1 × TBE as an electrophoresis buffer and then stained with ethidium bromide (0.25 ng/µl) to determine the approximate size of the products relative to a size standard.

Detection of polymorphic EST markers

Five detection methods were compared for their ability to resolve polymorphisms in the EST-specific fragments. Firstly, unlabelled PCR products were electrophoresed on 20-cm 6% non-denaturing acrylamide gels and stained with ethidium bromide (Sambrook et al. 1989). To evaluate four semi-automated gel detection methods, we 5′ end-labelled the nested adaptor primer (AP2) with either 6-FAM, HEX (Life Technologies) or NED (PE Biosystems), and PCR products were electrophoresed in either a denatured or a non-denatured form on an ABI Prism™ 377 DNA sequencer (ABI 377) using 36-cm well-to-read gels. Firstly, PCR products were electrophoresed denatured through 5% Long Ranger acrylamide gels (FMC BioProducts) on an ABI 377 using standard denaturing electrophoretic conditions. To detect single nucleotide substitutions and heteroduplexes, we then electrophoresed the PCR products through three different non-denaturing gels on an ABI 377 including; standard 6% non-denaturing acrylamide gels (Sambrook et al. 1989); a high concentration (low cross-linked) (14%T, 0.15% C) acrylamide gel (Hauser et al. 1998); and 1 × MDE™ gels (FM C BioProducts). The non-denatured samples (0.7 µl) were added to 1 µl of Ficoll loading dye (17% Ficoll, 17 m*M* EDTA, 10.5 mg/ml Dextran Blue) and 0.25 µl of GS-ROX 2500

Table 1 EST loci, primary and nested EST primers and heterozygosity information for the *P. radiata* and *P. taeda* parents of three mapping pedigrees

EST locus[a]	GenBank acc. no.	Heterozygosity in parents[b,c]				ESTSP 1 & 2[e]	Primer sequence: 5′ to 3′
		pr	*qtl*	*base*	Enzyme[d]		
Pr1793′	AA220868	FI*	N	N	A, D, S	F1	GGCTCTCTGTTATGGCGATTCTCAGAT
						F2	TAGAAATATAGGCACAGAATGTGGCCG
Pr4CL	–	N	FI	FI	S	R1	TTCTCTGGAACAGGTACTGCAGGAAAG
						R2	AAGGGGCTGTGAAAAATGAACTCATCT
PrCAD	U62394	FI*	Y	FI	D, E, S	F1	GTCCGTTACAGATTTGTGGTGGATGTT
						F2	GCTTTTGTGAGGTTGAAACAATTCAGA
PrCHS1	U90341	FI*	Y	FI	A, D, S	F1	CCAAGTTATGCGGGCTGTTAGTATCT
						F2	CAGTCGTAAATTTGCTTGAGCAGGTCT
PrFLL	U92008	N	Y	Y	S, D	F1	AGATAATGGTTGGGACATTGAGGGTGT
						F2	CGTCAGCTCTGTCACTTGGAGAAAAGT
PrCEL1	U76725	Y	N	N	D, S	R1	ATCACAAAATGCCCTCATTGTCAAAAG
						R2	CTCATTGTCAAAAGGGTAGACCCACAA
PrCEL2	U76756	FI*	N	N	D	R1	CGTTGCGGTGCCTATCAATAGTAAGAC
						R2	TTGCTCAAATGACAGTAAAACTCACAAA
PrE29	AF049068	FI	Y	Y	A, D, S	R1	TAGCCAATACCAATGAAATTGCACACC
						R2	TGCACACCTAATAAGTGGGATGTTTGA
PrE74	AF049067	FI	FI	FI	A, D, S	R1	ATGGCCTTTCAAACAGAGAAACCAAAG
						R2	AGAGAAACCAAAGCGACAAGGAAACAG
PrE79	AF049066	Y	FI	FI	D, S	R1	GATAACTAAAACAGTTGCAGGCCGTGA
						R2	AGAACAAGAAATGAAAAGCAGCGAAGG
PrE87	AF049069	Y	N	N	S	R1	GTTCATCGAGGTCAGGATCACTCTCAT
						R2	TCGGTTTCATAGCTTCTGAAATCATGG
PrGER1	AF049065	Y*	N	FI	E, S	R1	GGCCTTCCGTTCATGAAAATAGATGAC
						R2	TTACCATCAGAGAAACAAGGAGGATGC
PrMADS2	U42400	Y	N	N	D	R1	GCCCCATTCTTTCCTTCTTTCTTTCTT
						R2	CCTCTCTCGGCCTCTTCCTTATTCTTC
PrMADS3	U76726	N	FI	Y	D	F1	ATACATGCAAGGATGGTGGGTTTGATA
						F2	ACTTCAATCACAACAAAAGCCCAAAGC
PrMADS4	U90345	FI	Y	FI	A, D, S	R1	CCCACAGACTCCAAATACAGTATGCTTC
						R2	TGCTTCTACAATCATGCTCCAAGTCTC
PrMADS5	U90346	FI	FI	Y	A, D	R1	TGCCCTCAAAACAATCAATCAGAGACT
						R2	TAACCACTTCTCGAACCCAGAATTTGA
PrMADS6	U90347	FI	FI	FI	A, D, S	R1	CGCTCTCGATCTTTCTCATTTGTGTCT
						R2	CTCTCCAAGCAGGCAAATCAGAGACTA
PrMADS9	U90344	N	N	N	D, S	R1	TTCTCGATCCTTTTCATCTGGGTCTTT
						R2	TCTGGCAAATCAAGATCGAATTGAAAC
PrMC2	U90343	FI	Y	FI	A, S	R1	GACGTTCCCGTTTGATACCTGAATTTT
						R2	TTGATACCTGAATTTTGGCCATTCTCA
PrUAE	–	Y*	FI	FI	D, S	F1	TGTAATGGACGGGAAGGGGATAATAGA
						F2	AAAGAGCATGGCAATGGTATTGAACAG
PtIFG_0602J[f,g]	H75115	Y	N	Y	D	R1	AGGCTCGTTACCTCCAGTCTCTCCTTA
						R2	CGGCGTACAGATATTCTCTCCTCTTCC
PtIFG_0669S[f,g]	H75159	FI*	N	Y	D, S	F1	AGCTCTTCGTAGGAAGGTTGGCTCAG
						F2	CCTTCTCCATCATCGGTGAAATCATAA
PtIFG_0670E[f,g]	H75160	Y	N	N	D, E, S	R1	ACACCGAAACCGAAGAGTTAAGCTGAT
						R2	CACCCAACAGTTCTATGCAATCCTCAA
PtIFG_1917J[f,g]	H75124	N	Y	Y	D, S	F1	TGGGTATCTCAAAAGCGCACAATAATT
						F2	CGCACAATAATTCTTCTTAACCAGGCG
PtIFG_2006S[f]	H75042	FI	FI	FI	A, D, S	F1	CATGTTCTCCATGTTCGGCTTCTTC
						F2	AGTTGAAAATCTCTACGACCACTTGGC
PtIFG_2020E[f]	H75044	N	FI	N	A, S	R1	TCATAATAATGCACATTGACTCCAGCG
						R2	GAATTTTTCTTCATCTCTGCCCCAAAC
PtIFG_2442E[f]	H75198	N	Y	N	D	F1	TGCTCCTACGTTGGATGATTTTCTCAC
						F2	GGGCACCAGCAGCATAAATAAATTTCT
PtIFG_2564E[f]	H75212	Y	Y	Y	A, D, E	F1	CCCTTCAACATGTCATCGCTGATAATAA
						F2	TAATCGAGGTGTTGCTGCTGGGTTAC
PtIFG_2610E[f,g]	H75220	FI*	Y	Y	A, D, S	F1	TCACAACTGATGATGGAGGACATAGGA
						F2	TGATGGAGGACATAGGAGTCTCGCTAC
PtIFG_2723E[f,g]	–	N	Y	N	A, D	R1	ATTCTTCCTCTGCAACTCCAGGAAAAC
						R2	AACTCCCATTCCTGAAGGTGCAAGT
PtIFG_2986E[f,g]	–	FI*	Y	Y	D,S	R1	ATCTGCCCTTCTTGAGATACCTTCCAA
						R2	CAATTCTTCCCACTGGAAACTGAAGAC
PtIFG_3026E[f]	–	N	FI	FI	D	F1	CCTACCGACAAGGAAGAATTTGGAGAC
						F2	CATAGATCCAGACTTGAAGGACATGGG
PtLTP	U10432	FI	FI	FI	A, D	R1	TAAGGTATGTAGCGCATGGAGTCATGG
						R2	TCATCATTGTAACTACAAGCCCAACCA

Table 1 (continued)

EST locus[a]	GenBank acc. no.	Heterozygosity in parents[b,c]				ESTSP 1 & 2[e]	Primer sequence: 5′ to 3′
		pr	*qtl*	*base*	Enzyme[d]		
Pt1C5G	AA556423	FI	Y	FI	D, S	R1	CTCCATGAACTCCACCATCTCTTCGTA
						R2	CACAAACACACTAAAATGGCAGAACCA
Pt4C7C	AA556497	N	Y	Y	A, D	F1	GGACCTCAAGAATGGCAAAGTCACTAA
						F2	TTGTGGACAATGATTTTCTAGCTGTGG
Pt2C6G	AA556633	FI	N	N	D	R1	AGAAATGTAGACCTGCTCGGAGGAATC
						R2	AAAAATCCTCTGCAATCTCTCCATGGTT
Pt2NA12G	AA556728	Y	Y	Y	D	F1	TATTTGGGGCAATGGTTCTTTCTTTTC
						F2	ACCAGCGAATTCATTCTTTTGGGATAA
Pt6C3A	AA556828	Y	Y	Y	D, S	R1	CAAGAAGGAATGTGGGATGACCAAG
						R2	GTTCGCCACTGAAAAATCAGGTAGACA
Pt3C1C	AA556531	Y*	N	FI	D, S	F1	AGAGAAGAGCGTTCTGGATGAGAGACC
						F2	ACTGTTCATGCGACTGTAGACGGTTTC
Pt3C1G	AA556534	N	Y	Y	D	F1	GGGAATACAAAGAGAAGAGCGTTCTGG
						F2	CTGCCGTAATGAGACTGGCTGTAAGAT
Pt2NA3E	AA556733	N	FI	FI	S	F1	AGGAAATAACAAGGGAACGGTGTTCTG
						F2	CAGAAGCTGCTCTCCTCCAGAAAATAA
PtCOMT-1[h]	U39301	FI*	FI	FI	A, D, E, S	R1	CCCCTGCAATTCCTCTTCCTCGCTCTC
						R2	GCCAGGCCGTTCATGTTCGAATCCATC
PtAGP[h]	U09554	Y*	FI	FI	A, D, S	F1	AGATTCATTATTGGCCCATCAGACAGA
						F2	CCCACTATTCGCAATAAACATGCATTC
Pt4CL2[h]	U12013 U39404	N	FI	Y	D, E, S	R1	GGCATGGCGATCGAAATACTGAAAATA
						R2	AAAACGTGTCTGACAAATGTAGCCTGC
PtDDCBER1	AF081678	Y*	N	Y	D, S	F1	GACGAGTACCTGAGCAATTTTGTGTGA
						F2	AAGCCAGGCTAGGTAGATCCTTCACAC
PtICL8[h]	U39807	FI*	Y	FI	A, E, S	F1	TGCTGGCCAAGTCTAGAATGTAAAACA
						F2	TATACGTGGGGATATTCCATCAGACCA
PtLP3–1[h]	U52865	FI*	FI	FI	D, S	R1	GTTTCTCCTCCGACATCTTCTCCAAAG
						R2	ACCCGTGATGATATGGATAGAGACGTG
PtLP3	U67135	FI	N	N	A, S	F1	CCTCCTCCCCTTCTTGTGTTTGTTATC
						F2	GTTGCGTATCTCACATCCGTTCTTGAT
PtLP6[h]	U31309	N	N	Y	D	F1	TCAGTTCCGGTACATCTCAGCAAATTC
						F2	CGCTGTAATGACATACAATGTTTCGATG
PtLP15	AF013803	Y*	Y	FI	A, E	F1	CCACAGAAGGCATCACTGGTTTGTAAT
						F2	TTATAGGTCAGCGGTGGCTTCAAATTC
PtLP20	AF013804	Y	FI	FI	A	F1	GAGGTCCATTATCCAATCGAAAACCTG
						F2	TAAAACATTAAGGCATCGATTGGGTCA
PtLPPAL[h]	U39792	Y*	Y	N	A, D	F1	CCAGAAATATGTAAGGGTGGCAGATCA
						F2	CTGAAACATGTAAAAGTGGCGGATCAT
PbRAMS	U38186	Y	Y	Y	A, D, S	F1	CTGCTGGTAGTTATGTATGCTTTGAATTG
						F2	TGCTTTGAATTGTAATGGCTAAGTTGAT
PcGER1[h]	AF039201	FI	N	N	A, S	F1	AAGAGCAATGGATCAGAATTCCCATTA
						F2	TTGGCATTATTCAAAATGGTGTTTCAA
PmPINMIII	AF038949	N	FI	Y	D	F1	ATCGTCATTGTCGTTTATGGGCTGTAG
						F2	GTTGCGTGATTATGTGTAGTTCCGTGA
PsCHI4	U57410	Y	FI	FI	A, S	F1	TCCAACCTCGACTACAAAAACCAGAAG
						F2	CCAACCCATCCAATCGACTACTACTGT
PsyCHS[h]	X60754	FI*	Y	FI	A, D, E, S	R1	GATAACTGCTCTGGTCAACCGCATTAG
						R2	TAAATCCTTCATCATTCCTGCAGCCAT
PsyNG3PD	L32561	FI	FI	FI	A, D, E, S	F1	AACACCCAATGTCTCAGTTGTGGATCT
						F2	AGGCATCCTTGGATACACTGATGAAGA
PsyGPD	L26923	FI	FI	FI	A, D, E, S	F1	TGAAAGTGGTTGCTTGGTATGACAATG
						F2	GGGGTTATTCACAGAGGGTTGTTGATT
PthCAB	X13407	FI	FI	FI	D, E, S	R1	CTTACGGTTCTTCGCATGGTGATTC
						R2	CCACTTTCTTGACGAGCTCATTCTGCT

[a] The EST loci names contain a two- or three-letter prefix relating to the species of origin, followed by a locus identifier. The prefixes: Pr, *P. radiata*; Pt, *P. taeda*; Pb, *P. banksiana*; Pc, *P. caribaea*; Pm, *P. monticola*; Ps, *P. strobus*; Psy, *P. sylvestris*; Pth, *P. thunbergii*; [b] Y, Heterozygous in at least one parent; FI, fully informative with three or four alleles between the parents; N, not polymorphic; [c] *pr*, parents of a *P. radiata* mapping pedigree (850.055 × 850.096); *qtl* and *base*, parents for each of two *P. taeda* mapping pedigrees (Devey et al. 1994; Groover et al. 1994); [d] Restriction enzymes generating allelic polymorphism: A, *Alu*I; D, *Dra*I; E, *Eco*RV; S, *Ssp*I ; [e] Primers with the suffix F1 or F2 amplify fragments 3′ of the EST, while primers with the suffix R1 or R2 amplify fragments 5′ of the EST. F1 and R1 primers were used in the first round PCR; F2 and R2 primers are nested within the product of the first round of PCR and were used in the second-round PCR; [f] PtIFG-RFLP probes that had previously been mapped in *P. taeda* and *P. radiata* pedigrees (Devey et al. 1994, 1996, 1999; Groover et al. 1994); [g] ESTs screened for polymorphisms in the *qtl* and *base P. taeda* mapping pedigrees using either restriction enzyme analysis or denaturing gradient gel electrophoresis (DGGE) (Harry et al. 1998; B. Temesgen and G. Brown, unpublished data); [h] ESTs screened for polymorphisms in a *Pinus pinaster* and a *P. sylvestris* mapping pedigree using SSCP analysis (Plomion et al. 1999); *Indicates that the EST locus was genetically mapped in this pedigree using the described above

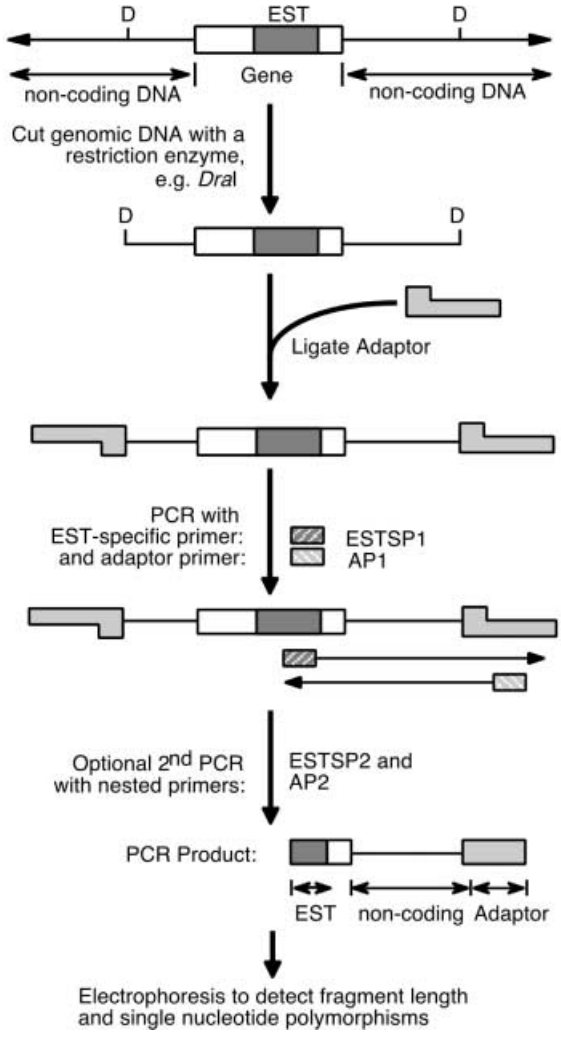

Fig. 1 Schematic representation of PCR-based EST polymorphism detection. Genomic DNA is cut with one of four blunt-end cutting restriction enzymes, then a GenomeWalker (Clontech) or a vectorette adaptor (Riley et al. 1990; Siebert et al. 1995) is ligated onto the fragments. An EST-specific primer is used in combination with an adaptor primer to amplify non-coding DNA either 5′or 3′ of the EST. For complex genomes, a second round of PCR using nested primers is required to obtain specific products

size standard (PE Biosystems) and then electrophoresed at 50 W, and a temperature of 37°C for 6 h. The electrophoresis data were analysed using GeneScan Analysis® and Genotyper® software (PE Biosystems).

Genetic mapping of polymorphic EST markers in *P. radiata*

To evaluate the inheritance pattern of polymorphic EST markers in *P. radiata*, we screened ten megagametophytes from each parent and 8 full-sib progeny with 10 EST primers . In addition, 18 polymorphic EST markers were amplified in 45 full-sib progeny, and the markers were genetically mapped onto existing RAPD, SSR and AFLP-based genetic linkage maps (P.L. Wilcox, unpublished data) using Mapmaker Macintosh V2.0 (Lander et al. 1987). The criteria for accepting linkage between loci were LOD ≥ 3, θ ≤ 0.4.

Identifying allelic variation in *P. radiata*

For 5 polymorphic EST markers the parental alleles from two *P. radiata* parents were excised from a 6% non-denaturing acrylamide gel, added to 100 µl of elution buffer (0.5 *M* ammonium acetate, 1 m*M* EDTA, pH 8.0), and incubated at 37°C overnight. The eluted DNA was then diluted 200-fold, and a 5 µl aliquot was re-amplified using a nested EST-specific primer (ESTSP2) and a nested adaptor primer (AP2) using the reaction conditions described above. PCRs were performed for 40 cycles with the following cycle profile: 30 s at 94°C, 30 s at 60°C and 60 s at 72°C. The PCR products were purified using a QIAquick PCR Purification Kit (Qiagen) and were either sequenced directly using the ESTSP2 and the AP2 primer, or the purified products were cloned into a pGEM-T vector (Promega), electroporated into electro-competent cells (DH10B) using a BioRad Gene Pulser and sequenced using an M13 universal primer (Sambrook et al. 1989).

Results

The procedure used for detecting EST locus polymorphisms is outlined in Fig. 1. The protocol features digestion of genomic DNA, ligation of a GenomeWalker (Clontech) or a vectorette-like adaptor (Riley et al. 1990; Siebert et al. 1995) and amplification using 1 EST-specific primer (ESTSP) and 1 adaptor primer (AP). Where possible, PCR primers were designed within the 5′ or 3′ UTR region (or near to the start or stop codon within coding DNA) so that amplification would occur towards a restriction enzyme site in the genomic DNA, either 5′ or 3′ of the EST.

All 60 EST primers amplified fragments in *P. radiata* and *P. taeda*

PCR primers were designed to 60 ESTs obtained from a range of *Pinus* species that included 20 ESTs from *P. radiata*, 32 ESTs from *P. taeda*, 3 ESTs from *P. sylvestris*, and 1 EST each from *P. banksiana*, *P. caribaea*, *P. monticola*, *P. strobus*, and *P. thunbergii* (Table 1). Restriction digests of genomic DNA from parents of a *P. radiata* mapping pedigree (850.055 and 850.096), and from four parents from the *qtl* and *base P. taeda* mapping pedigrees (Devey et al. 1994; Groover et al. 1994) were ligated to a GenomeWalker adaptor (Clontech) and screened for PCR amplification products using a single EST-specific primer (ESTSP1) and an adaptor primer (AP1) (Fig. 1). Four restriction enzymes, which varied in their expected cut-site frequency, were chosen for the genomic digests to increase the likelihood of amplifying a subset of fragments that ranged in size from 100 bp to 1 kb. The enzymes selected included a frequent 4-bp cutter (*Alu*I, cut site: AG↓CT), two AT-rich 6-bp cutters (*Dra*I, TTT↓AAA and *Ssp*I, AAT↓ATT) and a less frequent 6-bp cutter (*Eco*RV, GAT↓ATC).

Following amplification, PCR products were electrophoresed on 1.5% agarose gels and stained with ethidium bromide. Initially, all of the primer and enzyme combinations tested produced a smear of fragments in both *P. radiata* and *P. taeda*. It was assumed that the

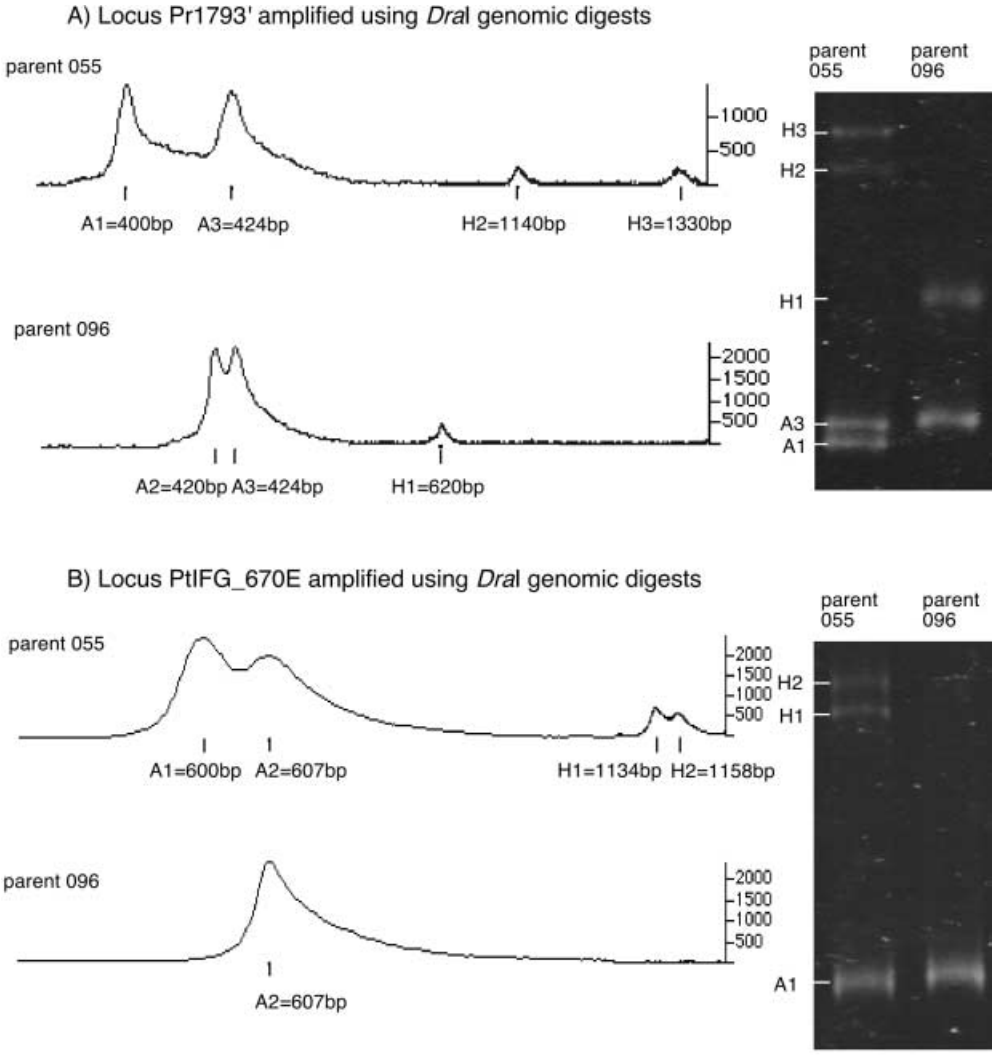

Fig. 2A, B GENOTYPER (PE Biosystems) electropherogram profiles and ethidium bromide stained 6% non-denaturing acrylamide gels for loci Pr1793′ (**A**) and PtIFG_0670 E (**B**) from *P. radiata* parents 850.055 and 850.096. Genomic DNA was digested with *Dra*I and amplified as described by two rounds of PCR using EST-specific primers 1793F1 and 1793F2, and PtIFG_0670ER1 and PtIFG_0670ER2, respectively. (Table 1). Amplified alleles are labelled *A1* to *A3*, and corresponding heteroduplexes are labelled *H1* to *H3*, respectively. Heteroduplexes were not detected using denaturing gel conditions and were not present in amplifications using haploid megagametophyte DNA (data not shown)

smears were due to non-specific amplification, so a second round of PCR amplification was performed using a nested EST-specific primer (ESTSP2) and a nested adaptor primer (AP2) to reduce the number of amplicons. Following the second amplification, of the 60 EST primers tested, 56 amplified single products with at least one of the four digests in both *P. radiata* and *P. taeda*. The remaining 4 EST primers amplified products from two or more loci in either *P. radiata* or *P. taeda*. The fragments ranged in size from 80 bp to 1.5 kb, and for many of the EST loci the sizes of the fragments seen on agarose gels were comparable in both species with all four restriction enzymes. As positive controls, 12 primers were designed within known restriction enzyme fragments, and in all cases a product of the expected size was seen in either *P. radiata* or *P. taeda*.

In both *P. radiata* and *P. taeda*, we found that the amplification frequency and fragment size was related to the restriction enzyme used for the genomic digests. In both species, approximately 50 (83%) of the EST primers amplified fragments from genomic digests with each of the two AT-rich 6-bp cutters, *Dra*I and *Ssp*I, and the median size of the fragments was 360 bp and 400 bp, respectively. In contrast, 35 (58%) of the EST primers amplified fragments with genomic digests from the 4-bp cutter, *Alu*I, and the median size of the fragments was 270 bp. Only 22 (37%) of the EST primers amplified fragments with the *Eco*RV genomic digests, and the median size of the fragments was 510 bp.

To test the reproducibility of this technique, we performed restriction digests and PCR amplifications for ten EST primer pairs in triplicate for two *P. radiata* trees. All 10 EST primers amplified identical products in each replicate when analysed using 6% non-denaturing acrylamide gels stained with ethidium bromide.

Detection of length polymorphisms allowed 40% of the ESTs to be mapped in *P. radiata*

Genomic digests from the two *P. radiata* parents were screened for fragment length polymorphisms at 60 EST loci using an ABI 377. Fluorescently labelled PCR products were electrophoresed through denaturing 5% Long

Ranger acrylamide gels (FMC BioProducts) and analysed with GENESCAN software (PE Biosystems). Fragment length polymorphisms were detected in at least one parent for 24 (40%) of the ESTs tested, and in both parents for 5 ESTs. These length polymorphisms ranged in size from 1 to 30 bp and enabled us to genetically map the ESTs in this pedigree.

Detection of both SNPs and fragment length polymorphisms allowed over 70% of the ESTs to be mapped in *P. radiata*

To increase the likelihood of detecting polymorphic EST loci, we evaluated four non-denaturing acrylamide gel systems for their ability to detect both fragment length and nucleotide substitution polymorphisms in *P. radiata*. The formation of heteroduplexes, was promoted by first denaturing the PCR products and then slowly cooling them to 42°C prior to electrophoresis on all four non-denaturing gel systems.

When the PCR products were electrophoresed on 6% (20 cm) non-denaturing acrylamide gels and subsequently stained with ethidium bromide, polymorphisms were detected in at least one parent for 42 (70%) of the ESTs, and in both parents for 20 (33%) ESTs (e.g. Fig. 2). Thus, by using standard, non-denaturing acrylamide gels and heteroduplex analysis, the overall level of polymorphism increased by 30%, and the number of fully informative markers (those that segregate four genotypes in the progeny) increased four-fold. The additional polymorphisms were detected as either homo- or heteroduplexes, or both, and resulted from nucleotide substitutions that affected the mobility of the fragments in a non-denaturing gel (see sequence data below). In most cases, the heteroduplexes migrated at a slower rate, and the polymorphisms were resolved more clearly, than the corresponding homoduplexes (Fig. 2).

To evaluate a more automated gel detection method, we electrophoresed fluorescently-labelled PCR products through 6% non-denaturing acrylamide gels run on an ABI 377. Using this detection method, all of the polymorphisms that were initially resolved on 6% non-denaturing acrylamide gels stained with ethidium bromide (including both the homo and heteroduplexes) could also be detected on an ABI 377 (Fig. 2). In addition, for 9 of the ESTs tested, additional polymorphisms under 10 bp in apparent size were also detected in at least one of the parents (Fig. 2). As a result, the number of mappable EST markers increased to 45 (75%), and the number of fully informative markers to 26 (43% of the ESTs tested) (Table 1). Interestingly, in each pedigree approximately 90% of the polymorphic EST loci were detected using genomic digests from the restriction enzymes *Dra*I and *Ssp*I.

Two alternative high-resolution acrylamide gel systems were compared for their ability to resolve polymorphisms on an ABI 377. One was a high concentration, low cross-linked, non-denaturing acrylamide gel (Hauser et al.

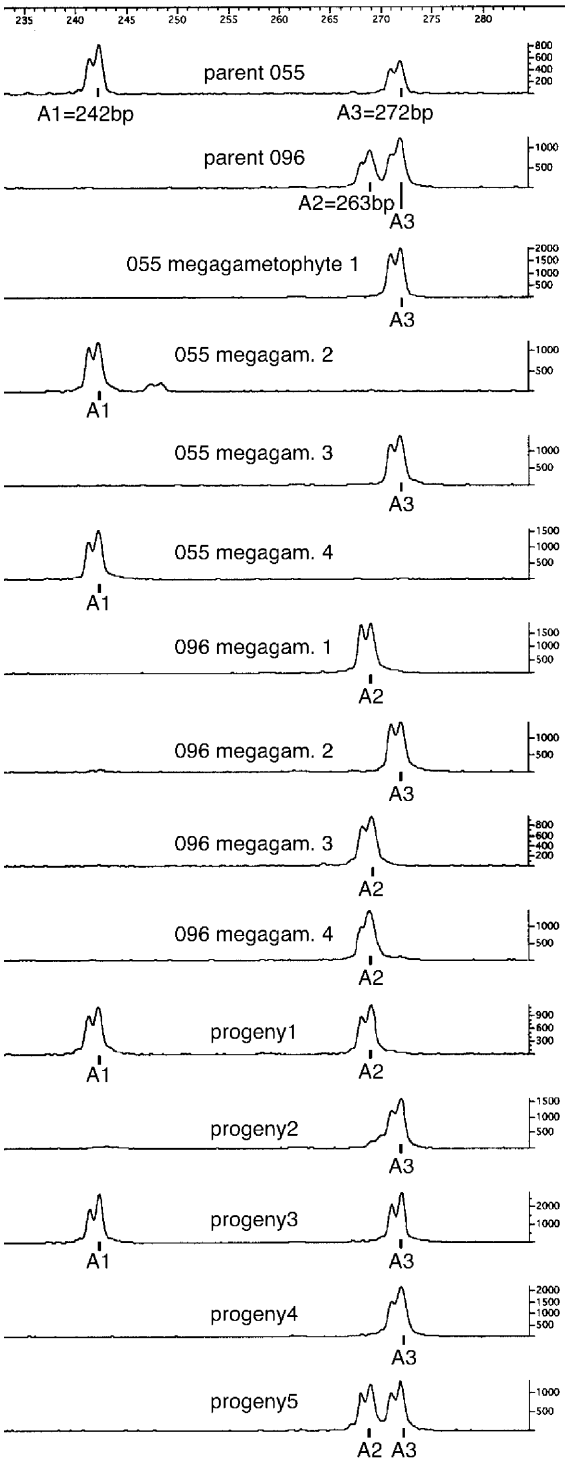

Fig. 3 GENOTYPER (PE Biosystems) electropherogram profiles for locus Pr1793′ from *P. radiata* parents 850.055 and 850.096, four haploid megagametophytes from each parent and 5 diploid progeny. Genomic DNA was digested with *Alu*I and amplified as described above using EST-specific primers 1793F1 and 1793F2 (Table 1). Amplified alleles ranged in size from 242 bp to 272 bp and are labelled *A1, A2* and *A3*. Only one allele was amplified from each megagametophyte. The 5 progeny segregated as either homozygotes (progeny 2 and 4) or heterozygotes (progeny 1, 3, and 5) for the locus. The profiles shown were generated using denaturing gel conditions, but comparable profiles were also seen using 6% non-denaturing gels on an ABI 377 (see Fig. 2)

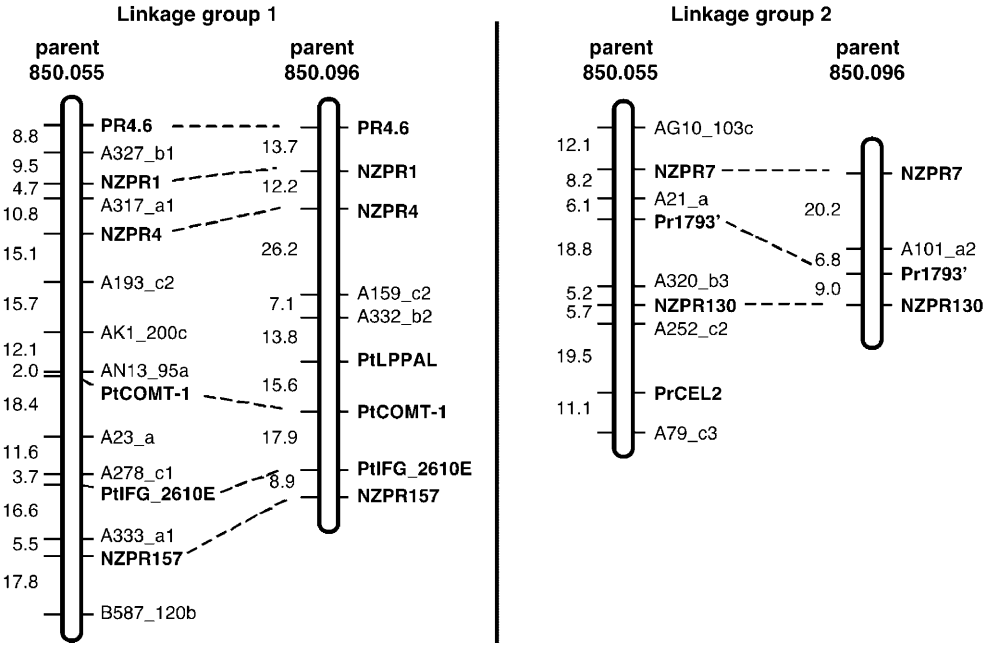

Fig. 4 Example of alignment between *Pinus radiata* parents 850.055 and 850.096 of sex-specific linkage groups using EST and SSR markers (*bold print*). RAPD and AFLP markers that were previously mapped in this pedigree (P.L. Wilcox, unpublished data) are also shown. Map distances to the *left* of each linkage group are expressed in Kosambi centiMorgans

1998), and the other was a 1 × MDE™ gel (FMC BioProducts). Although the electropherogram profiles varied slightly depending on the gel system used, no new polymorphisms were revealed by either of the two high-resolution gel systems on an ABI 377.

Equivalent levels of polymorphism were detected in *P. taeda*

To assess the level of polymorphism in *P. taeda*, we screened the four parents from the *qtl* and *base P. taeda* mapping pedigrees (Devey et al. 1994; Groover et al. 1994) for polymorphisms with all 60 EST primers using 6% non-denaturing gels on an ABI 377. Of the 60 EST loci tested, 44 (73%) were heterozygous in at least one parent of the *qtl* pedigree, while 22 (37%) were heterozygous in both parents (Table 1). Similarly, for the *base* pedigree, 46 (77%) loci were heterozygous in at least one parent, while 27 (45%) were heterozygous in both (Table 1).

Genetic Mapping of EST markers

To investigate the segregation and inheritance pattern of polymorphic EST markers in *P. radiata*, we screened ten megagametophytes, and 8 full-sib progeny from each parent with 10 EST primers. All 10 primers amplified products that were heterozygous in one or both of the parents, were present in the haploid megagametophytes and inherited in the full-sib progeny in a Mendelian manner (Fig. 3). Eighteen EST loci (including the 10 ESTs described above) were genotyped in 45 progeny of the *P. radiata* mapping pedigree and placed onto existing RAPD, SSR and AFLP-based genome linkage maps (Fisher et al. 1998; P.L. Wilcox and M. Yoon, unpublished data). Of the 18 EST loci 10 were heterozygous in both parents and provided alignment between homologous linkage groups derived from each parent (Fig. 4).

The sequences of the polymorphic alleles revealed abundant length and nucleotide polymorphisms

To confirm that the targeted region of the genome had been amplified and to determine the nature of the polymorphisms seen, we excised the parental alleles from 5 polymorphic EST loci (Pr1793′, PrCAD, PrCHS1, PtIFG_2610E, and PtCOMT-1) from non-denaturing acrylamide gels and sequenced them. Sequence analysis revealed that each set of nested primer pairs amplified products that contained an ESTSP2 primer sequence (in the 3′ or 5′UTR), the 3′ or 5′ end of the original EST sequence, a region of unknown (presumably untranslated or non-coding) DNA, the remains of a restriction enzyme site and the end of the GenomeWalker adaptor (Clontech) sequence (Fig. 5). Comparison of allelic sequences showed that for 3 of the 5 EST loci (Pr1793′, PrCHS1, PtIFG_2610E), the length polymorphisms seen on both the denaturing and non-denaturing acrylamide gels were due to insertions or deletions in the non-coding DNA adjacent to the EST. For all 3 ESTs, the parental alleles also contained between 1 and 16 single nucleotide substitutions per 250 bp of sequence (ignoring the insertions and deletions) (Fig. 5). These substitutions did not change the fragment length but may have contributed to the mobility differences seen between the homo- and heteroduplexes on the non-denaturing acrylamide gels. For the other 2 EST loci that were polymorphic on the non-denaturing acrylamide gels but not on the denatur-

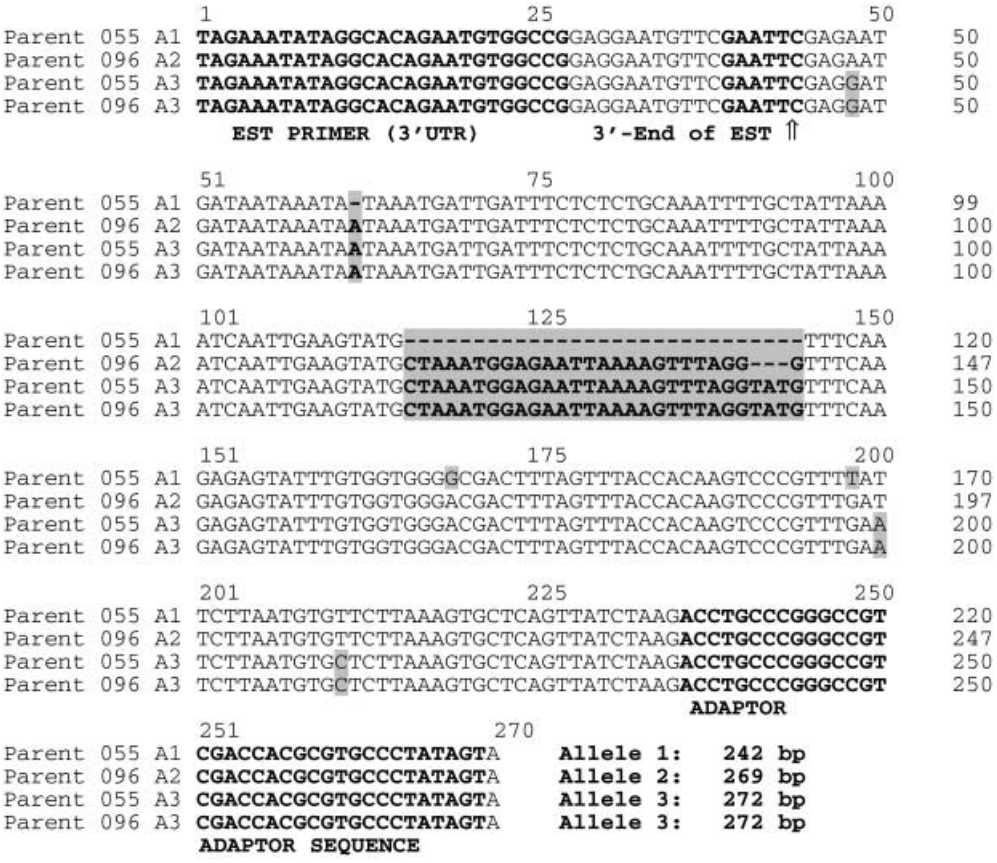

Fig. 5 The parental alleles (*A1*, *A2* and *A3*) shown in Figure 3 were sequenced. Sequences are shown in the 5′ to 3′ direction. The regions corresponding to the EST-specific primer 1793′F2 and the adaptor sequence, at either end of the fragment, are shown in *bold*. Insertions and deletions in each of the four fragments are shown in *bold* and are *highlighted grey*. Five additional SNPs are shaded in *grey*

ing gels (PrCAD and PtCOMT-1), we found that the parental alleles contained between two and five single base pair substitutions with no change in overall fragment length (data not shown).

Discussion

We have developed a high-throughput method for genetically mapping ESTs based on fragment length polymorphisms and SNPs in coding and non-coding sequences adjacent to ESTs. The method is compatible with standard ABI 377 DNA analysis systems and can be used with a range of denaturing and non-denaturing gel systems.

PCR primers were designed to 60 ESTs from a range of *Pinus* species that included *P. radiata*, *P. taeda*, *P. sylvestris*, *P. banksiana*, *P. caribaea*, *P. monticola*, *P. strobus* and *P. thunbergii*. In many cases the nested primers were designed within the 5′ and 3′ untranslated regions, yet in all cases products were amplified in both *P. radiata* and *P. taeda*. This suggests that a high level of sequence conservation exists within the transcribed DNA among *Pinus* species. The fragments ranged in size from 80 bp to 1.5 kb, and for many of the ESTs the position of the restriction enzyme sites outside of the EST were conserved between *P. radiata* and *P. taeda*. These results suggest that orthologous, rather than paralogous genes were amplified in both species. This is an important consideration if EST markers are to be used for aligning genome linkage maps across related species.

In both *P. radiata* and *P. taeda* we found that the amplification frequency and fragment size was related to the restriction enzyme used for the genomic digests. Approximately 83% of the EST primers amplified fragments from genomic digests with each of the two AT-rich 6-bp cutters, *Dra*I and *Ssp*I, and the median size of the fragments was 360 bp and 400 bp, respectively. In contrast, only 58% of the EST primers amplified fragments from genomic digests with the more frequent cutting enzyme, *Alu*I, and the fragments were smaller, with a median size of 270 bp. In most cases, the decrease in amplification frequency with the enzyme *Alu*I could be attributed to the presence of a cut site between the 2 EST primers and the consequent lack of a template for the nested EST primer in the second PCR. Only 37% of the EST primers amplified fragments with genomic digests from the rarer cutting enzyme *Eco*RV and, as expected, the fragments were larger with a median size of 510 bp. The lower amplification frequency with the *Eco*RV digests may have been due to the restriction sites being too distant from the EST priming sites for reliable amplification using *Taq* polymerase under the PCR conditions employed.

The requirement for a second nested PCR to amplify single-locus products may have been due to the large and complex genome of *Pinus* species (1C = 21–31 pg, ap-

prox. 2.0–3.0 × 10^{10} bp) (Wakamiya et al. 1993). In most cases, amplifications with a single EST and adaptor primer were sufficient to amplify single-locus products when full-length cDNA from *P. radiata* was used as a template instead of genomic DNA (data not shown). Therefore, it is likely that a second PCR would not be required in organisms with less complex genomes, providing that the EST does not belong to a large multi-gene family.

The number of detectable polymorphisms was affected by the gel and detection system used. When PCR products were electrophoresed on denaturing 5% Long Ranger acrylamide gels, using an ABI 377, fragment length polymorphisms were detected in 40% of the EST loci in two *P. radiata* parents. However, when PCR products were analysed using non-denaturing conditions that favoured SNP and heteroduplex analysis, the overall level of polymorphism increased to approximately 75%. This increase was presumably due to base pair changes that altered the mobility of the homo- and/or heteroduplexes on non-denaturing gels. This hypothesis is supported by the DNA sequences of 2 EST markers that were polymorphic on a non-denaturing acrylamide gel but not on a denaturing gel. In this example, the parental alleles contained between 2 and 5 bp substitutions in AT- and GC-rich regions. Such regions have been shown to affect DNA curvature and electrophoretic mobility on non-denaturing acrylamide gels (Bolshoy et al. 1991; Goodsell and Dickerson 1994).

We found there were several advantages in using the laser-based fluorescent detection method over conventional ethidium bromide staining. A larger number of samples could be assayed per gel (up to 288 samples could be screened per gel if three colours and 96 lanes were used on an ABI 377). PCR products ranging in size from 80 bp to 1.5 kb could be effectively resolved on the same gel, and allele scoring and analysis was semi-automated using existing GENESCAN and GENOTYPER software (PE Biosystems). In addition, because a single labelled universal adaptor primer is used for every PCR, the costs associated with using fluorescently labelled primers is minimised.

One *P. radiata* and two *P. taeda* pedigrees were screened for polymorphisms with all 60 ESTs. We found that the level of polymorphism in each pedigree was comparable, with an average of 45 polymorphic ESTs per pedigree (i.e. 75% of the ESTs were polymorphic), and of these an average of 25 (42%) were fully informative in both parents. We found that approximately 90% of the polymorphisms could be detected using genomic digests from the AT-rich 6-bp cutting restriction enzymes, *Dra*I and *Ssp*I. Thus, a similar level of polymorphism could be detected in the future by screening ESTs with genomic digests from these two enzymes alone, which would effectively halve the time and costs associated with screening EST primers. In total, 59 of the 60 ESTs were heterozygous and mappable in at least one of the three pine pedigrees tested (Table 1). These results suggest that in *Pinus* species, insertions, deletions and base pair substitutions are frequent in the DNA immediately adjacent to expressed genes. These results are supported by the sequences of 5 polymorphic markers, from two *P. radiata* parents, which revealed between 1 bp and 16 bp substitutions per 250 bp of sequence (ignoring insertions and deletions).

The observed inheritance pattern of 18 polymorphic EST loci in a *P. radiata* pedigree revealed that the polymorphisms were due to allelic differences within a single locus and were not due to independent amplifications of different members of a gene family. These 18 EST loci were genetically mapped onto existing linkage maps for each *P. radiata* parent. Ten of the loci were fully informative and provided common loci for alignment of the sex-specific linkage groups (e.g. Fig. 4). ESTs derived from related *Pinus* species could be genetically mapped using this technique and this will enable a more rapid transfer of genomic information from one *Pinus* species to another.

We are currently using this technique to construct genome linkage maps of expressed *Pinus* genes. These EST maps will enable workers to align existing linkage maps in *P. radiata* and other pine species and will allow ESTs to be more efficiently identified that are associated with traits of interest in a range of species. Cytogenetic studies have shown that *Pinus* has been cytologically very stable over evolutionary time (Pederick 1970). Futhermore, genetic linkage studies using isozymes in several pine species (Conkle 1981), and more recent investigations using orthologous RFLP and EST markers for genetic mapping in *P. taeda*, *P. radiata*, and *P. elliotii* (Brown et al. 2000; Devey et al. 1999) suggest that the genome organisation within *Pinus* species is highly conserved. Once a high-density EST linkage map is constructed in a single *Pinus* species it may well serve as a reference gene map for all pines.

In summary, we have developed a high-throughput method for genetically mapping ESTs that reveals abundant fragment length polymorphisms and SNPs in genomic DNA adjacent to ESTs. The method is compatible with ABI 377 electrophoresis systems but can be used effectively in a range of non-denaturing gel systems. The procedure developed here should be generally applicable to any species for which EST or genomic sequences are available. It can also be used where sequence information is available only from a related species.

Acknowledgments We thank Amanda Warr, Minsoo Yoon and Andrea Langman from ***Forest Research*** for assistance with the *P. radiata* and *P. taeda* polymorphism screens and Matthew Lord and Joanne Dobson from ***SignaGen***™ for running the ABI 377 gels. We also thank Craig Echt, Philip Wilcox and David Harry for comments on previous versions of this manuscript. We thank David Neale, Garth Brown and Daniel Bassoni from the Institute of Forest Genetics for supplying the DNA from the *P. taeda* mapping pedigrees, and the sequences of PtIFG-RFLP probes. This work was funded by the New Zealand Foundation for Research, Science and Technology.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–10

Bolshoy A, McNamara P, Harrington RE, Trifonov EN (1991) Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. Proc Natl Acad Sci USA 88: 2312–2316

Brown GR, Kadel WE, Bassoni DL, Keihne KL, van Buijtenen JP, Neale, DB (2000) Orthologous markers for comparative mapping in the genus *Pinus*. http://dendrome.ucdavis.edu/Synteny

Conkle MT (1981) Isozyme variation and linkage in six conifer species. In: Conkle MT (ed) Isozymes of North American forest trees and forest insects. USDA Forest Ser Gen Tech Rep PSW-48: 11–17

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

Devey ME, Bell JC, Smith DN, Neale DB, Moran GF (1996) A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers. Theor Appl Genet 92: 673–679

Devey ME, 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

Fischer SG, Lerman LS (1983) DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. Proc Natl Acad Sci USA 80: 1579–1583

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

Goodsell DS, Dickerson RE (1994) Bending and curvature calculations in B-DNA. Nucleic Acids Res 22: 5497–5503

Groover A, Devey M, Fiddler T, Lee J, Megraw R, Mitchell-Olds T, Sherman B, Vujcic S, Williams C, Neale D (1994) Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics 138: 1293–1300

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

Hauser M-T, Adhami F, Dorner M, Fuchs E, Glössl J (1998) Generation of co-dominant PCR-based markers by duplex analysis on high resolution gels. Plant J 16: 117–125

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4: 403–410

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181

Messing J, Llaca V (1998) Importance of anchor genomes for any plant genome project. Proc Natl Acad Sci USA 95: 2017–2020

Orita M, Suzuki Y, Sekiya T, Hayashi K (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5: 874–879

Pederick LA (1970) Chromosome relationships between *Pinus* species. Silvae Genet 19: 171–180

Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce-Jacino M (1999) Mining SNPs from EST databases. Genome Res 9: 167–174

Plomion C, Hurme P, Frigerio J-M, Ridolfi M, Pot D, Pionneau C, Avila C, Gallardo F, David H, Neutelings G, Campbell M, Canovas FM, Savolainen O, Bodénès C, Kremer A (1999) Developing SSCP markers in two *Pinus* species. Mol Breed 5: 21–31

Riley J, Butler R, Ogilvie D, Finniear R, Jenner D, Powell S, Anand R, Smith JC, Markham AF (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res 18: 2887–2890

Rozen S, Skaletsky HJ (1998) PRIMER3. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html

Saier MH (1998) Genome sequencing and informatics: new tools for biochemical discoveries. Plant Physiol 117: 1129–1133

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

Schafer AJ, Hawkins JR (1998) DNA variation and the future of human genetics. Nat Biotechnol 16: 33–39

Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA (1995) An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res 23: 1087–1088

Wakamiya I, Newton RJ, Johnston JS, Price HJ (1993) Genome size and environmental factors in the genus *Pinus*. Am J Bot 80: 1235–1241