Theor Appl Genet (2001) 102:664–675

B. Temesgen · G.R. Brown · D.E. Harry · C.S. Kinlaw
M.M. Sewell · D.B. Neale

Genetic mapping of expressed sequence tag polymorphism (ESTP) markers in loblolly pine (*Pinus taeda* L.)

Received: 10 April 2000 / Accepted: 13 July 2000

Abstract The development and mapping of genetic markers based upon expressed sequence tag polymorphisms (ESTPs) in loblolly pine (*Pinus taeda* L.) are reported. The new markers were generated by PCR-amplification of loblolly pine genomic DNAs with primers designed from sequenced cDNAs. The cDNA libraries were constructed from RNAs expressed in the needles of loblolly pine seedlings or in the xylem from young trees. DNA polymorphisms were identified by analyzing the amplified products for differences in fragment size or restriction sites, or by examining mobility differences using denaturing gradient gel electrophoresis (DGGE). DGGE revealed more DNA polymorphisms than the other two methods. Fifty six ESTPs were mapped using either of two mapping populations and positioned onto a loblolly pine consensus genetic map. Unlike many other markers commonly used in forestry, ESTPs can be used as orthologous markers for comparative mapping, to map genes of known function, or to identify candidate genes affecting important traits in loblolly pine.

Keywords Expressed sequence tag (EST) · Denaturing gradient gel electrophoresis (DGGE) · Genetic marker · Genetic linkage map · Loblolly pine (*Pinus taeda* L.)

Communicated by P.M.A. Tigerstedt

B. Temesgen · G.R. Brown · D.E. Harry · C.S. Kinlaw
M.M. Sewell · D.B. Neale (✉)
Institute of Forest Genetics, Pacific Southwest Research Station,
USDA Forest Service, c/o Department of Environmental
Horticulture, University of California, Davis, CA 95616, USA
e-mail: dneale@dendrome.ucdavis.edu
Tel.: +1-530-754-8431, Fax: +1-530-754-9366

Present addresses: B. Temesgen,
California Department of Justice, DNA Laboratory,
626 Bancroft Way, Berkeley, CA 94710, USA
D.E. Harry, Nicholas Turkey Breeding Farms,
19449 Riverside Drive, P.O. Box Y, Sonoma, CA 95476, USA

M.M. Sewell, Oak Ridge National Laboratory,
Environmental Sciences Division, P.O. Box 2088, Oak Ridge,
TN37831-2008, USA

Introduction

Genetic linkage maps are useful tools for exploring plant genomes, and they have been constructed for various tree species, including conifers (Cervera et al. 2000). DNA markers used to construct such maps in forestry include random amplified polymorphic DNAs (RAPDs) (Nelson et al. 1994; Yazdani et al. 1995), amplified fragment length polymorphisms (AFLPs) (Remington et al. 1999), and simple sequence repeats (SSRs) (Devey et al. 1996). In large and complex genomes, like those of pines (Kinlaw and Gerttula 1993; Kinlaw and Neale 1997), such markers typically target non-coding (e.g. repetitive) regions and cannot directly pinpoint genes controlling important traits (Neale 1998). Even restriction fragment length polymorphisms (RFLPs) detected using cDNA probes in *Pinus radiata* and *Pinus taeda* are thought to reveal polymorphisms in regions flanking genes rather than in the coding regions of the genes themselves. In fact, none of the DNA markers hitherto used in forestry combine all of the following desirable attributes: the direct targeting of expressed genes (e.g., cDNAs), technical simplicity, high reproducibility among laboratories and across genetic backgrounds, codominance, and multiallelism.

We began developing PCR-based markers with these features from expressed sequence tags (ESTs) of loblolly pine (Harry et al. 1998). Many of the EST primer pairs amplified genomic DNAs isolated from loblolly pine, and the resulting PCR products revealed Mendelian polymorphisms after digestion with various restriction enzymes, a technique known as PCR-RFLP (Tragoonrung et al. 1992). PCR-RFLPs are technically simple, but they are time consuming and inefficient for large-scale projects, as revealing polymorphisms may require screening with many restriction enzymes. In addition, PCR-RFLPs typically reflect the presence or absence of specific restriction sites, so that any single restriction enzyme usually reveals only two alleles (Harry et al. 1998).

EST polymorphisms (ESTPs) can also be revealed, without the additional manipulation of PCR products, by

gel-based techniques such as denaturing gradient gel electrophoresis (DGGE; Fischer and Lerman 1983; Myers et al. 1987). DGGE reveals differences in the mobility of DNA fragments caused by partial melting of duplexes in the presence of increasing concentrations of the denaturant (Sheffield et al. 1992). Partial melting occurs in the region of lowest melting temperature (the lowest melting domain) and inhibits further migration; where sequence differences between alleles fall within the lowest melting domain, and alter melting characteristics, alleles can be resolved by DGGE (Sheffield et al. 1992). Since the lowest melting domain of a DNA fragment may span several hundred base pairs, the power of DGGE over that of the PCR-RFLP method to scan larger regions for polymorphism, as well as to detect multiple alleles of a locus, is clear (Temesgen et al. 2000).

We report the development and genetic mapping of 56 expressed sequence tag polymorphisms (ESTPs) in loblolly pine. ESTPs were generated by PCR-amplification of samples mainly from two loblolly pine pedigrees (Devey et al. 1991; Groover et al. 1994) with primers designed from cDNA sequences. DNA polymorphisms were identified by examining fragment lengths, restriction sites, and electrophoretic patterns detected using DGGE. Some of the cDNA clones were previously used as probes for RFLP mapping, and in such cases the map positions of the ESTPs were compared with their corresponding RFLP loci. This study demonstrates the feasibility of generating a relatively large number of ESTPs for genetic mapping in loblolly pine. This strategy may help to map identified genes and provide anchor loci for comparative maps. Once associated with quantitative trait loci (QTLs), ESTPs may also help to identify candidate genes controlling important traits in loblolly pine and other conifers.

Materials and methods

cDNA clones, sequence analysis, and primer design

ESTPs were developed from cDNA clones selected from two different libraries. The first library was constructed from RNA expressed in needles from loblolly pine seedlings and was derived using random priming (Devey et al. 1991). A total of 85 cDNAs were selected from this library for sequencing based on the criteria described by Harry et al. (1998), including RFLP banding patterns on Southern blots of loblolly pine genomic DNA. Nucleotide sequences were initially determined by manual methods, but later an ABI 377 automated sequencing apparatus was used. The second cDNA library was constructed from RNA expressed in the xylem of young loblolly pine trees and was synthesized by oligo-d(T) priming and directional cloning. Sequences from 75 xylem cDNAs included in this study were obtained by automated sequencing, and each contained a 3′ untranslated region (3′ UTR). Unlike cDNAs from the seedling tissues, those from the xylem tissues were not selected for low or single copies using the Southern hybridization data. The nucleotide sequences of all clones were compared against known gene sequences contained in GenBank using BLASTN and FASTA (see Table 1).

PCR primers had similar physical properties and were designed as described by Harry et al. (1998). Initially, most sequence polymorphisms were detected using restriction digests, so primers were designed to produce longer PCR products (>400 bp), preferably containing introns. Later, as DGGE was emphasized, the strategy was changed to generate shorter PCR fragments, preferably lacking introns. To do this, a reverse primer was first selected near the poly-(A) region, and then a corresponding forward primer was located about 200–300 bp upstream from it.

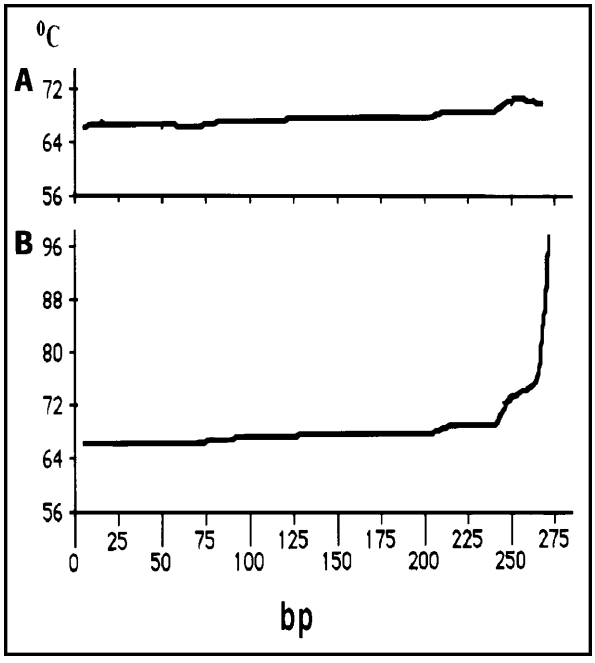

Fig. 1A, B The melting profile of EST0674. **A** The melting profile for the portion of PtIFG0674 cDNA sequence (272 bp), graphed using 50% probability. **B** Same as **A** except that a 39-bp GC clamp was attached to the 3′ end of the truncated sequence

To increase the efficiency of detecting DNA polymorphisms using DGGE, predicted melting profiles of the DNA fragments were first analyzed using the computer program MacMelt (Bio-Rad, Hercules, Calif., USA). After truncating a fragment to encompass only the sequence delimited by a given pair of primers, the MacMelt program was used to plot melting temperatures along the fragment's length. Melting profiles encompassing at least two regions with melting temperatures differing by >15°C were included. In many instances, MacMelt revealed acceptable melting profiles with no primer modifications (see Table 1). For some, however, the predicted profiles showed similar melting properties along the entire length of the fragment. In such cases, the melting temperature at one end of the fragment was increased by adding a GC-clamp to one of the primers (see Fig. 1).

PCR-amplification and detection of DNA polymorphisms

PCR reactions were performed as previously described (Harry et al. 1998), except that a Perkin-Elmer 9600 DNA Thermal Cycler (Norwalk, Conn.), or a PTC-100, MJ Research, Inc. (Watertown, Ma.) was used. Amplified DNA fragments were screened for DNA polymorphisms by three different methods: (1) differences in product length, (2) PCR-RFLPs, or (3) DGGE. To detect different fragment sizes or PCR-RFLPs, 4–6 µl of the PCR products were analyzed in 2% agarose gels (Harry et al. 1998). DGGE was performed using a D GENE apparatus (Bio-Rad, Hercules, Calif., USA) (Temesgen et al. 2000), except that perpendicular DGGEs were eliminated. ESTs were screened for polymorphism between the parents of two mapping populations on a 15–45% parallel denaturing gradient gel. In several cases, gradients were adjusted for better resolution of alleles. Putative polymorphisms were subsequently confirmed by analyzing six F_2 progeny before mapping the loci, using up to 95 individuals per mapping population.

Table 1 PCR primers derived from cDNA clones of loblolly pine

Clone name[a]	Primer pairs (forward and reverse, 5′→3′)	Product size		Population, detection method[b]	LG/ cM[c]	Best possible protein similarity[d]	Score[d]	E value[d]
	cDNAs from seedling needle tissues[e]	cDNA	Genomic					
0048[f]	GAC ATT TGG ACT CTT CCT GGC AAC ATG AGA ATC ACG GGC AC	270	270	Q, DGGE	2/72	Aquaporin	107	1e-26
0066[f, g]	gcc cgc ccc cgc cgc ccg gcc cgc gcc ccg cgc cgc- ccg cTG GCT GTT GGA ACC CAA G GGT ACC AGG GTA GGG TTT GC	385	385	B, DGGE	2/88.3	P65 protein Ectomelia virus	32	0.91
0107	GCA GGA CCT TCT GGA CAA TC AGG TGG AGA AAG CCA AGC TC	450	450	B, DGGE	1/112.3	ATP synthase B′ chain	83	2e-95
0136[f]	TGG CCA TCT ATT TCT GGT GG CAA TTG TCC CAT CCC CAT C	400	800	NP[h]	–	Protein kinase-calciumdependent	290	6e-78
0149	GGT GAC GCA ATC ACC TTT G CCA ATC CAA CCG ATA CTT GG	410	550	Q, RE (*Dde*I)	21/0	Apospory associated protein	96	3e–20
0459	TTA AAC GCT GCT GGT CAG G TGG AAA GAG CTG GAA GAT GG	450	550	Complex[h]	–	Ubiquitin	74	2e-13
0464	TGT CAC TGC CCA GAG CTA TTC ATC ACA GCC GCT CCA AAA C	450	730	B, LV	2/58	Aquaporin	65	7e-11
0500	GGC GAG TTG GCT TTC ATT C CAG CGA GGT ACC AGA TTT GC	400	500	B, DGGE	9/73.8	Plastocyanin	61	1e-12
0606[f]	TCA TGA GGG GAG AAA GTT GG GGG CAG TCA CTT GAA CTT CG	360	550	Q, DGGE	6/98.6	RepJ	32	3.9
0624	CAC AAT TGC CAG ATG GGT C CTT CTC TAG CAA CGA TCC GG	620	940	B, RE (*Hin*P1 I)	9/55.2	Protein kinase	120	2e-27
0669[f]	GTA TTC CTG CTT TCG TTC GG CGA GTC CCT GAT AGA CCC TG	370	600	NP[h]	–	Photosystem II 22-kDa protein precursor/psbS gene	94	7e-19
0670	gcc cgc cgc gcc ccg cgc cgg ccc ggc gcc ccc ggg- cgg CAG AAG ACC AGA GCC AGG AC GAC AAA CGC CAA GGG AAT C	400	450	NP[h]	–	Pectin esterase 3	49	6e-06
0674[f]	ATA ATT GGG GAG ATG CGG AG gca ggc ggc gcg ggg cgc ggg ccg ggc ggc ggg ggc- ggA GCA GAA GGA AGA CAG AGT GG	270	270	Q, DGGE	14/0	Cyclophilin	99	2e-20
0708[f]	TGT TGA CGA TCA AGC TGC C TCA AGC CAT CAG CGA AGT AC	310	350	Smear/ complex[h]	–	Translationally controlled tumor protein	202	2e-51
0739	TAC AAA GGC TCG GGT ATT GG GTC ATG CCT AAC AAG CCC TG	350	350	Q, DGGE	6/67.6	Phospholipase D	29	4.0
0893	GGA CTG AAG GGA TCT AGC TGG CAG CCC AAA TTC CAT CGT C	450	620	Q, DGGE	5/41.3	Nonspecific lipid transfer protein	157	2e-38
0967	AAG GGG AGG AGC AAG TGA AG CTT GTT CAG GAA GTC TGC CC	350	350	NP[h]	–	Glutathione S-transferase	36	0.039
0975	GAC GAA GAT GCG AAG AAA GG ATA CGG AAG CGT ACC TGG G	400	400	NP[h]	–	Ribosomal B2 protein	80	4e-15
1165	CTG AGG GGA TGT GCA TGA C AAA CCA GCA GCC ACA TGA G	250	380	B, DGGE	6/50.8	Metallothionein-like protein	55	8e-08
1454	ACA TCA ATC AAG TTG GCC TTG ACG ACC ATC TCC AAC CAC TC	350	350	B, DGGE	5/75.9	–	–	–
1576	TGG TAT GTG GAG GGA AGG C TAC AGC GTT TGC TCC TCC TG	480	480	Q, RE (*Dde*I)	1/59.7	Regulatory factor RFX1	33	0.55

Table 1 (continued)

Clone name[a]	Primer pairs (forward and reverse, 5′→3′)	Product size		Population, detection method[b]	LG/ cM[c]	Best possible protein similarity[d]	Score[d]	E value[d]
1584	CGA AGC AAA GGA TGT CAC G TGT TGA GGT GGG GAT TGG	340	340	NP[h]	–	Chalcone reductase	57	9e-11
1599	CAG GAT CAT ATG CTG AAG CG ACT CGC CAA TTT GCT CTA GC	520	520	NP[h]	–	ATP synthase delta chain chloroplast	40	0.001
1623	TTT CTC AGG TGG GAG AGG TG TCG CAT CCA TGT GCG TAG	350	950	P, DGGE[i]	–	NS1-associated protein	102	2e-21
1626	TAA GAA GGC GGC GGT ACA G AAA CCA GCA GCC ACA TGA G	550	550	B, RE (*Hinf*I)	6/51	Metallothionein-like protein	62	8e–10
1635–3′	AAG AAG ACA ACG AGC AAC GG GCC CAC TCG AAT CAC AAA A	650	870	Q, RE (*Hin*P1 I)	10/48.8	Ribulose bisphosphate carboxylase, small subunit	145	3e-36
1643	AAT GGA GGA TGC CGT TAC AG AAC CAC TCT CGA ATC CCC AC	490	650	B, DGGE	10/56	ABI1 gene product (Protein Phosphatase)	59	5e-09
1750	TGT TTA CGT TCT TGA CGC GG TAG CAA GCA CTC TGA CTG TGG	300	300	B, DGGE	16/24.7	Ferredoxin I	54	7e-17
1917	gcc cgc ccc cgc cgc ccg gcc cgc gcc ccg cgc cgc-ccg ATA TCC GTC GCC TGG TTA AG GAT TCT CAA AGC AGC CCA AG	320	450	Smear/ complex[h]	–	Chloroplast transit peptide	108	1e-23
1934	GAC GAA GTT GGT GGC GTA G TTC TGT TTG TGC GCC TAC TG	850	850	B, RE (*Bsa*J I)	3/2.5	Light harvesting complex chlorophyll a/b binding protein	97	5e-21
1950	AAA CCA GCA GCC ACA TGA G TAT TAA GAA GGC GGC GGT AC	350	450	Q, DGGE	6/54.7	Metallothionein-like protein	62	3e-09
1955	AGC CAA TGC ACC AAG AAG G ATC CAA CAA CAG AAC CCC TC	290	290	B, RE (*Bsa*J I)	5/82.3	Metallothionein-like protein	29	0.017
1956	GAA GCT AGC GAA GGC TTT GG GGG TGT GAC CAT ATA ACA CCG	410	500	Q, RE (*Alu* I)	9/7.1	ubiquitin protein ligase	36	3e-16
2009	CAC AGT TCC CCA CAG CAA C ACA AGC GGT TCA GTG GCT C	400	600	B, DGGE	6/107.7	60S ribosomal protein L10A	126	2e-29
2053[f]	TGA AAC TGC GGA TGG CTC CCG GAA TCG AAC CCT AAT TC	310	360	Q, DGGE	Not mapped[h]	Folate binding protein	85	2e-21
2111	AGT CTT GGC CTT TTC TTG GG CAC CAA ACC TAA CCG AGA CG	350	350	GU[h]	–	–	40	0.002
2166	CTG CTG TTG AGC TTG TGT ACG TGC CCG TGT AAA GAT GAC AG	400	400	Q, RE (*Hinf*I)	4/56.5	Pyruvate dehydrogenase	40	0.002
2253	CCA ATT TGC ACT TTG CCC CCA AAG CCC AAA TCC ATG	370	370	Q, DGGE	1/59.9	Fructose-bisphosphate aldolase 1, chloroplast	58	1e-08
2274	TGA TCA GAG AGC TGG TGC AG AGA TGA GCA TCA GGT CAG CC	400	600	B, DGGE	4/86	adenylyl cyclase	30	3.0
2290	AGC TTG CAG CAT CAA CCG GAA CCA AAC AGC TTC AGG ACC	550	840	Q, RE (*Hin*P1 I)	9/57.5	Light harvesting complex chlorophyll a/b binding protein	53	2e-07
2358[f]	GTT AAC CCT CGA GGA GAC ATG GCT TCC ACA GTC CAC AAT CTG	330	330	Q, DGGE	22/0	Phenylalanine tRNA synthetase	30	3.6
2393[f]	ACA CGT CTG TCA TCT CAT GGG ccc ggc cgc cgg cgg CCC TGA ACC-AGC ACA AGT ATC	280	350	S, DGGE[i]	–	BcDNA GH07269	34	0.99
2541[f]	GCT GCT GAC ATG AGC TTT TG AAG TCC CAC ACA TCG AGG TAC	300	410	NP[h]	–	6-Phosphpfructo-2-kinase/ fructose-2,6-bisphosphatase	160	2–39
2610	CTG AGG GGA TGT GCA TGA C AAA CCA GCA GCC ACA TGA G	350	350	B, DGGE	6/50.9	Metallothionein-like protein	40	1e-04

Table 1 (continued)

Clone name[a]	Primer pairs (forward and reverse, 5′→3′)	Product size		Population, detection method[b]	LG/ cM[c]	Best possible protein similarity[d]	Score[d]	E value[d]
2615[f]	CAC TCT TTA TTC TTG CCC TTC G TCG GTT AGG TAA CGA CTG GAC	320	400	Q, DGGE	11/49.5	P69 2–5A synthase I	33	1.6
2723	GCA CAC CAA AGC AGC AGC AAT GGG ACT TGC ACC TTC AG	550	850	P/A	–	Actin de-polymerizing factor	55	5e-11
2781[f]	GAT GAT GCC CTG AAG AGC C ATG GAA CCA AAG GAG ATG CC	450	450	B, DGGE	8/64.7	Glucose-induced repressor	43	0.002
2786[f]	CTG CCA AGA CAG AGA AAC TGT C CCA TTG TCT GTT GTT GCC TC	300	300	NP[h]	–	Transport protein	39	0.048
2889	ACG CCA GCT CTG ACT ACC AG GTT TCT TCT CGT GGT GCT CG	530	750	B, RE (*Rsa*I)	23/0	ABA/ ripening inducible protein	73	4e-13
2986	TCG CAG GTG AAG TTG TGA AG TCT AAA GGC CCA CCC TAA CTC	350	1000	Complex[h]	–	Histone H2A	40	0.005
	cDNAs from xylem tissue[e]							
8462	TGC CCC TAT TGT ACA CAG GTC GGA CCA TGC CAA CAT CAA AC	260	260	GU[h]	–	Glutaredoxin	168	3e-41
8510	TTC AAG GCC GTT CCA TTC CAT TAA CCC AAA GGC CCA G	240	240	Q, DGGE	11/51.1	RNA-binding glycine rich protein	81	5e-15′
8562	AAG ACG GCT GTG TGG TTC TG TAG CAG GAG ATG TTG GTT CCC	200	200	Complex[h]	–	Chitinase	144	4e-21
8564	CAC CAG GGC AAA AAG TTG G GCA GTT ATA GGT TTC CTG GCC	230	230	B, DGGE	6/55	Ribosomal protein 60S L2	147	4e-35
8565	ATT TGT GGC TGC GGA AAG CAC CAA GTA CAC CAC AAC ACC	200	220	Smear[h]	–	Glucose-regulated	160	9e-43
8569	TCG ACC AGT GTT GCA GAG G TCG TCT CCC TCA TCA CCT TC	210	210	Q, DGGE	2/24	Tubulin	242	1e-63
8598	CAC AGA TGG TGA CGG AGA AC GGT GAA CGC TTG TTG AAC G	190	190	Complex[h]	–	Coatomer delta subunit	62	8e-15
8613	ATT CAC CCT TCC CCT GAC C gcc ggg ccc gg cTT AGC CCA AAC AGG-GAA AGC	210	210	B, DGGE	3/67.6	Embryogenesis-associated protein late embryogenesis LE14-goshi	113	5e-25
8643	TGG TTT CCA CTG TTG CAG C GAG TTT GAA GCC CTT GAC TGG	260	260	GU[h]	–	Cysteine proteinase	98	3e-20
8650	ccc ggc cgc cgg ccc CTC TCA AGG AGC-CGA ACA AG ACC CTT GGA TTT TCC CGT C	220	220	S, DGGE[h]	–	RNA binding protein	127	4e-29
8702	GTT GCA GAA AAG GGT GGC AGT CGC ACT TGC TCC AGT TC	290	350	Q, LV/ DGGE	6/77.9	Thioredoxin	69	3e-11
8704	TAA GGC ACA TGG CAA GCT C CAG AAG AAT GAG CTG GCA CC	250	250	NP[h]	–	Pyruvate dehydrogenase	56	7e-16
8714	TGA ACT GCT GAA TCT TGG GC GAC GGC ATC TCA AGA ATA CCC	260	260	Smear[h]	–	Initiation factor 5A-2	76	7e-15
8721	ccc ggc cgc cgg ccc TGC CCA AGT GCA-ATC ATG CGC TTG CTC AGA TCA TGG	290	290	S, DGGE[i]	–	Oligosaccharyl transferase	87	77e-17

Table 1 (continued)

Clone name[a]	Primer pairs (forward and reverse, 5′→3′)	Product size		Population, detection method[b]	LG/ cM[c]	Best possible protein similarity[d]	Score[d]	E value[d]
8725	AGC GCT GAA TGA TGT CTT GG gcc ggg ccc ggc CCA AAC TTA CAC- CAT GCT CG	260	260	Q, DGGE	9/1.0	COP9 complex subunit 4	106	8e-23
8728	CCA AAG CCC AAA TCC ATG CCA ATT TGC ACT TTG CCC	370	370	Complex[h]	–	Aldolase plastid	60	3e-17
8744	TGC AGA TCC TAT CTT TGG GG CAA GGT CCT CGT ACC CAA AC	340	420	Complex[h]	–	Cell-wall glycosylatable polypeptide	67	7e-11
8758	GAG TTC GCA TCA CAC GTT ACC AAT TGC CCC ATT GAT TCT GG	350	350	Complex[h]	–	Protein kinase guanine nucleotide binding	89	1e-17
8777	ccc ggc cgc cgg AAG AGC TTT GCA AGC GTG AG AAC GTC CAA ATC TGT CCA GG	170	170	NP[h]	–	Ribosomal protein 60 S L18 A	199	1e-50
8790	TTG GGG AAG CAA GCA TTC CTT TGT TTC CAC CCT CCG AC	250	900	Complex[h]	–	Hypothetical ORF	96	9e-20
8796	TCA GAG TCA GCT AGG CGT TG TTC ATC CAA GGC GTT GAA C	400	400	Complex[h]	–	ATP synthase vacuolar	77	5e-14
8886	TTC CGG AAG GTG TGG TGG AGT CAC TCC CTG TCA CCG AC	310	310	B, LV	7/81.4	Kinase I nucleoside diphosphate spiol	78	3e-24
8887	TGG GGT TGG TGA GAT ACT GC CAT ATA TTG GGA AAA CGT TCG C	320	500	B, LV	4/89.1	RNA polymerase II subunit RPB10 homolog	73	1e-12
8898	GGG ATG GCA ACA ACA AAA AG ATG GGG GTG CAG CAT AAA C	330	1400	B, DGGE	4/120	Testis mitotic checkpoint	119	9e-27
8907	TCC TCA ACA GGC AGA CCT TC TGG GTG AGA CTT ACG AAC AGG	300	300	Q, DGGE	8/0	Peroxidase cationic	67	6e-11
8917	CCT TTC AGA GGC TTG TTA GGG ACA TTG CGT TTC TGC GTA GG	310	550	Smear[h]	–	Histone H3.2	149	9e-36
8939	ACG TGG ACG AGC AGT CAA AG AAC CAC GAG CTT GGC ATG	300	300	Q, DGGE	2/84.3	Ribosomal protein 40S S16	195	2e-19
8962	TCC TGA AGA GGA GGA GGA GG AAC CAT GCA ATT GGA GCC	350	350	Smear[h]	–	Kinetochore Skp1p homolog	69	2e-11
8972	TTG GTC CCC TTG TTG GAG GCC TCC ATT CGA CTC ACT TG	310	310	B, DGGE	6/18.3	Protein 2 major intrinsic plasma membrane	136	1e-31
8993	CCA TCT CCA CTG ATG CGT AC GAA GAG CGG CAT ATC CGA C	250	600	Complex[h]	–	Disease resistance response protein PEA	121	2e-27
9008	ggc gcc cGG TAA ACT GGG ATG GAT TGC TCT CGG ATA GGG CAA TAT GC	290	290	Q, DGGE	5/71.5	ATP-binding protein	242	2e-63
9022	CGG TGT GTT TCA TGT GCT G GGA TTT GCA TTT TGC ATG CC	210	210	B, DGGE	2/36.7	Translation factor SUI1/GOS2 protein	88	2e-17
9034	AGC CCG GTC CAC TTA TCA AG TGC ACA CAG AAT CAG AAT TGG	300	300	Q, DGGE	Not mapped[h]	CAEEL 19.6-kDs protein C23G10.2 in chromosome III	71	5e-12
9036	CAG GAC GAA TGA GAT ACC TGC GTC ATC CGA TAC AAC CTC AAT C	250	350	Q, DGGE	8/45.8	Ribosomal protein L37	61	5e-20
9044	cgc ggc ccA ACT GGA GGA AAA GCA CGA C CAT CGC ATC AGT CAT ACT CAC C	280	280	B, DGGE	6/53.1	Ribosomal protein 40 S S27	105	1e-22
9047	CTA AAG GCT TTC CTA TGC GC GTC ATT TTT CCA CTT ATC CCT CC	300	300	NP[h]	–	Cytochrome P450	71	6e-13
9050	gcg ccC AGC TCT GGA AGC TCT CTC C TTT TGA CAG TCT GAA ACG GC	250	250	Complex[h]	–	ABA and salt stress response 40 g	73	6e-13
9053	ggc ggT GCA TGA TGA CGG CTC TAT G CCA CCG AAA TAT ATG CCT GTC	280	280	Q, DGGE	1/38.1	Translationally controlled tumor homolog (TCTP)	108	2e-23

Table 1 (continued)

Clone name[a]	Primer pairs (forward and reverse, 5′→3′)	Product size		Population, detection method[b]	LG/ cM[c]	Best possible protein similarity[d]	Score[d]	E value[d]
9055	ATA GGG CGC ACA ACA GCT C AGG AGC CTT CTT CTT TCC TGG	300	1000	GU[h]	–	Ribosomal protein 30S S13	89	5e-31
9061	ggc gcT CGG GCT GTA ACG AAT ATG C CCT CCA TCA CAG ACC AAC AG	250	250	NP[h]	–	Fe(II) ascorbate oxidase	69	1e-11
9064	GAG GGG AAC ACT AAT CAA GCG TGC CTC AAA TGC TTG CTT TC	270	270	Smear[h]	–	Ribosomal protein L14	93	1e-18
9076	AGA ATT TAC TGG CCG CTC G CTC TAT TGC AAA AAT GTG CCA C	250	250	B, DGGE	11/39.1	Phophoglucomutase	98	3e-20
9088	GCC CTT TTG GGT GTG CGT AGG AG AGC CCT GAA CAT AAT TTC AGG	290	290	NP[h]	–	DNA-damage-repair/toleration protein DRT102	64	7e-10
9092	ggc cgg gTC ACT GAC CTT AAC GTC CC AGC TAA AGT TGG CTG GCA TC	270	400	B, DGGE	5/43	Non-specific lipid transfer protein	89	7e-18
9098	ggc ggg GTG GGC TTG CTA TAA ATG C AGT GCA TCG TTC ACA ATT CTC	280	280	NP[h]	–	Selenium-binding protein		
9099	ggc cgg cAG ACA CAT TAA GCT GGG GTC C TTT CAG GGG AGT TGC AAT AAG	300	300	GU[h]	–	Ribosomal protein S12	69	2e-11
9102	CCC AGA GAT CTT CCG CTA TG ggc cgc ggc AGA AAG GAG CAT TTC CCG AC	240	240	Q, DGGE	1/84.6	laccase	63	2e-10
9103	cgc ggg cTC TCA GGA GAC CAG TGA GAG C GCA TAG AGA GGA ACC TTT CCC	260	260	Smear[h]	–	Ribosomal protein 60 S L30	68	2e-11
9113	AGG AAA AGG TTC TCC AAG CG gcc ggc cgA CAG CTT AGG CAT TAC AGC CC	300	300	Q, DGGE	8/80.2	S-adenosylmethionine synthetase 2	95	2e-19
9123	ccg gGA CGT CTC TCT GCT ATC GTC G TTC CAA AGT TCC CAG GTC AG	280	280	NP[h]	–	Amino peptidase	105	1e-22
9132	cgg ccg CTC CAT CTG GTG GGT TGA AG AAT CTA TGC CGT GGA ACA GC	270	270	Complex[h]	–	Ribosomal protein L7	120	4e-29
9134	TAT GGT TGT GGA GAG GCT TG GAG GAA TCT TCT GCC ATT AGC	300	300	NP[h]	–	Profilin	107	4e-23
9136	gcc ggc cgg GCC CTT GTC CAA GAC TGT AAG CCG ATT TGA AAC ACA CAT CC	260	260	P, DGGE[i]	–	ribosomal protein S11	150	4e-36
9151	TAG TGA GCC CTG GAG CGT AC GCA GAA TCT CAG CAG CAA TG	290	290	Q, DGGE	7/20.1	Cucumber basic protein, a blue copper electron transport protein	91	3e-18
9155	gcg ccc ggT TGG TGT TAG GCA GTC ATG G TGA AAT TCC CAG CCC ATA TC	260	550	P, DGGE[i]	–	EST *Arabidopsis*	80	2e-14
9156	ggc ccg cgc ccg TAA GCT TCG TGC AAC AGG AG GAC AAT CCC TCT AAA CCT CGC	280	400	Q, DGGE	9/13	Ribosomal protein 40S S14	132	1e-30
9157	TTC CAG TTT CCC TGA GCA TC ccg cgg gcc gcg gcc AAT ACG CTG CTT-AAT CGT GTC	275	275	B, DGGE	4/8.2	26S protease regulatory subunit	51	5e-06

[a] For simplicity, only locus identity fields were reported for clone names

[b] Loblolly pine mapping populations used in this study and methods utilized to detect the DNA polymorphisms: B: *base*; Q: *qtl*; P/A: presence or absence of PCR products; LV: length variant; RE: PCR-RFLP; DGGE: denaturing gradient gel electrophoresis

[c] LG: linkage group; cM: centimorgan; these values were also presented in Fig. 2

[d] Putative identities determined from BLASTN AND FASTA search of GenBank; Score: the score values obtained; E value: expected values

[e] cDNA clones constructed from either seedling needle (clones #s <3000) or xylem tissues (clones #s >8000)

[f] Primers were designed from sequences of seedling cDNA clones that were obtained by using an ABI 377 automated sequencing apparatus

[g] GC-clamps were added to 5′ ends of one of the primers (GCs are distinguished as lower case; e.g., 0066)

[h] PCR products dropped from further analysis for various reasons: complex, smear, NP (Non-polymorphic), GU (genetics un-interpretable), not mapped (could not be mapped)

[i] Not polymorphic in the *base* and *qtl* mapping population but were polymorphic either in *prediction* (another loblolly pine population, unpublished) or in *slash pine* (Brown et al., in preparation)

Germplasm and genetic mapping

ESTP segregation data were obtained from either the *base* (Devey et al. 1994) or *qtl* (Groover et al. 1994) loblolly pine mapping populations. Polymorphisms in either population were located on a consensus genetic map using markers and methods described in Sewell et al. (1999). Primers that were monomorphic in both the *base* and *qtl* populations were subsequently tested for polymorphisms in two other populations: the *prediction* pedigree (also from loblolly pine, Sewell et al., in preparation), and a *P. elliotti* pedigree (Brown et al., in preparation). Mapping data from the latter populations will be presented elsewhere.

Nomenclature and database

The nomenclature for genetic loci and markers follows the outline on the Genome Resources page in Dendrome, the forest tree genome informatics Web site (http://dendrome.ucdavis.edu/Data/locusname.html). Although we previously named the PCR-based codominant markers as sequence-tagged sites (STSs; Harry et al. 1998), the name "ESTP" was introduced to emphasize that they are based on expressed (cDNA) sequences. For convenience, the full locus names of both the ESTP and RFLP loci were shortened. For example, in IFGBAS_PtIFG_2253_a, an RFLP locus on the loblolly pine map (Devey et al. 1994), both the Experiment (IFGBASE) and Source (PtIFG) fields were omitted and only the Accession Number and locus fields (2253_a) were reported (see Table 1).

Results

Amplification of genomic DNA

PCR primers were based on EST sequences from two distinct sources. Differences in library construction, primer placement, and methods to detect polymorphisms prevent direct comparisons between the seedling and xylem libraries, although several general observations were made.

In total, 58% (50 of 85) of primer pairs from the loblolly pine seedling library amplified genomic DNAs (Table 1). PCR products from genomic DNA ranged in size from 270 to 1000 bp, averaging 515 bp. Fifty six percent of these genomic fragments were measurably larger (55–735 bp) than those from their corresponding cDNAs, indicating the presence of one or more introns.

Among primers from the xylem library, 76% (57 of 75) amplified genomic DNAs. PCR products ranged in size from 170 to 1400 bp, averaging 344 bp (Table 1). Introns were rarely observed using these primer sets–only 12 primer pairs produced measurably larger genomic products than fragments obtained from the corresponding cDNAs. Primers located closer to the 3′-end of the cDNA sequence tended to amplify genomic DNA more reliably than did primers located further upstream. For example, when the forward (upstream) primer was within 350 bp of a reverse primer near the poly-(A) region, approximately 95% of the primer sets amplified genomic DNA. However, less than 60% of primer sets, selected from regions between around 200 bp from the poly-(A) region and 600 bp (upstream), amplified genomic DNAs (data not shown). We surmise that smaller DNA fragments and/or those near the poly-(A) regions amplify more reliably, possibly because of the decreased likelihood of introns in smaller fragments and near the 3′-end regions.

DNA polymorphisms

Parents of the *base* and *qtl* mapping pedigrees of loblolly pine were screened for ESTPs using three methods. All ESTs were tested for both length variation on 2% agarose gels and mobility differences using parallel DGGE, typically using a single 15–45% gradient gel and the appropriate acrylamide concentration (either 6 or 10%). Perpendicular DGGE, used previously to determine the optimal denaturant concentration for each EST (Temesgen et al. 2000), was omitted because the majority (>90%) of polymorphisms could be detected on a 15–45% parallel gradient. PCR-RFLP assays were performed on only a subset of seedling ESTs in addition to those reported earlier (Harry et al. 1998), because of the cost and time associated with screening PCR products with numerous restriction enzymes. Polymorphisms showing Mendelian segregation were shown for 36 (70%) seedling ESTs and 29 (50%) xylem ESTs (Table 1).

Some ESTs were polymorphic by more than one method and/or in both mapping populations**.** Confirming an earlier observation (Harry et al. 1998), only five product length polymorphisms were detected (EST464, EST1955, EST8702, EST8886 and EST8887) using 2% agarose gels. Note that EST8702 was converted to a DGGE marker (see below). Detection methods for each ESTP are shown in Table 1.

One major advantage of using xylem sequences for developing ESTs was that the majority of them were predictive of their corresponding complete genomic sequence (i.e., 80% of xylem primer pairs did not amplify introns). Using the MacMelt computer program, the melting properties of the PCR fragments could be visualized before primer synthesis. In this manner, primers could be placed to optimize melting profiles, with or without a GC clamp added to the 5′ end of the forward or reverse primer. In some cases, primers from seedling ESTs that did not amplify introns were re-analyzed as well. For example, EST0674 was not polymorphic by any method until its melting profile was assessed and a GC clamp incorporated (Fig. 1).

Previously, DGGE was shown to reveal more DNA polymorphisms than are revealed by PCR-RFLPs (Temesgen et al. 2000). It is also not surprising that DGGE reveals additional polymorphisms relative to other types of allelic variants. For example, two length variants (300 and 320 bp) were observed for EST8702 in the *base* and *qtl* populations (Fig. 2A). Subsequent analyses using DGGE revealed three alleles in the *base* population and four alleles in the *qtl* population (Fig. 2B). Among individuals in the *base* population, lanes 2 and 3 appear as identical heterozygotes based upon length variants (Fig. 2A), but a third allele is revealed using DGGE (Fig. 2B).

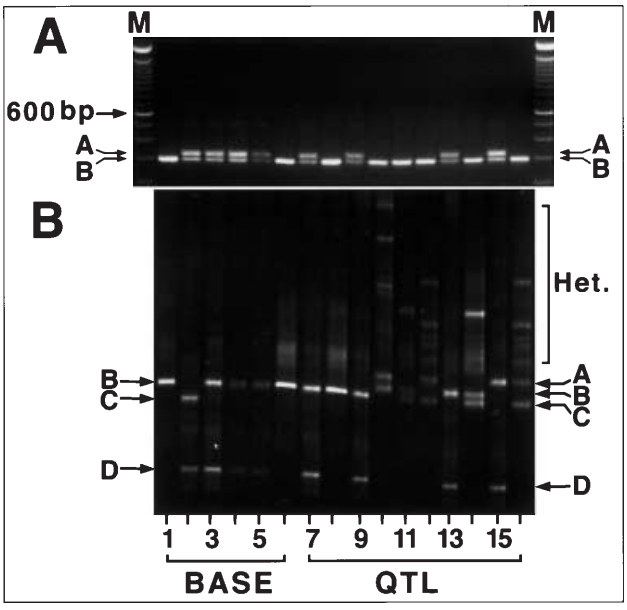

Fig. 2A, B Analysis of the EST8702 primer pair. Lanes 1–6 are grandparents and parents from the *base* map population (Devey et al. 1994), while lanes 7–12 are those from the *qtl* population (Groover et al. 1994): (*1*) and (*7*) maternal grandmother, (*2*) and (*8*) maternal grandfather, (*3*) and (*9*) seed parent, (*4*) and (*10*) pollen parent, (*5*) and (*11*) paternal grandmother, (*6*) and (*12*) paternal grandfather. Lane *M* is a 100-bp DNA ladder. F_2 progeny from the *qtl* mapping population are also shown (lanes 13–16). **A** PCR products after electrophoresis in 2% agarose gel and staining with ethidium bromide. The two alleles in both the *base* and the *qtl* are designated as *A*/*B*. DNA marker fragment sizes (*M*) are given in base pairs (bp). **B** Same as **A** except that the PCR fragments were separated using DGGE. Different alleles in each population are labeled *A*–*D*. Heteroduplexes (*Het*.) are indicated by brackets

Likewise among individuals in the *qtl* population, lanes 13 and 15 appear as identical heterozygotes in agarose gels (Fig. 2A) and, again, a third allele is revealed using DGGE (Fig. 2B). Similarly, among individuals that appear to be identical homozygotes in agarose gels (Figs. 2A, B, lane 14 vs 16), DGGE again reveals differences. In addition, DGGE reveals DNA fragments representing heteroduplex molecules whose mobility is retarded relative to homoduplexes (Fig. 2B, Temesgen et al. 2000). Heteroduplexes can be used to enhance both allelic classifications and genetic interpretations (Temesgen et al. 2000).

Several primer sets were dropped from subsequent genetic analyses for various reasons: eight primer sets yielded smeared amplification products, 12 primer sets produced complex patterns using DGGE, and five primer sets yielded patterns that could not be interpreted genetically.

Of the primer sets, 21% (23 of 108) yielded distinct products with single bands, but no DNA polymorphisms were detected in either the *base* or the *qtl* mapping populations. The search for DNA polymorphisms in these samples was broadened by including DNA samples from the parents of two other populations. Three additional polymorphisms were detected in the *prediction* population in loblolly pine, and another three were detected in the *P. elliotti* pedigree (Brown et al., in preparation) (Table 1).

Genetic mapping of ESTPs

Segregation data were obtained for 56 ESTPs from the *base* and *qtl* mapping populations of loblolly pine. Thirty two ESTPs were derived from seedling cDNAs (1 size variant, 11 PCR-RFLPs, and 20 DGGE) and 24 were derived from xylem cDNAs (2 size variants and 22 DGGE). Two ESTs (EST2053 and EST9034) could not be mapped. Some ESTs were polymorphic by different methods and/or in both pedigrees and, in such cases, segregation data from only one population were reported (usually, those representing fully informative polymorphisms were chosen). All 56 loci were positioned on a loblolly pine consensus map (Fig. 3; Sewell et al. 1999). Many of the seedling cDNAs used to generate the ESTPs were previously used as RFLP probes (Devey et al. 1994; Groover et al. 1994; Sewell et al. 1999). Comparing the two types of markers revealed that several ESTPs (EST0066, EST0606, EST0624, EST0893, EST1454, EST1576, EST1635–3′, EST1643, EST1750, EST1934, EST1955, EST2009, EST2253, EST2274 and EST2610) mapped closely to their RFLP counterparts. Estimated map locations for corresponding ESTP and RFLP markers were always very similar (Fig. 3), with small differences attributable to sampling or estimation. In fact, no recombinant gametes were observed. On the other hand, six ESTPs (EST0149, EST1626, EST1950, EST2166, EST2781 and EST2889) were not linked to their corresponding RFLPs, and EST2615 was separated from its corresponding RFLP locus by about 28 cM. This is not surprising as pine genomes are complex with multigene families, and cDNA probes reveal multiple loci (e.g., PtIFG1626, Fig. 3; Kinlaw and Neale 1997; Sewell et al. 1999); the RFLPs and ESTPs might, therefore, have detected unlinked members of a gene family. In addition, DGGE revealed polymorphisms for some ESTs (EST0107, EST0464 and EST0500) that were monomorphic using RFLPs. Several redundant clones from seedling tissues (e.g., PtIF0893 and PtIF2540) were identified on the basis of their RFLP patterns, and only one of the ESTs was mapped. In some instances, however, RFLP patterns did not clearly indicate clones that belonged to the same gene family, and ESTPs mapped redundantly. In one such case, clones PtIFG1165, PtIFG1626, PtIFG1950 and PtIFG2610 are of the same gene family but this was detected only after mapping their respective ESTs. cDNAs from the xylem library have not yet been mapped using RFLPs.

Discussion

EST sequencing projects of conifers have started to generate a considerable amount of primary sequence data

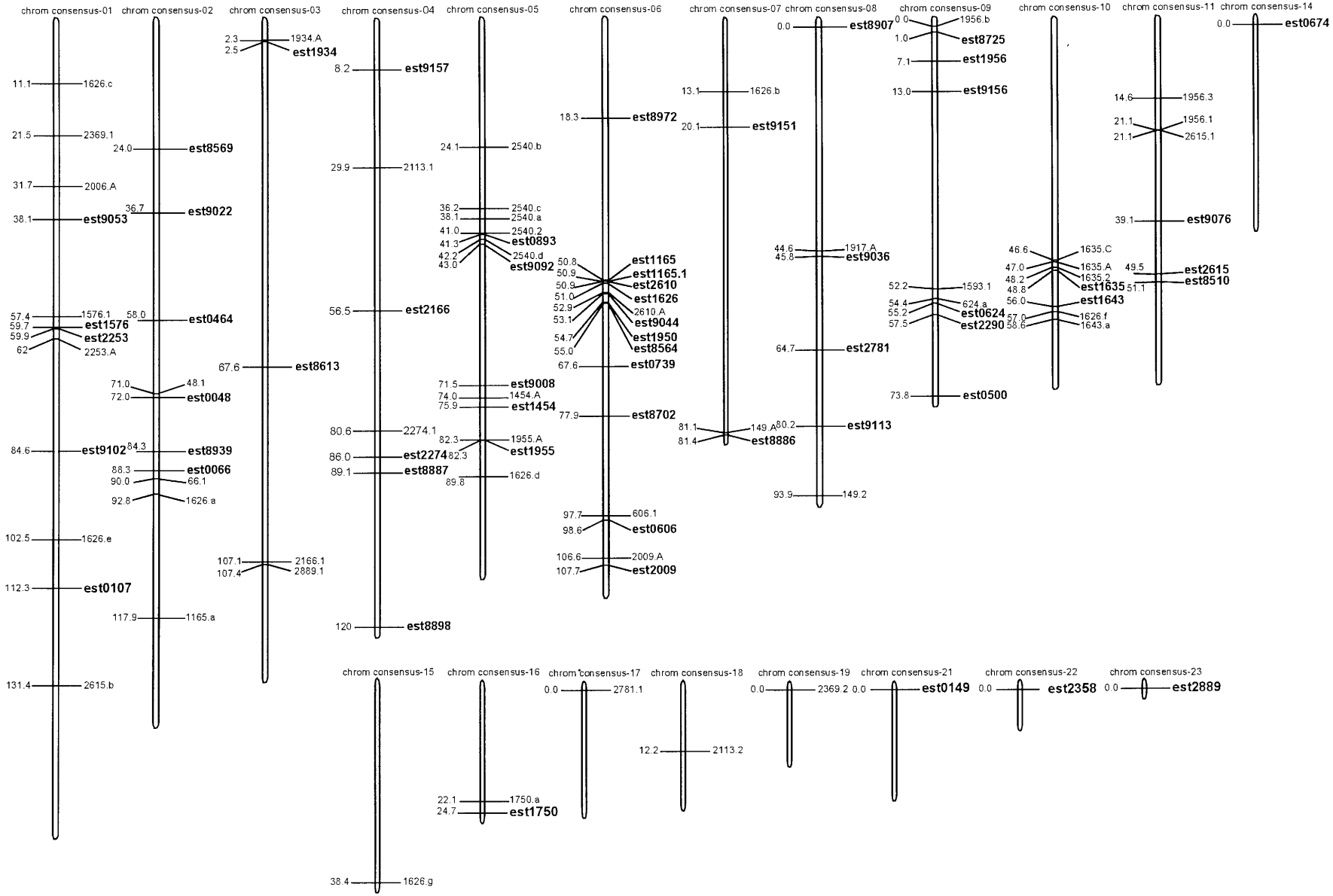

Fig. 3 Mapping of ESTPs to loblolly pine consensus genetic map. The ESTPs were located on a consensus map using marker data from both (*base* and *qtl*) mapping populations following methods described in Sewell et al. (1999). Also shown are RFLPs that were mapped with cDNAs and used to generate the corresponding ESTPs. Estimated map distances are given as cM on the left, and linkage groups are given at the top of each group

(Allona et al. 1998). However, determining the genomic location of ESTs remains a challenge. For many vertebrate species, including humans, and for some plant species, ESTs can be mapped physically using resources such as bacterial or yeast artificial chromosomes (BACs, YACs), cytogenetic analyses such as fluorescent in situ hybridization (FISH), or by using radiation hybrid analysis (Fonstein and Haselkorn 1995; Antonarakis 1998). Because technologies such as these are rudimentary in conifers, the genomic location of ESTs must be determined through linkage analysis. Harry et al. (1998) outlined a foundation for developing codominant PCR-based genetic markers from loblolly pine ESTs. This approach has been expanded, modified, and streamlined to provide an effective means of generating ESTPs and localizing them onto the consensus genetic map of loblolly pine (Sewell et al. 1999).

Once cDNA sequences are available, the rate-limiting step in developing ESTPs is in detecting DNA polymorphisms. Several strategies were adapted in order to streamline this process and increase overall efficiency. First, PCR primers were selected to have similar physical properties (e.g., T_m, length, and GC content) so that PCR reactions could be done using standardized conditions (Harry et al. 1998). Other authors (Tsumura et al. 1997; Plomion et al. 1999) used several sets of conditions. We had previously determined an optimum gradient of denaturant for individual primer sets using a perpendicular gradient gel (Temesgen et al. 2000). We have since realized that eliminating this step would simplify the overall process and also enable simultaneous screening of multiple ESTs, saving time and increasing overall efficiency.

Another step that helped improve success was the use of primers from the 3′ end of DNA sequences. More than 95% of primer sets from xylem sequences with 3′ UTRs amplified genomic DNAs, whereas only about 60% of primers from random seedling sequences successfully amplified genomic samples. The latter figure was achieved after repeated attempts were made to select and test additional primers. The final results (Table 1) seem to contradict these observations. From a total of 50 seedling primers that amplified genomic DNAs, 64% of them were mapped, whereas only 42% primers from xylem sequences could finally be mapped (Table 1). However, sequences from the seedling library had previously been more-thoroughly characterized, evaluated using RFLP banding patterns from Southern blots, and were extensively analyzed using both PCR-RFLP and DGGE. Furthermore, the simultaneous use of multiple mapping populations enhanced the detection of more genetic polymorphisms.

Differences were observed in the occurrence of introns contained in the amplified genomic fragments using primers from the two libraries. PCR primers designed to amplify about 300-bp fragment near the 3′ UTR of genes usually did not amplify introns. Only about 20% of primers based on the xylem library included introns, whereas more than 50% of those from the seedling library appeared to contain introns. The lack of introns in or near 3′ UTRs is consistent with observations in other species (e.g., humans; Wilcox et al. 1991).

The presence of introns within amplified genomic fragments poses a potential tradeoff in developing ESTPs. Relative to coding sequences, introns may exhibit a higher rate of base substitution, revealing more DNA polymorphisms. In this sense, introns would seem desirable for developing ESTPs. On the other hand, introns of unknown size, sequence, and location can lead to unpredictable results, such as failure to amplify genomic DNAs and the inability to predict and analyze melting profiles. Because of the general lack of introns in the 3′ UTRs, we found that sequences with 3′ UTRs (from the xylem library) were advantageous for developing ESTPs compared to using those from the random coding regions (seedling library).

The emphasis on 3′ sequences for primer selection also simplified the detection and interpretation of genetic polymorphisms. Kinlaw and Gerttula (1993) demonstrated that >70% of cDNAs from the seedling library belong to gene families. Amplification products arising from multiple members of a gene family can result in uninterpretable banding patterns or smearing during DGGE. Harry et al. (1998) restricted EST marker development from the seedling library to those of relatively simple families. Restrictions on complexity were not imposed for xylem ESTs, but subsequent Southern hybridizations have shown that most belong to more complex gene families. Where comparisons can be made, it is evident by the simple DGGE profiles that primers from the 3′ end tended to selectively amplify single members of a family.

To-date, little effort has been directed toward generating PCR-based EST markers for conifers. A semi-automated method enabling the mapping of ESTs on the basis of length variation is being developed for *P. radiata* and *P. taeda* by the PCR-amplification of regions flanking genes and by a laser-based fluorescence detection technique (Cato et al. 2000). Perry and Bosquet (1998) detected many length variants in black spruce [*Picea mariana* (Mill.) B.S.P.] by using only agarose gels. However, these authors sampled many more unrelated individuals than were included in the present study. In contrast, most DNA polymorphisms that we have detected involved base substitutions rather than short insertions or deletions, requiring the use of such methods as PCR-RFLPs or DGGE. The rarity of length variants was established even in fragments containing introns. Only 2 of 27 seedling ESTs and 3 of 13 xylem ESTs that contained introns detected length variants.

The efficiency of DGGE could be improved by analyzing and manipulating the predicted melting profiles of specific DNA fragments using computer programs such as MacMelt. MacMelt helps to assess how melting profiles would be affected using different combinations of PCR primers, and it also helps to assess whether GC clamps would be advantageous (Myers et al. 1985a,b). GC clamps prevent complementary DNA strands from completely denaturing as they migrate into regions of increasing denaturant concentration within a gel. In comparison to the 4–6 bp restriction site assayed by PCR-RFLP, hundreds of base pairs included in the lowest melting domain can be scanned simultaneously by adding a GC clamp with as few as 3–15 Gs and Cs, justifying our shift from PCR-RFLPs to DGGE.

Other gel-based methods such as single-strand conformation polymorphism (Orita et al. 1989) analysis have also recently been employed in pines (Plomion et al. 1999) and could be used to test for polymorphisms that were undetected or poorly resolved by DGGE.

Loblolly pine cDNAs hybridize to genomic DNA from a variety of other pines and conifers (Ahuja et al. 1994), and EST-based primers are being used to amplify genomic DNA from other conifer species (unpublished data). Therefore, such amplified fragments could be useful as orthologous markers (Lyons et al. 1997) and facilitate comparative mapping in conifers. Currently, anchor loci are being generated for the genus *Pinus* using primers from loblolly pine sequences (Brown et al., in preparation). Because these strategies are based on expressed genes, they may provide an opportunity to identify candidate genes affecting adaptive or commercially important traits.

Acknowledgements This work was supported by grants from the USDA National Research Initiative Plant Genome 93–37300–8785 and the Dendrome Database project. We thank the Weyerhaeuser Company for providing plant materials. Names of commercial manufacturers are provided for identification only, and inclusion does not imply endorsement by the USDA.

References

Ahuja MR, Devey ME, Groover AT, Jermstad KD, Neale DB (1994) Mapped DNA probes from loblolly pine can be used for restriction fragment length polymorphism mapping in other conifers. Theor Appl Genet 88:279–282

Allona I, Quinn M, Shoop E, Swope K, Cyr SS, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R, Whetten RW (1998) Analysis of xylem formation in pine by cDNA sequencing. Proc Natl Acad Sci USA 95:9693–9698

Antonarakis SE (1998) Ten years of genomics, chromosome 21, and Down syndrome. Genomics 51:1–16

Cato SA, Gardner RC, Kent J, Warr AA, Echt CS, Richardson TE (2000) A rapid PCR-based method for genetically mapping ESTs. Theor Appl Genet (in press)

Cervera MT, Plomion C, Malpica C (2000) Molecular markers and genome mapping in woody plants. In: Jain SM, Minocha SC (eds) Molecular biology of woody plants. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 375–394

Devey ME, Jermstad KD, Tauer CG, Neale DB (1991) Inheritance of RFLP loci in a loblolly pine three-generation pedigree. Theor Appl Genet 83:238–242

Devey ME, Fiddler TA, Liu B-H, Knapp SJ, Neale DB (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

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

Fonstein M, Haselkorn R (1995) Physical mapping of bacterial genomes. J Bacteriol 177:3361–3369

Groover A, Devey ME, Fiddler T, Lee J, Megraw R, Mitchell-Olds T, Sherman B, Vujcis S, Williams C, Neale DB (1994) Identification of quantitative trait loci influencing wood specific gravity in 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

Kinlaw CS, Gerttula SM (1993) Complex gene families of pines. In: Proc 22nd Southern Forest Tree Improvement Conference (Atlanta, GA.) The National Technical Information Service, Springfield, VA 22161. pp 14–17

Kinlaw CS, Neale DB (1997) Complex families in pine genomes. Trends Plant Sci 2:356–359

Lyons LA, Laughlin TF, Copeland NG, Jenkins NA, Womack JE, Obrien SJ (1997) Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nature Genet 15:47–56

Myers RM, Fischer SG, Lerman LS, Maniatis T (1985a) Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. Nucleic Acids Res 13:3131–3145

Myers RM, Fischer SG, Maniatis T, Lerman LS (1985b) Modification of the melting properties of duplex DNA by attachment of a GC-rich DNA sequence as determined by denaturing gradient gel electrophoresis. Nucleic Acids Res 13: 3111–3130

Myers RM, Maniatis T, Lerman LS (1987) Detection and localization of single base changes by denaturing gradient gel electrophresis. In: Wu R (ed) Methods Enzymol 155: 501–527

Neale DB (1998) Molecular genetic approaches to measuring and conserving adaptive genetic diversity. Proc Int Sympos on in situ Conservation of Plant Genetic Diversity, November 4–8, 1998, Antalya, Turkey, pp 385–390

Nelson CD, Kubisiak TL, Stine M, Nance WL (1994) A genetic linkage map of longleaf pine (*Pinus palustris* Mill.) based on random amplified polymorphic DNAs. J Hered 85:433–439

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-stranded conformational polymorphisms. Proc Natl Acad Sci USA 86:2766–2770

Perry DJ, Bousquet J (1998) Sequence-tagged-site (STS) markers of arbitrary genes: development, characterization and analysis of linkage in black spruce. Genetics 149:1089–1098

Plomion C, Hurme P, Frigerio JM, 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

Remington DL, Whetten RW, Liu B-H, O'Malley DM (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theor Appl Genet 98: 1279–1292

Sewell MM, Sherman BK, Neale DB (1999) A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. Genetics 151:321–330

Sheffield VC, Beck JS, Nichols B, Cousineau A, Lidral AC, Stone EM (1992) Detection of multiallele polymorphisms within gene sequences by GC-clamped denaturing gradient gel electrophoresis. Am J Hum Genet 50:567–575

Temesgen B, Neale DB, Harry DE (2000) Use of haploid mixtures and heteroduplex analysis enhance polymorphisms revealed by denaturing gradient gel electrophoresis. BioTechniques 28:114–122

Tragoonrung S, Kanazin V, Hayes PM, Blake TK (1992) Sequence-tagged-site-facilitated PCR for barley genome mapping. Theor Appl Genet 84:1002–1008

Tsumura Y, Suyama Y, Yoshimura K, Shirato N, Mukai Y (1997) Sequence-tagged-sites (STSs) of cDNA clones in *Cryptomeria japonica* and their evolution as molecular markers in conifers. Theor Appl Genet 94:764–772

Wilcox A, Khan AS, Hopkins JA, Sikela JM (1991) Use of 3′ untranslated sequences of human cDNAs for rapid chromosome assignment and conversion to STSs: implications for an expression map of the genome. Nucleic Acids Res 19:1837–1843

Yazdani R, Yeh F, Rimsha J (1995) Genome mapping of *Pinus sylvestris* (L.) using random amplified polymorphic DNA markers. For Genet 2:109–116