Molecular Breeding **5:** 21–31, 1999.

Developing SSCP markers in two *Pinus* species

C. Plomion[1,*,†], P. Hurme[2,†], J-M. Frigerio[1], M. Ridolfi[1], D. Pot[1], C. Pionneau[1], C. Avila[3], F. Gallardo[3], H. David[4], G. Neutelings[4], M. Campbell[5], F.M. Canovas[3], O. Savolainen[2], C. Bodénès[1] and A. Kremer[1]

[1]*INRA, Laboratoire de Génétique et Amélioration des Arbres Forestiers, BP45, 33610 Cestas, FRANCE (*author for correspondence; e-mail plomion@pierroton.inra.fr);* [2]*Department of Biology, University of Oulu, 90571 Oulu, Finland;* [3]*Departamento de Bioquimica y Biologia Molecular, Facultad de Ciencias, Universidad de Malaga, Campus de Teatinos s/n 29071, Malaga, Spain;* [4]*Laboratoire de Physiologie des Parois Végétales, UFR de Biologie - Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq, France;* [5]*Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK;* [†]*These authors have contributed equally to this work*

Received 23 February 1998; accepted in revised form 14 August 1998

Key words: SSCP, *Pinus pinaster*, *Pinus sylvestris*, linkage map

Abstract

This study demonstrates the feasibility of generating sequence-based markers in *Pinus* species, from data available in electronic databases. Nucleotide sequences from 23 partially or fully characterized cDNAs or genomic sequences of pines were used to design PCR primers for amplifying targeted fragments of genomic DNA from Maritime and Scots pine. Various template DNA and $MgCl_2$ concentrations, annealing temperatures, and buffer compositions were used to optimize the PCR amplifications. The polymorphism of 16 sequences was then investigated in a tree-generation inbred pedigree of Maritime pine and in a two-generation pedigree of Scots pine, using single-stranded DNA conformation polymorphism (SSCP) on polymerase chain reaction (PCR) products. The level of polymorphism was shown to be independent of (1) fragment size, (2) the presence or absence of introns in the amplified product and (3) temperature during electrophoresis. Mendelian segregation was tested for 5 SSCP markers in each species. Chromosomal locations of five genes were identified by linkage analysis with previously mapped markers in a genetic map of Maritime pine. The use of SSCP is recommended for constructing a transcriptional map for comparative mapping studies among pines and to provide useful 'candidate genes' for characterizing quantitative trait loci.

Introduction

Over the past six years, genetic linkage maps have been constructed for a growing number of *Pinus* species. These maps have been mainly constructed with random-amplified polymorphic DNA (RAPD, [52]) markers [13, 21, 25, 37, 38, 41, 53]. One large impact of the RAPD technique has been to increase the number of coniferous species amenable to mapping activities.

An important characteristic of the conifer genome is its large size compared to other plants species. For example, in the genus *Pinus*, genome sizes range from 21pg/C to 31pg/C [49]. In a range of coniferous species, DNA-DNA reassociation kinetics experiments have indicated that the ratio of the repetitive DNA fraction to the single- and low-copy fraction was 3:1 [10, 24]. In such a complex genome, Plomion *et al.* [41] showed, using a dot blot experiment, that most RAPD markers amplified in repetitive (i.e. mostly non-coding) DNA regions. These features make this marker technique of little value for characterizing quantitative trait loci (QTL) through colocalization analysis with candidate genes. In addition, RAPDs suffer from their lack of comparability across closely related species, a major disadvantage

for comparative mapping study. Alternatively, the sequence data obtained in *Pinus* species by characterization of full-length cDNAs, or else obtained in newly developed cDNA sequencing projects (R. Sederoff, personal communication), constitute a virtually unlimited source of PCR-based genetic markers corresponding to the transcribed genome. This resource should promote the development of sequence-based markers for known genes and further their use for establishing transcriptional maps for comparative mapping and characterizing QTLs.

From partially or completely sequenced genes available in databases, primer pairs can be designed to establish expressed sequence tags (ESTs). Genetic marker information can be easily disseminated by publishing primer sequences and amplification conditions. For mapping such ESTs, different techniques can be used. The simplest PCR-RFLP or cleaved amplified polymorphic sequence (CAPS) procedure consists of digesting the amplified product with a specific restriction enzyme and visualizing the products directly on agarose gel by ethidium bromide staining [48]. Other techniques like TGGE (thermal gradient gel electrophoresis, [43]), or DGGE (denaturing gradient gel electrophoresis, [32]) provide sensitive detection assays for sequence variation. These techniques are based on differences in the stability of the DNA under specific gel conditions. They are technically quite demanding and require highly controlled conditions. A more recent method for mutation detection is based on cleavage patterns generated by a structure-specific thermostable endonuclease [6]. This cleavase fragment length polymorphic (CFLP) method is, however, expensive when genotyping hundreds of individuals. An alternative, inexpensive technique for finding polymorphisms within a DNA sequence is to use single-strand conformation polymorphism (SSCP) analysis [39, 40]. SSCP is based on the conformational differences of single-stranded DNA fragments that can be detected as mobility shifts in non-denaturing polyacrylamide gel electrophoresis. SSCP has been shown to be useful for detecting polymorphism in PCR products of large size in plants [4].

In this paper we demonstrate the efficiency of SSCP to obtain molecular markers in two pine species, Maritime pine (*Pinus pinaster* Ait.) and Scots pine (*Pinus sylvestris* L.), the most widely used conifer species for reforestation in southwestern (France, Spain, Portugal) and northern (Finland, Sweden) European countries, respectively. The specific objectives of our study were: (1) to define optimal PCR conditions to amplify a set of coding sequences, (2) to develop single-stranded DNA conformation polymorphism for these sequences, (3) to check their Mendelian inheritance in two pedigrees, and (4) to map those genes that displayed clear SSCP patterns in Maritime pine. The SSCP markers developed in this study will be further used in comparative mapping studies to establish the extent of synteny among pine genetic maps, and will provide a set of useful candidate genes to understand the biological meaning of QTLs.

Material and methods

Genetic material and DNA extraction

Maritime pine genomic DNA was extracted from needles and megagametophytes (haploid nutritive tissue surrounding the embryo) as described [12] with few modifications [41]. Scots pine DNA was extracted from fresh needles using the Qiagen DNeasy Plant Mini Kit. Individual trees used in this study originated from the breeding programmes of Maritime pine and Scots pine carried out in France and Finland, respectively. For Maritime pine, a three-generation inbred pedigree [41] was used for testing Mendelian segregation ratios and genetic mapping. This F_2 family consisted of two grandparents (accessions L146 and C10), one hybrid parent (accession H12) and 62 megagametophytes collected from 62 F_2 seeds of H12. For Scots pine, a two-generation outbred pedigree was used for testing Mendelian segregation ratios. This full-sib family consisted of two parents (accession E635C used as female and E1101 used as male parent) and 48 full-sibs.

Choice of genes

We concentrated our analysis on 23 conifer nuclear sequences: cDNA or genomic clones (Table 1) corresponding to genes of known function, or EST identified on the basis of sequence similarity to other genes. These were:

- genes of the phenylpropanoid pathway with emphasis on lignification genes (see [50], for a review): PAL (phenylalanine ammonia-lyase, [51]), COMT (cafeate *O*-methyltransferase, [27]), 4CL (4-coumarate:CoA ligase, [54]), and CAD (cinnamyl-alcohol dehydrogenase, [30]).

- ICL (isocitrate lyase), an enzyme that plays a role in the conversion of membrane lipids into carbohydrate.The primer pair was designed to specifically amplify one of the two ICL cDNAs (ICL8) recently described [33].
- CuZn-SOD (superoxide dismutase) is involved in the detoxification of active oxygen species. Two primer pairs were designed to specifically amplify both choloroplastic and cytosolic forms of this gene [20].
- GS (glutamine synthetase) catalyzes the ATP-dependant incorporation of ammonium to glutamate for glutamine biosynthesis. The primer pair was designed to specifically amplify an intron containing region in GS genomic DNA (unpublished).
- STS (stilbene synthase) and CHS (chalcone synthase), two closely related genes that correspond to key enzymes in the biosynthesis of flavonoids and stilbenes [15]. Although these genes showed high degree of identity (77.4%), the alignment of the sequences allowed us to design 'gene-specific' primer pairs.
- NIR (nitrite reductase) is one of the two enzymes along with nitrate reductase, involved in nitrate assimilation through reduction of nitrate to ammonium [36].
- Protochlorophyllide oxidoreductase (PCHILDE) catalyses the light-dependent reduction of protochlorophyllide to chlorophyllide in higher plants. The primer pair was designed to amplify the third intron of this gene [47].
- two cDNAs preferentially expressed in differentiating xylem of loblolly pine [29], ARA554 and ARA556, showing strong similarity with arabinogalactane proteins.
- six drought-inducible cDNAs [7]: lp2 (homologue to a root specific *S*-adenosyl methionine synthetase of *Pinus banksiana*), lp3-1 (similar to a water deficit-inducible tomato protein), lp4 (similar to a copper-binding protein in the Japanese lacquer tree), lp5 (cell wall putative protein), lp6 (chitinase homologue described in *Pinus strobus*), lp8 (putative calcium-binding protein).
- a *Pinus caribaea* cDNA corresponding to a germin-like protein (GLP) [11]. Germins and GLPs are a multigenic family of proteins showing differential expression during development [31] and in response to pathogen infection, osmotic shock or heavy metals. Germins and GLPs are part of a superfamily of proteins, the implication of which being suggested in the maintenance of water homeostasis in extreme conditions.
- two cDNAs (unpublished), ptxmyb413 and ptxmyb126, belonging to the Myb family of transcription factors. These two members are most abundantly expressed in differentiating xylem and therefore are likely to be involved in regulating some aspect of wood formation.

PCR amplification and optimization

The position and length of the primers were chosen according to their thermodynamic parameters using the Oligo Primer Analysis Software version 4.1 [44]. When intron positions were available, forward and reverse primers were designed in exons flanking one or several introns. This was the case for 4CL, CAD, GS and PCHILDE. For proteins that are post-translationally imported into the plastid (e.g. PCHILDE), we made sure that the transit peptide signal was avoided in the priming sites. The specificity of the primers was further checked by comparing the primer sequences to the sequences of the Non-redundant Nucleotide Sequence Database using the BLASTn programme [2], with the default parameters proposed by the NCBI. The primer pairs were first tested by PCR on bulk diploid and haploid Maritime pine and Scots pine DNA samples. The annealing temperatures (ranging from 44 °C to 72 °C) were optimized with a Robocycler Gradient 96 (Stratagene). The tested $MgCl_2$ concentrations ranged from 1 mM to 3 mM. Template DNA (20 ng or 40 ng) was amplified by PCR in a total volume of 15 μl, containing 0.2 μM of each primer, home made 10$\times$ buffer (refered to as buffer 1: 67 mM Tris-HCl pH 7.5, 1 ng BSA, 0.2% 2-mercaptoethanol, 16 mM ammonium acetate) or Gibco-BRL 10$\times$ buffer (refered to as buffer 2), 200 μM of each dNTP and 1 U of *Taq* DNA polymerase (Gibco-BRL). The conditions of amplification were as follows: preliminary denaturing (4 min, 94 °C), followed by 35 cycles consisting of denaturing (45 s, 94 °C), annealing (60 s, see Table 1 for the temperature), and extension (90 s, 72 °C), and a final extension (10 min, 72 °C). The fragments were separated on 2% agarose gels and stained with ethidium bromide. For some genes, touchdown PCR was used to increase the specificity of PCR. The annealing temperature in the initial cycle was set to 2.0 °C above the optimal annealing temperature obtained without touchdown. In subsequent cycles the annealing temperature was decreased in steps of 0.1 °C per cycle for

20 cycles and then kept at a constant temperature for 15 cycles.

SSCP conditions

For SSCP analysis, we used the protocol described [4]. Polymorphism was screened using DNA from the parents (2 parents for Scots pine, 2 grandparents and 1 hybrid parent for Maritime pine) and 6 progeny (6 F_1 for Scots pine and 6 haploid megagametophytes for Maritime pine). Briefly, 4 μl of PCR products were added to 8 μl of a solution containing 95% formamide, 10 mM NaOH, 0.05% of xylene cyanol, and 0.05% of bromophenol blue. The samples were heat-denatured for 6 min in 94 °C and quickly cooled on ice for 2 min. Then 10 μl of the sample was loaded onto a 0.75 mm × 16 cm × 18 cm non-denaturing acrylamide gel (0.5× MDE (mutation detection enhancement gel; Bioprobe Systems), 0.6× TBE). Electrophoresis was run in 0.6× TBE buffer at different electrophoretic conditions, for example variable temperatures and V/cm conditions (Table 2). The gel was pre-heated or pre-cooled for 30 min in the same temperature conditions used for SSCP analysis. After the gel was removed from the apparatus, it was silver-stained as described [3], scanned and dried for 2 h at 80 °C between plastic sheets.

Data analysis

Chi-square tests were performed to examine if the observed genotypic frequencies of SSCP loci deviated from expected Mendelian ratios. Segregation patterns in both Maritime pine and Scots pine pedigrees were analysed based on at least 48 progeny. For Maritime pine, cosegregation analysis between already mapped RAPD markers [41] allowed to define the map position of SSCP markers. Significant linkages were determined using a minimal linkage LOD of 5.0 and a maximum recombination fraction Θ of 0.40. The analysis was performed with the MAPMAKER version 2.0 for the Macintosh provided by S. Tingey (DuPont, Wilmington, DE) under the haploid model.

Results and Discussion

Choice of nucleotide sequences

A significant number of protein coding genes in conifers occur in complex gene families [22, 23]. The extent of multi-copy gene families was one of the selection criterion we considered in the choice of the sequences. Multicopy genes were avoided when the information was available in the literature. Although we thought this characteristic could affect the ability to amplify a single genomic DNA fragment and further to detect polymorphism, [46] have recently reported SSCPs for 22 fatty acid synthesis genes belonging to seven gene families. Other criteria used to base our choice consisted of our own interest to provide useful candidate genes in quantitative trait dissection analysis: (1) PAL, CAD, 4CL, COMT, ARA554, ARA556, ptxmyb413 and pxtmyb126, (2) GS and NIR, (3) Cu-Zn SOD, lp2, lp3-1, lp4, lp5, lp6, lp8 and GLP. These could provide good candidate for wood properties, yield components and drought stress response QTLs, respectively. For instance, in an ongoing experiment, we will test the hypothesis that at least some of the variation observed in lignin content can be accounted for by known structural and regulatory genes of the lignification pathway.

Optimization of PCR amplification

We designed 23 primer pairs from nucleotide sequences of reference species belonging to two subsections of pines [14]. A total of 16, 6 and 1 sequences originated from *Pinus taeda* (subsection Australes), *Pinus sylvestris* (subsection Sylvestres) and *Pinus caribaea* (subsection Australes), respectively (Table 1). For most genes, the primer pairs tested were able to amplify sequences from total genomic DNA of related pines, i.e. *Pinus pinaster* (subsection Australes) and *P. sylvestris*. In both species, the amplification failed in only one case (STS). While for most sequences only one fragment was amplified, several fragments were observed for lp3-1 and lp5 in *P. pinaster* and for lp6 in both species. No information on copy number was available for these sequences. In four cases (CAD, SODcyt, lp4, lp6) a genomic fragment was only amplified in *P. pinaster*, while for *P. sylvestris* the corresponding primer pairs did not yield any product. When only one fragment was obtained in both species, the lengths of the PCR products were similar, except for COMT and ptxmyb413. The nucleotide sequences of the cloned PCR products for these two genes were obtained in both *Pinus* species. These were similar to each other and to COMT and sequences of members of the Myb family found in public databases (data not shown).

Table 1. Description of the optimal PCR conditions used for specific amplification of different genes in *Pinus pinaster* and *Pinus sylvestris*. U, upper primer; L, lower primer.

Gene (accession) ((species))	primer sequences (5'->3')	expected length[a]	observed length[b]	T°[c] (range)	$MgCl_2$ concentration (mM)[d]
•4CL	U CCCCGTCAAATCTGGCTCCT	1127	1400	68	1.0 [(1)]
(U39404)	L GGGCGCTTACTCTGCACCAC	(genomic clone)		(60-68)	
((*Pinus taeda*))					
•COMT	U AACGCCAGAATTGTGATGAA	453	Pp=870	61	1.0 [(1)]
(U39301)	L GGAACTATGCGCCACCATTA	(cDNA clone)	Ps=1000		
((*Pinus taeda*))					
•PAL	U TAGCCAAGAAAACCCTGAG	452	452	Pp=52[TD]	Pp=2.5 [(2)]
(U39792)	L ACTGATAGCGTCGTAAACCA	(genomic clone)		Ps=54	Ps=2.0 [(2)]
((*Pinus taeda*))					
•CAD	U CGAATCCTGTGAAGTGGTT	870	Pp=900	54	Pp=3.0 [(2)]
Z37991	L CTCGCTACCAATCTCTGTT	(genomic clone)	Ps=no amplification	(50-64)	
((*Pinus taeda*))					
•ARA554	U CAGCAATGACGATGGTTTTA	306	306	54	2.0 [(2)]
(U09554)	L GGCTGGGGACGAGAAC	(cDNA clone)		(50-56)	
((*Pinus taeda*))					
•ARA556	U GAAACCCGGGGCTCTGAACT	455	455	56	2.0 [(2)]
(U09556)	L CCGCAGCGAGGAAAACAA	(cDNA clone)		(56-64)	
((*Pinus taeda*))					
•CHS	U ACTCCCCCTAATGCGGTTGA	370	370	64	2.0 [(2)]
(X60754)	L CTTGGCTGCGGCTTCTTTC	(genomic clone)		(54-64)	
((*Pinus sylvestris*))					
•STS	U CTGTAGGAATGGGCGTTGAT	776	no amplification		
(X60753)	L GATTGCTCCGTCGCTGTT	(cDNA clone)			
((*Pinus sylvestris*))					
•ICL	U GGGGTAACACTGGCAATGTT	678	1600	56	2.0 [(2)]
(U39807)	L GCACAGTACCCGCACCAA	(cDNA clone)		(54-60)	
((*Pinus taeda*))					
•PCHILDE	U CTGGGCCACTTTCTTCTATC	947	Pp=1000+1020	60	1.5 [(2)]
(X66727)	L CAAGCCCTCTCAGGTCAC	(genomic clone)	Ps=947		
((*Pinus taeda*))					
•NIR	U AAAGCAAGAGCCCTGAAAAT	303	303	56	2.0 [(2)]
(X74949)	L CCTGATGGCTCCAAAGTG	(cDNA clone)		(Pp=54-60	
((*Pinus sylvestris*))				Ps=50-60)	
•SODchl	U TCCGTTTGACAGGATTGACT	284	950	60	2.0 [(2)]
(X58579)	L CCCCAGGTCATCCTCTAACT	(cDNA clone)		(50-62)	
((*Pinus sylvestris*))					
•SODcyt	U TGGGTCTTCTTAAGGCTGTT	383	Pp=700	48	2.5 [(2)]
(X58578)	L GCCAAGATCATCAGGATCAG	(cDNA clone)	Ps=no amplification	(48-50)	
((*Pinus sylvestris*))					
•GS	U AAGTTGGGCTGCACGT	600	1300	58	1.5 [(2)]
(AJ00119)	L TTGGAAGCTGGCCTGC	(cDNA clone)		52-60	
((*Pinus sylvestris*))					
•GLP	U GGGAGGCCATGTCTGAAT	209	209	52	Pp=2.0 [(2)]
(AF039201)	L GGGTGAATATGGGGAGGTA	(cDNA clone)		(48-56)	Ps=1.5 [(2)]
((*Pinus caribaea*))					
•lp2	U GATTTGCCGTGCTGTTTGCT	631	1400	58	1.5 [(2)]
(reference #7)[e]	L TCGTGCTGAGTGGAGATGAGAA	(cDNA clone)		(58-66)	
((*Pinus taeda*))					
•lp3-1	U GGAGGAGAAACAGCACCAC	716	Pp=200+800+980	Pp=48-58	Pp=2.5 [(2)]
(U52865)	L CGGAAATCACACGAAAAGAA	(cDNA clone)	Ps=716	Ps=62[TD]	Ps=2.0 [(2)]
((*Pinus taeda*))					
•lp4	U ACATACACCAGTGGCAACGA	589	Pp=900	52	2.0 [(2)]
(reference #7)[e]	L GACGGACATGAACCCAATG	(cDNA clone)	Ps=no amplification	(50-52)	
((*Pinus taeda*))					
•lp5	U CCGGACTTCTGTTTGCTT	773	Pp= 450+510+500	Pp=63	Pp=2.0 [(2)]
(AF013805)	L ATGGCCACACAATACACTGT	(cDNA clone)	Ps=faint smear	Ps=58-64	Ps=2.5 [(2)]
((*Pinus taeda*))					
•lp6	U TTTCGCTTCCCTTCATTTTC	787	Pp= 220+300+370+570	56	2.0 [(2)]
(U31309)	L TTCCACATTCGGCATCTG	(cDNA clone)	Ps=no amplification		
((*Pinus taeda*))					
•lp8	U TGATTTCGGAGGCGGAGAT	461	461	64	1.5 [(2)]
(AF013802)	L GCGCTGAGAGCTGAGGTAGG	(cDNA clone)		(50-64)	
((*Pinus taeda*))					
•ptxmyb413	U TCACCCCAGGAAGAACAAT	269	Pp=660	56	Pp=1.0 [(2)]
(unpublished)	L AATGAGGTTGTTGGTGGC	(genomic clone)	Ps=540	(Pp=51-56	Ps=1.5 [(2)]
((*Pinus taeda*))				Ps=48-58)	
•ptxmyb126	U AGTGGGATTGAAGAAGGGT	500	450	52	Pp=1.2 [(2)]
(unpublished)	L CCAACGCAATCTACAACT	(genomic clone)		(Ps=48-56	Ps=1.5 [(2)]
((*Pinus taeda*))				Pp=51-54)	

a) Expected length of product in base pair. It is indicated whether a genomic clone or a cDNA clone was used to design the primer pairs.
b) Observed length of product in base pair (Pp=*Pinus pinaster* ; Ps=*Pinus sylvestris*). Comparison between expected and observed length indicates whether an intron was amplified.
c) Indicates the optimal annealing temperature in °C as well as the temperature range where the fragment was clearly amplified. It is also indicated whether a touch down (TD) protocole was used.
d) upper case 1 and 2 correspond to 10X buffer #1 and 10X buffer #2, respectively (see the Material and Methods section).
e) see the references.

Table 2. Polymorphism obtained under different electrophoretic conditions. C, complicated pattern; M, monomorphic fragment; P, polymorphic fragment.

Gene	size (bp)	presence of intron	occurence of " P "	*Pinus pinaster*			*Pinus sylvestris*		
				4°C	15°C	37°C	4°C	15°C	37°C
4CL	1400	Yes	3	M 15h,250V	M 15h,250V	M 15h,250V	P 15h,250V	P 15h,250V	P 15h,250V
COMT	Pp=870 Ps=1000	Yes	1	M 15h,200V	M 15h,200V	P 5h,200V	M 15h,200V	C 15h,200V	C 20h,250V
PAL	452	No	4	P 5h,200V	P 5h,150V	M 5h,210V	P 5h,200V	P 5h,150V	M 5h,210V
ARA554	306	No	4	M 5h,150V	P 5h,150V	P 5h,210V	M 5h,150V	P 5h,150V	P 5h,210V
ARA556	455	No	1	C 5h,200V	C 5h,200V	M 5h,200V	C 5h,200V	C 5h,200V	P 5h,200V
CHS	370	No	5	P 5h,150V	P 5h,150V	M 5h,150V	P 5h,150V	P 5h,150V	P 5h,150V
ICL	1600	Yes	4	P 17h,250V	P 15h,250V	M 15h,250V	P 17h,250V	P 15h,250V	M 15h,250V
NIR	303	No	3	M 5h,150V	M 5h,150V	M 5h,200V	P 5h,150V	P 5h,150V	P 5h,200V
SODchl	950	Yes	4	P 16h,200V	M 14h,200V	M 15h,200V	P 16h,200V	P 14h,200V	P 15h,200V
GS	1300	Yes	2	P 15h,250V	P 15h,250V	M 15h,250V	M 15h,250V	M 15h,250V	M 15h,250V
GLP	209	No	2	P 4h,150V	P 5h,150V	M 5h,200V	M 4h,150V	M 5h,150V	M 5h,200V
lp2	1400	Yes	3	M 15h,250V	M 18h,250V	P 15h,250V	P 15h,250V	P 18h,250V	M 15h,250V
lp3-1	Ps=716	No	2	/	/	/	C 16h200V	P 15h150V	P 15h250V
lp4	Pp=900	Yes	1	C 16h,200V	P 14h,200V	C 15h,200V	/	/	/
lp8	461	No	0	C 5h,200V	C 5h,150V	C 5h,210V	C 5h,200V	C 5h,150V	C 5h,210V
ptxmyb126	450	Yes	3	M 5h,200V	M 5h,150V	P 5h,210V	P 5h,200V	M 5h,150V	P 5h,210V
occurence of P				6	7	4	8	9	8
total occurence of P			42/96		17/48			25/48	

Optimization of SSCP conditions

Of the 16 fragments tested in SSCP, 8 were polymorphic in a single species, 7 showed polymorphism in both, while one was monomorphic in the two tested pedigrees. As shown in Table 2, the level of polymorphism was significantly higher for Scots pine than for Maritime pine (25 vs. 17 polymorphisms). However, while a maximum of four alleles could segregate in the F_1 pedigree of *P. sylvestris*, only two could segregate in the megagametophytes of the *P. pinaster* hybrid tree. In addition, it has been shown from allozymes data, that *P. pinaster* had a lower level of polymorphism (R. Petit, personnal communication).

As mentioned by several groups [19, 45], fragment length can affect the separation of single-stranded DNA. The sizes of the amplified fragments ranged from 209 to 1600 bp. The 8 fragments smaller than 461 bp generated 22 polymorphisms, while the 8 fragments larger than 461 bp produced 20 polymorphisms. Thus, the ability to detect SSCP appeared not to be affected by fragment length in our study. The temperature in SSCP electrophoresis has been reported to be another important factor of variation [18]. To optimize the conditions for SSCP in our study, electrophoresis was run at three different temperatures. As shown in Table 2, this factor could affect particular SSCP patterns; for example, a polymorphism observed under the three temperatures for SODchl in Scots pine could only be detected at 4 °C for Maritime pine. But, in general, the level of polymorphism seemed to be independent of temperature, since the polymorphism level was almost similar at 4 °C, 15 °C and 37 °C, i.e. 14, 16 and 12 polymorphisms, respectively (Table 2). Therefore, the electrophoresis conditions of each SSCP have to be decided one by one. In rice, Fukuoka *et al.*

[16] noticed that lowering the temperature increased the resolution, especially for smaller fragments. This was not confirmed by our results: for the eight smaller fragments (<461 bp), 8, 7 and 7 polymorphisms were detected at 4 °C, 15 °C and 37 °C, respectively. The use of glycerol (weak denaturant) in MDE gel has been shown to affect the relative mobilities between fragments [19, 28] and could be further used to increase the efficiency of the SSCP as demonstrated by Fukuoka *et al.* [16].

Surprisingly, the presence of intronic region(s) in the amplified fragments did not yield a higher rate of polymorphism. Introns correspond to non-coding regions of eukaryotic genes. As a result, the accumulation of mutations in intron is selectively neutral [17]. It is therefore expected that introns will exhibit a higher mutation rate and degree of polymorphism than associated exons of the gene. This difference in mutation rate has the consequence that, for a given gene, two individuals with highly conserved exon sequences may differ in the sequences of the introns. Introns therefore represent good targets for detecting DNA sequence polymorphisms. Applied to our SSCP analysis, this expectation should be translated in a better efficiency of detecting polymorphism when PCR primers are amplified a region spanning one or more polymorphic introns. Introns were clearly amplified for eight sequences: 4CL, COMT, ICL, SODchl, GS, lp2, lp4 and ptxmyb126; the mean number of polymorphisms was 21 for both species. For eight other sequences (PAL, ARA554, ARA556, CHS, NIR, GLP, LP3-1 and lp8), no intron was amplified, and the mean number of polymorphism was 21 for both species.

Mendelian inheritance of SSCP fragments

A total of 10 sequences (5 in each species) was used to assess the mode of inheritance of SSCP markers. Two, three or four bands were obtained on SSCP gels and the segregation observed in the offspring was consistent with the patterns observed in the parents (Figure 1). No significant departures from Mendelian segregation ratios were observed in a χ^2 test of fit (Table 3). The electrophoretic profile could be genetically interpreted following Bodénès *et al.* [4]. As an example, we first report the segregation observed in *P. pinaster* for SODchl. Three bands (B) were present in both grand-parents and in the hybrid parent (B1, B2 and B3) (see Figure 1). In the haploid progeny 34 samples presented the bands B1 and B2, while 29 samples presented the bands B1 and B3. Hence, on the one hand, bands B1 and B2 and, on the other, bands B1 and B3 were associated and were segregating in the offspring. We concluded that the female and male grandparents were heterozygous, one allele corresponding to bands B1 and B2, and the other to bands B1 and B3. Their diploid genotype was denoted (12; 13). Other segregation obtained in *P. pinaster* were of similar interpretation with 1, 2 or 3 bands observed in both grandparents and the hybrid parent (Table 3). For *P. sylvestris*, the segregation patterns were more complex due to the use of diploid DNA samples and therefore the possibility of 4 alleles segregating in the F_1 progeny. We report here the interpretation for ARA556. Four bands (B1, B2, B3 and B4) and three bands (B1, B2, and B5) were present in the female and male parents, respectively (Figure 1). The offspring also exhibited these bands but in different combinations resulting in four genotypic groups (Figure 1), where 8 individuals presented B1, B2, and B5, 17 individuals presented B1 and B2, 13 individuals presented bands B1, B2, B3 and B4 and, finally, 10 individuals presented B1, B3, B4 and B5. The second group showed that both B1 and B2 were associated. By transitivity, the third and fourth groups showed that B3 and B4, and then B1 and B5 were also associated. Hence, we conclude that (1) the female parent was heterozygous, one allele corresponding to bands B1 and B2 and the other to bands B3 and B4; genotype noted (12; 34); (2) the male parent was also heterozygous, one allele corresponding to bands B1 and B2 and the other corresponding to bands B1 and B5; genotype noted (12; 15). For the other SSCP patterns the allelic compositions in the parents were inferred using the same kind of analysis. As noticed by Bodénès *et al.* [4], we observed that the two strands corresponding to a given allele were always distal on the gels. This previous observation greatly facilitated our own genetic interpretation. In one case (i.e. 4CL for *P. sylvestris*), the segregation pattern clearly showed that one allele was not amplified (null allele in the male parent), indicating that presumably a mutation had occurred in the priming site(s).

Mapping genes in Pines

The three-generation inbreed pedigree of Maritime pine was previously used to construct a saturated linkage map [41] based on RAPDs and proteins assayed on haploid megagametophytes. Significant linkage (LOD>5) between already localized markers and SSCPs were established for five tested genes: ICL and

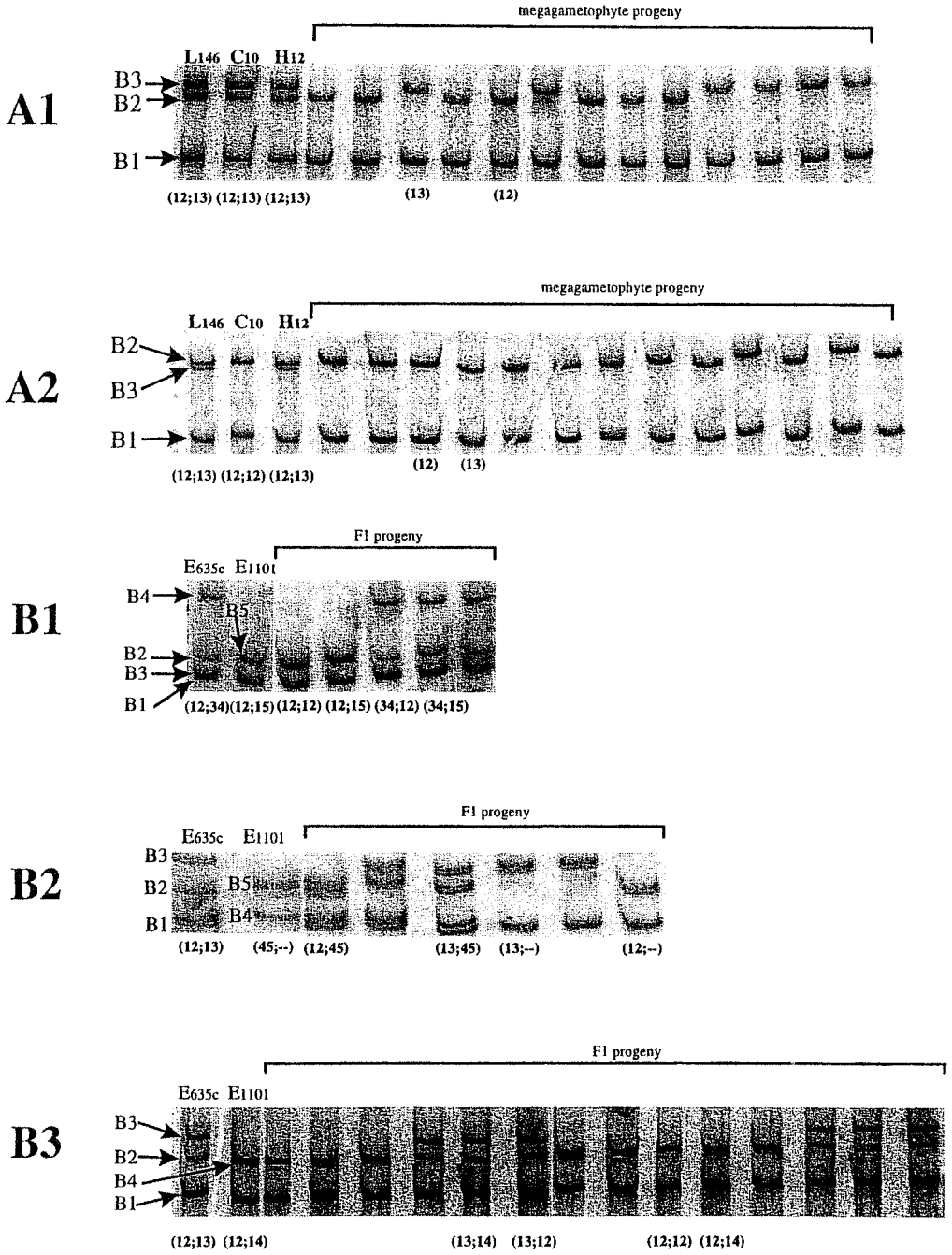

Figure 1. Silver-stained SSCP gels showing polymorphism. Separation conditions are indicated in tables 2 and 3. **A.** Example of SSCP patterns in *Pinus pinaster*. The first three lanes contain DNA amplified from both the female (L146) and male (C10) grand-parents and the hybrid parent (H12), respectively. Other lanes contain DNA amplified from a set of megagametophytes collected in seeds of the hybrid parent. **A1.** SODchl. L146, C10 and H12 genotypes are (12; 13). Megagametophytes present the haploid genotypes (12) or (13). **A2.** GS. C10, L146 and H12 genotypes are (12; 12), (12; 13) and (12; 13), respectively. Megagametophytes present the haploid genotypes (12) or (13). **B**: Example of SSCP patterns in *Pinus sylvestris*. The first two lanes contain DNA amplified from both the female (E635C) and male (E1101) parents, respectively. Other lanes contain DNA amplified from a set of F_1. **B1.** ARA556. E635C and E1101 genotypes are (12; 34) and (12; 15), respectively. F_1 offspring present the diploid genotypes (12; 15), (12; 12), (34; 12) and (34; 15). **B2.** 4CL. E635C and E1101 genotypes are (12; 13) and (45; – –), respectively. F_1 offspring present the diploid genotypes (12;45), (12; – –), (13;45) and (13; – –). **B3.** ICL. E635C and E1101 genotypes are (12; 13) and (12; 14), respectively. F_1 offspring present the diploid genotypes (12; 12), (12; 14), (13; 12) and (13; 14).

Table 3. Allelic interpretation of SSCP patterns for five sequences in each species.

	grand-parental and/or parental genotypes [a]	segregation classes	observed ratio	expected ratio	χ^2 test[b]	df
Pinus pinaster	C10 / L146 / H12					
ICL (15°C)	(13 ; 13)/(12 ; 12)/(12 ; 13)	12 /13	32/25	28.5/28.5	0.86ns	1
SODchl (4°C)	(12 ; 13)/(12 ; 13)/(12 ; 13)	12/13	34/29	31.5/31.5	0.4ns	1
GLP (15°C)	(11 ; 11)/(12 ; 12)/(11 ; 12)	11/12	30/29	29.5/29.5	0.01ns	1
PAL (15°C)	(11 ; 12)/(11 ; 11)/(11 ; 12)	11/12	28/34	31/31	0.60ns	1
GS (15°C)	(12 ; 12)/(12 ; 13)/(12 ; 13)	12/13	30/32	31/31	0.06ns	1
Pinus sylvestris	E635C / E1101					
4CL (15°C)	(12 ; 13)/(45 ; --)	(12 ; 45)/(12 ; --)/(13 ; 45)/(13 ; --)	7/16/11/14	12/12/12/12	3.8ns	3
ICL (15°C)	(12 ; 13)/(12 ; 14)	(12 ; 12)/(12 ; 14)/(13 ; 12)/(13 ; 14)	4/7/18/12	12/12/12/12	5.2ns	3
ARA556 (37°C)	(12 ; 34)/(12 ; 15)	(12 ; 15)/(12 ; 12)/(34 ; 12)/(34 ; 15)	8/17/13/10	12/12/12/12	3.8ns	3
CHS (4°C)	(12 ; 12)/(12 ; 13)	(12 ; 12)/(12 ; 13)	27/21	24/24	0.75ns	1
PAL (15°C)	(13 ; 23)/(23 ; 23)	(13 ; 23)/(23 ; 23)	21/26	23.5/23.5	0.53ns	1

a : (--) indicates a null allele
b : critical value of χ^2 at P=0.05 (1df :3.84, 3df :7.81)

SODchl in linkage group 5, PAL and GS in linkage group 2, GLP in linkage group 7. The GS gene was mapped to the same location as a protein marker identified as a glutamine synthetase [42]. The map position of these genes is available at the following web site: http://www.pierroton.inra.fr/genetics/pinus/.

A transcriptional map will be a valuable tool for comparative mapping study between conifers. It will also provide 'candidate genes' in quantitative trait dissection analysis. RFLP markers based on cDNA probes could play this role. RFLP methods are well suited for comparative mapping studies, because the same hybridization probes can be used in comparison among species. Mapped DNA probes from loblolly pine cross-hybridized to other members of Pinaceae and therefore could be used to construct RFLP maps for other related species [1]. However, although RFLPs are unlimited, they require elaborate laboratory techniques (development of specific probe libraries, use of radioisotopes, southern blot hybridization procedures and autoradiography), which make them labor intensive, time consuming and costly. The RFLP technique also requires relatively large amounts of DNA and, of course, the availability of genomic or cDNA libraries with useful probes and, finally, the maintenance of clones. The task of screening RFLPs in parents and progeny of one cross was estimated to take about 3 years [35]. In addition, conifer DNA contents are so high that Southern blots typically contain many bands making the banding patterns too complex to interpret [9]. These particular features have made RFLP genetic mapping a special challenge in conifers. Indeed, RFLP-based maps have not been developed except for two *Pinus* species: *P. teada* [9] and *P. radiata* [8].

PCR-based markers resulting from targeted amplification of coding genomic DNA constitute an alternative technique to establish a transcriptional map. If informative primer sequences can be derived from sequences of known function, then current *Pinus* maps that are mostly based on anonymous markers (e.g. RAPD, AFLP), might be easily completed. Identifying polymorphism in these sequences will remain the limiting step for gene mapping. Direct genomic sequencing would certainly be the most powerful but is a labour-intensive and costly method. Among the variety of more rapid screening methods that have been developed mostly in human genetics, SSCP has been used quite often in animal molecular genetics [5, 26, 34] and only seldom in plants [4, 16, 46]. However, due to its simplicity, low costs and the high percentage of sequence variants detected, this technique should be more widely used for mapping known genes. We also investigated the possibility of generating CAPS markers. Ten restriction endonucleases with 4, 5 and 6 bp restriction sites (*Alu*I, *Bfa*I, *Dnp*I, *Hin*fI, *Hpa*II, *Msp*I, *Rsa*I, *Taq*I, *Mbo*II, *Ava*I) were utilized on GLP, PAL and ptxmyb126 products from Maritime pine. No polymorphism was detected in this experiment.

The establishment of a transcriptional genetic map will depend on nucleotide sequence data. While many angiosperm genes have been cloned and sequenced, few gymnosperm genes are available in public databases. We therefore encourage researchers involved

in cloning and characterizing conifer genes or ESTs to publish their sequences, and we will attempt to determine their map location in *P. pinaster* and *P. sylvestris*.

Acknowledgements

This work was supported by grants from the European Union (BIO4-CT97-2125 and FAIR-CT95-0781) and from France: Centre Technique Permanent de la Sélection, and Ministère de l'Education Nationale de l'Enseignement Supérieur et de la Recherche (ACC-SV3). We also gratefully acknowledge support from PICASSO, integrated action 97-114 between France and Spain, and grants from the Research Council for Environment and National Resources to O.S. The Finnish Research Institute is acknowleged for allowing use of the Scots pine material.

References

1. Ahuja MR, Devey ME, Groover AT, Jermstad KD, Neale DB: Mapped DNA probes from loblolly pine can be used for restriction fragment length polymorphism mapping in other conifers. Theor Appl Genet 88: 279–282 (1994).
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. J Mol Biol 215: 403–410 (1990).
3. Bassam BJ, Caetano-Anolles G, Gresshoff PM: Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 196: 80–83 (1991).
4. Bodénès C, Laigret F, Kremer A: Inheritance and molecular variations of PCR-SSCP fragment in Pedunculate oak (*Quercus robur* L.) . Theor Appl Genet 93: 348–354 (1996).
5. Borsa P, Coustau C: Single-stranded DNA conformation polymorphism at the Rdl locus in *Hypothenemus hampei* (Coleoptera: Scolytidae). Heredity 76: 124–129 (1996).
6. Brown MAD, Oldenburg M, Lyamichev V, Heisler L, Grotelueschen J, Lyamichev N, Kozyavkin S, Fors L, Dahlberg J, Smith L, Olive DM: Mutation detection by cleavase fragment length polymorphism analysis. Focus 18: 2–5 (1996).
7. Chang S, Puryear JD, Dias MADL, Funkhouser EA, Newton RJ, Cairney J: Gene expression under water deficit in loblolly pine (*Pinus taeda*): isolation and characterization of cDNA clones. Physiol Plant 97: 139–148 (1996).
8. Devey ME, Bell JC, Smith DN, Neale DB, Moran GF: A genetic map for *Pinus radiata* based on RFLP, RAPD and microsatellite markers. Theor Appl Genet 92: 673–679 (1996).
9. Devey ME, Fiddler TA, Liu BH, Knapp SJ, Neale BD: An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. Theor Appl Genet 88: 273–278 (1994).
10. Dhillon SS: DNA in tree species. In: Bonga JM, Durzan DJ (eds) Cell and Tissue Culture in Forestry. vol. 1, pp. 298–313. Martinus Nijhoff (Kluwer Academic Publishers), Dordrecht, Netherlands (1987).
11. Domon JM, Dumas B, Laine E, Meyer Y, David A, David H: Three glycosylated polypeptides secreted by several embryogenic cell cultures of Pine show highly specific serological affinity to antibodies directed against the wheat germin apoprotein monomer. Plant Physiol 108: 141–148 (1995).
12. Doyle JJ, Doyle JL: Isolation of DNA from fresh plant tissue. Focus 12: 13–15 (1987).
13. Echt C, Nelson CD: Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.) Theor Appl Genet 94: 1031–1037 (1997).
14. Farjon A: Pines: Drawing and Description of the Genus *Pinus*. Brill, Leiden, Netherlands, 220 pp. (1984).
15. Fliegmann J, Schroder G, Schanz S, Britsch L, Schroder J: Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. Plant Mol Biol 18: 489–503 (1992).
16. Fukuoka S, Inoue T, Miyo A, Monna L, Zhong HS, Sasaki T, Minobe Y: Mapping of sequence-tagged sites in rice by single-strand conformation polymorphism. DNA Res 1: 271–277 (1994).
17. Futuyama DJ: Evolutionary Biology, 2nd edn. Sinauer Associates, Sunderland, MA (1986).
18. Glavac D, Dean M: Optimization of the single-strand conformation polymorphism (SSCP) technique for detection of point mutations. Human Mut 2: 404–414 (1993).
19. Hayashi K, Yandell DW: How sensitive is PCR-SSCP? Human Mut 2: 338–346 (1993).
20. Karpinski S, Wingsle G, Olsson O, Hällgren J-E: Characterization of cDNAs encoding cuZn-superoxide dismutase in Scots pine. Plant Mol Biol 18: 545–555 (1992).
21. Kaya Z, Neale DB: Linkage mapping in Turkish red pine (*Pinus brutia* Ten.) using random amplified polymorphic DNA (RAPD) genetic markers. Silvae Genet 44: 110–116 (1995).
22. Kinlaw CS, Gerttula S, Alosi MC: The non specific lipid transfer protein gene family of loblolly pine is complex. Plant Mol Biol 26: 1213–1216 (1995).
23. Kinlaw CS, Harry DE, Sederoff RR: Isolation and characterization of alcohol dehydrogenase cDNAs from *Pinus radiata*. Can J For Res 20: 1343–1350 (1990).
24. Kriebel, H B: DNA sequence components of the *Pinus strobus* nuclear genome Can J For Res 15: 1–5 (1985).
25. Kubisiak TL, Nelson CD, Nance WL, Stine M: RAPD linkage mapping in a longleaf pine × slash pine F1 family. Theor Appl Genet 90: 1110–1127 (1995).
26. Lagziel A, Lipkin E, Soller M: Association between SSCP haplotypes at the bovine growth hormone gene and milk protein percentage. Genetics 142: 945–951 (1996).
27. Li L, Popko JL, Zhang XH, Osakabe K, Tsai CJ, Joshi CP, Chiang VL: A novel multifunctional *O*-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine. Proc Natl Acad Sci USA 94: 5461–5466 (1997).
28. Liu Q, Sommer SS: Parameters affecting the sensitivities of dideoxy fingerprinting and SSCP. PCR Meth Appl 4: 97–108 (1994).
29. Loopstra CA, Sederoff R: Xylem-specific gene expression in loblolly pine. Plant Mol Biol 27: 277–291 (1995).
30. McKay JJ: A mutation in lignin biosynthesis in loblolly pine: genetic, molecular and biochemical analyses. Ph. D. thesis, North Carolina State University, Raleigh, NC, 144 pp. (1996).
31. Membré N, Berna A, Neutelings G, David A, David H, Staiger D, Sáez Vásquez J, Raynal M, Delseny M, Bernier F: cDNA sequence, genomic organization and differential expression

of three *Arabidopsis* genes for germin/oxalate oxidase-like proteins. Plant Mol Biol 35: 459–469 (1997).
32. Meyers RM, Maniatis T, Lerman LS: Detection and localization of single base changes by denaturing gradient gel electrophoresis. Meth Enzymol 155: 501–527 (1986).
33. Mullen RT, Gifford DJ: Regulation of two loblolly pine (*Pinus taeda* L.) isocitrate lyase genes in megagametophytes of mature and stratified seeds and during postgerminative growth. Plant Mol Biol 33: 593–604 (1997).
34. Mutebi JP, Black WC, Bosio CF, Sweeney WP, Craig GB: Linkage map for the Asian tiger mosquito [*Aedes* (*Stegomyia*) *albopictus*] based on SSCP analysis of RAPD markers. J Hered 88: 489–494 (1997).
35. Neale DB, Tauer CG, Gorzo DM, Jermstad KD: Restriction fragment length polymorphism mapping of loblolly pine: methods, applications and limitations. In: Proceedings of the 20th Southern Forest Tree Improvement Conference, Charleston, SC, pp. 363–372 (1989).
36. Neininger A, Seith B, Hoch B, Mohr H: Gene expression of nitrite reductase in Scots pine (*Pinus sylvestris* L.) as affected by light and nitrate. Plant Mol Biol 25: 449–457 (1994).
37. Nelson CD, Kubisiak TL, Stine M, Nance WL: A genetic linkage map of longleaf pine (*Pinus palustris* Mill.) based on ramdom amplified polymorphicd DNAs. J Hered 85: 433–439 (1994).
38. Nelson CD, Nance WL, Doudrick RL: A partial genetic linkage map of Slash pine (*Pinus elliotti* Englem var. *elliottii*) based on random amplified polymorphic DNA's. Theor Appl Genet 8: 145–151 (1993).
39. Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T: Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms Proc Natl Acad Sci USA 86: 2766–2770 (1989).
40. Orita M, Suzuki Y, Sekya T, Hayashi K: Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction Genomics 5: 874–879 (1989).
41. Plomion C, Bahrman N, Durel C-E, O'Malley DM: Genomic mapping in *Pinus pinaster* (Maritime pine) using RAPD and protein markers. Heredity 74: 661–668 (1995).
42. Plomion C, Costa P, Bahrman N: Genetic analysis of needle protein in Maritime pine. 1. Mapping dominant and codominant protein markers assayed on diploid tissue, in a haploid-based genetic map. Silvae Genet 46: 161–165 (1997).
43. Riesner DG, Steger U, Wiese M, Wulfert M, Heibey M, Henco K: Temperature-gradient electrophoresis for the detection of polymorphic DNA and for quantitative polymerase chain reaction. Electrophoresis 13: 632–636 (1992).
44. Rychlik W, Spencer WJ, Rhoads RE: Optimization of the annealing temperature for DNA amplification in vitro. Nucl Acids Res 18: 6409–6412 (1990).
45. Sheffield VC, Beck, JS, Kwitek AE, Sandstrom DW, Stone EM: The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. Genomics 16: 325–332 (1993).
46. Slabaugh MB, Huestis GM, Leonard J, Holloway JL, Rosato C, Hongtrakul V, Martini N, Toepfer R, Voetz M, Schell J, Knapp SJ: Sequence-based genetic markers for genes and gene families: single-strand conformation polymorphisms for the fatty acid synthesis genes of *Cuphea*. Theor Appl Genet 94: 400–408 (1997).
47. Spano AL, He Z, Timko MP: NADPH: protochlorophyllide oxidoreductases in white pine (*Pinus strobus*) and loblolly pine (*P. taeda*). Mol Gen Genet 236: 86–95 (1992).
48. Tragoonrung S, Kanizin V, Hayes PM, Blake TK: Sequence tagged-site-facilitated PCR for barley genome mapping. Theor Appl Genet 84: 1002–1008 (1992).
49. Wakamiya I, Newton RJ, Johnston JS, Price HJ: Genome size and environmental factors in the genus *Pinus*. Am J Bot 80: 1235–1241 (1993).
50. Wetten R, Sederoff R: Phenylalanine ammonia-lyase from loblolly pine. Plant Physiol 98: 380–386 (1992).
51. Whetten R, Sederoff R: Lignin biosynthesis. Plant Cell 7: 1001–1013 (1995).
52. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA: DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl Acids Res 18: 6531–6535 (1990).
53. Yazdani R, Yeh F, Rimsha J: Genomic mapping of *Pinus sylvestris* (L.) using random amplified polymorphic DNA markers. For Genet 2: 109–116 (1995).
54. Zhang XH, Chiang VL: Molecular cloning of 4-coumarate:coenzyme A ligase in loblolly pine and the roles of this enzyme in the biosynthesis of lignin in compression wood. Plant Physiol 113: 65–74 (1997).