Theor Appl Genet (2001) 103:1223–1231

R. Schubert · G. Mueller-Starck · R. Riegel

Development of EST-PCR markers and monitoring their intrapopulational genetic variation in *Picea abies* (L.) Karst.

Received: 29 April 2000 / Accepted: 25 August 2000

Abstract Fifteen cDNA sequences are reported for the European coniferous forest tree species Norway spruce [*Picea abies* (L.) Karst.], including the results of similarity searches in public electronic databases. The sequences were subsequently employed for the design of specific primer pairs and PCR-based amplification of genomic fragments. For seven primer pairs, polymorphic EST-PCR markers were identified among 18 trees. Their mode of inheritance was verified by analysing single-tree offspring and studying segregation among haploid endosperms in comparison to diploid tissue. Codominant inheritance was indicated for six markers, while one marker was apparently dominant. Variation of the six codominant EST markers was tested by genotyping 110 randomly selected trees in a Bavarian Norway spruce population. For comparison, the same trees were genotyped at 18 enzyme coding gene loci. There were 3.33 alleles per locus for EST markers and 3.00 for isoenzyme gene markers. In general, a trend to more even frequency distributions and larger intrapopulational variation, including observed heterozygosities, was indicated more for EST markers than for isoenzyme gene markers. The benefits of these newly developed EST-PCR markers are outlined with respect to population genetics and ecological genetics.

Keywords cDNA sequencing · EST-PCR markers · Isoenzyme gene markers · Intrapopulational variation · *Picea abies*

Communicated by H.C. Becker

R. Schubert (✉) · G. Mueller-Starck · R. Riegel
Technical University of Munich,
Center of Life and Food Sciences, Department of Plant Sciences,
Section of Forest Genetics, Am Hochanger 13,
D-85354 Freising, Germany
e-mail: schubert@genetik.forst.uni-muenchen.de
Fax: +49 08161-714861

Introduction

Norway spruce [*Picea abies* (L.) Karst.] is one of the most important tree species from both an ecological and an economical point of view (for detailed informations see Schmidt-Vogt 1986, 1987, 1991). Its prehistoric origin was probably East Asia from where the species migrated to Europe, now covering continuous areas in Central and Southeast Europe, Northeast Europe as well as Siberia. The species has been widely sown and planted in Central Europe, particularly during the last two centuries, so that large areas with artificial forests exist outside the natural habitat. Like other conifers, Norway spruce contains an extremely large nuclear genome (approx. $30–40\times10^9$ bp per haploid nucleus according to Govindaraju and Cullis 1991), which is organized in 2n=24 chromosomes.

Until the early 1990s, population genetic surveys and genome analysis as well as the development of genetic maps for Norway spruce were constrained by the availability of a only small number of morphological markers, monoterpenes and isoenzyme gene markers (Geburek and von Wuehlisch 1989; Lagercrantz and Ryman 1990; Baradat et al. 1991; Mueller-Starck and Ziehe 1991). During the last few years, however, the development of different classes of DNA-based markers has significantly advanced our knowledge of *P. abies* (Pfeiffer et al. 1997; Scheepers et al. 1997; Scotti et al. 1998, 2000). On the basis of DNA markers, the first genome maps were constructed for Norway spruce (Binelli and Bucci 1994; Bucci et al. 1997; Paglia et al. 1998). These maps are still incomplete since the markers coalesce into more than the 12 linkage groups expected, obviously demonstrating the experimental difficulties encountered when attempting to saturate a large genome of a non-domesticated species with markers.

Genetic markers are defined as heritable polymorphic characters that simply reflect differences in DNA sequences directly at the nucleotide level or indirectly at the level of gene expression. The selection of appropriate markers is essential since no marker presently used in

forest tree species has turned out to be absolutely perfect in answering all biological questions (see Neale and Harry 1994). The majority of DNA-based markers which are currently applicable for the characterization of the nuclear genome of *P. abies* represent multilocus approaches [randomly amplified polymorphic DNA (RAPD) and amplified fragment length polymorphic DNA (AFLP)], visualizing many anonymous bands simultaneously. The fundamental drawbacks of these technologies are their limited comparability among different studies and the dominant mode of inheritance of the marker bands (for detailed informations see Sunnucks 2000). In contrast to dominantly inherited markers, codominant markers allow the identification of two different alleles at a single heterozygous locus of a diploid organism. Assays confirming the level of heterozygosity or the number of alleles present per locus are essential tools in forest genetics, since the proportion of genetic variation has been found to be high within populations of the major indigenous tree species in comparison with annual plants (Mitton 1983; Hamrick and Godt 1989; Mueller-Starck 1991; Mueller-Starck et al. 1992). Genetic diversity has been verified to be correlated with fitness (Mitton and Grant 1984; Allendorf and Leary 1986) and, consequently, has been considered as an important determinant of the stability of long-lived ecosystems under complex environmental conditions (for review see Harwood and Amos 2000). In order to characterize unknown nuclear genome regions of Norway spruce, we designed additional multiallelic DNA markers with the main emphasis on codominantly inherited single-locus polymorphisms. Once developed, these markers would allow the unambiguous identification of all genotypes and a more efficient utilization of the information potential of the population studied. The generation of such markers involves gene discovering in order to address transcribed genome parts, polymerase chain reaction (PCR)-based amplification of the respective genomic regions and population tests to detect nucleotide polymorphisms, including the analysis of their inheritance.

Using an elicitor-induced cDNA library of Norway spruce, we recently started analysing expressed sequence tags (ESTs) (Schubert et al. 1997; Nageisky 1999). We report here 15 cDNA sequences, mostly corresponding to known genes, and PCR primer pairs for their amplification. For seven of them genetic markers were developed by verifying size differences or restriction site polymorphisms. The intrapopulational variation of the newly generated codominant EST-PCR markers was quantified in a population of 110 trees, and the results were compared with reference data from isoenzyme gene markers.

Materials and methods

Plant material

Leaf buds and corresponding single tree seed samples were collected from 18 trees (Forest of Kranzberg, located near Freising, Bavaria, 490 m asl) during October 1995, immediately after the trees had been cut down. Cones were induced at room temperature to release their seeds. Based on a neighbouring population, additional leaf buds were sampled during February 1999 from 110 randomly selected growing trees (age class: 50 years). Buds and seeds were stored at –20°C and +4°C, respectively, until use.

cDNA library

For the construction of a spruce cDNA library in the plasmid pSPORT1 see Galliano et al. (1993). In brief, poly(A)$^+$ RNA was isolated from photomixotrophic suspension cells which had been treated with an elicitor preparation of the fungus *Rhizosphaera kalkhoffii* Bub. Following enzymatic synthesis of cDNA, directional cloning was carried out with the Superscript Plasmid System (Gibco/BRL, Karlsruhe, Germany).

DNA extraction

Megagametophytes (endosperm tissues) were isolated by removing both the inner and outer seed coats as well as the embryo. Individual megagametophytes and single leaf buds were used to extract total genomic DNA according to the method of Doyle and Doyle (1989). For this purpose, the cell lysis buffer employed was supplemented with 1 mg/ml RNase A (Qiagen, Hilden, Germany).

The DNA of randomly picked recombinant pSPORT1 plasmid derivatives was isolated from 5-ml liquid bacterial cultures (cultivated over night at 37°C) by means of the Plasmid DNA Isolation Mini Kit and tip 20 columns as has been recommended by the manufacturer (Qiagen, Hilden, Germany).

cDNA sequence analysis and gene identification

Recombinant plasmid DNA was sequenced on both strands by an oligonucleotide walking strategy, employing the Cy5-AutoRead Sequencing Kit and the Cy5-dATP Labeling Mix (Amersham Pharmacia Biotech, Freiburg, Germany). Using electrophoretic separation on 6 *M* urea/5.5% (w/v) Hydrolink (Biozym, Hess. Oldendorf, Germany) long ranger gels (large plate, 800 min, 800 V, 60 mA, 25 W, 55°C, 0.5×TBE), we detected fluorescently labeled dideoxy chain terminated fragments using an ALFexpress automated laser sequencer (Amersham Pharmacia Biotech). Clone-specific oligonucleotides were synthesized by Gibco/BRL and MWG Biotech (Ebersberg, Germany). For the amino acid sequences deduced, online gene identification (http://www.ncbi.nlm.nih.gov/BLAST/) was performed by means of the BLASTX algorithm (Altschul et al. 1990).

Development of EST-PCR markers

Using the HIBIO DNASIS programme (version 2.5, Hitachi Software Engineering Europe, Olivet Cedex, France), we employed cDNA sequences in order to design clone-specific primer pairs (see Table 2). The reverse primer was in all cases originally labeled at the 5′ site by the dye Cy5 for fluorescent detection (all oligonucleotides and labeling reactions provided by MWG Biotech). Later, the Cy5-label was omitted from reverse primers P2 and P22 because the polymorphisms which were detected for marker PA0002 and PA0055 are easily scoreable on agarose gels.

PCR using individual primer pairs was performed in a total volume of 25 µl, containing 2.5 µl 10× PCR-buffer, 1.5 m*M* $MgCl_2$, 0.1 m*M* each dNTP, 10 pmol each primer, 0.5 U Platinum® *Taq* DNA polymerase (all reagents provided by Gibco/BRL) and approximately 40 ng template DNA. Three distinct temperature profiles were tested by means of two different thermocycler instruments. Amplification programme JAN1 (Uno instrument, Biometra, Göttingen, Germany, lid temperature at 94°C) consisted of an initial denaturation of DNA and activation of the *Taq* polymer-

ase at 94°C (5 min), annealing of the primers at 60°C (5 min) and primer extension at 72°C (1 min), followed by 42 cycles of 94°C (1 min), 60°C (1 min) and 72°C (1 min), a final elongation was carried out for 9 min at 72°C. The thermocycler (PTC-100 instrument, MJ Research, USA, lid temperature at 94°C) conditions for programme RR2 were an initial denaturation at 94°C (4 min) followed by 34 cycles, each consisting of denaturation (94°C, 1 min), primer annealing (50°C, 1 min) and primer extension (72°C, 1 min). For programme RR3 (PTC-100 instrument, as-mentioned), an initial denaturation (see RR2) was carried out, followed by 30 cycles each with 1 min at 94°C, 1 min at 55°C and 1.5 min at 72°C. Both programmes were terminated with a final elongation (RR2: 5 min; RR3: 10 min) at 72°C.

Following denaturation, genomic PCR products were directly analysed for the presence of DNA length polymorphism using the ALFexpress instrument. Fluorescently labeled fragments were detected by electrophoresis on 7 *M* urea/6% (w/v) acrylamide gels (small plate, 120 min, 1500 V, 38 mA, 34 W, 50°C), and their sizes were calculated by the ALLELE LINKS version 1.0 software (Amersham Pharmacia Biotech) using comparisons with both internal and external size standards. For marker PA0043, however, the above-mentioned Hydrolink gel matrix was employed to resolve all alleles present using the small plate. Monomorphic PCR products were digested with different restriction enzymes (*Alu*I, *Bsp*I, *Dra*I, *Hae*III, *Hha*I, *Hinf*I, *Msp*I, *Nde*II, *Rsa*I, *Taq*I and *Xba*I) according to the instructions of the manufacturer (Quantum Appligene, Heidelberg, Germany). The digestion products were detected by the ALFexpress instrument during electrophoresis (small plate with 7 *M* urea/6% (w/v) acrylamide gels). Alternatively, digested PCR products were run on standard agarose gels in 0.5× TBE buffer.

Inheritance analysis of EST-PCR markers

In the case of a putatively heterozygous pattern in the PCR reactions of bud samples (diploid DNA panel), segregation in at least 12 corresponding haploid megagametophytes (haploid DNA panel) was studied. Megagametophytes represent the female meiotic products in conifers. When there was a 1:1 segregation of the two marker bands in the megagametophyte samples together with the appropriate band combination in diploid tissue, codominant inheritance was assumed. The occurrence of two or more marker bands that appeared together in one and the same haploid megagametophyte sample was interpreted to indicate the annealing of the PCR primer pair with multiple genomic targets.

Isoenzyme analysis

The following 18 enzyme coding gene loci with known codominant inheritance were surveyed (enzyme commission code numbers are given in brackets): aconitase ACO-A (EC 4.2.1.3), alanine aminopeptidase AAP-B (EC 3.4.11.2), aspartate aminotransferase AAT-A, -B, -C (EC 2.6.1.1), glutamate dehydrogenase GDH-A (EC 1.4.1.2), isocitrate dehydrogenase IDH-A (EC 1.1.1.42), leucine aminopeptidase LAP-A,-B (EC 3.4.11.1), malate dehydrogenase MDH-C (EC 1.1.1.37), menadione reductase MNR-A, -B (EC 1.6.99.2), NADH dehydrogenase NDH-A (EC 1.6.99.3), phosphoglucose isomerase PGI-B (EC 5.3.1.9), phosphoglucomutase PGM-A (EC 2.7.5.1), 6-phosphogluconate dehydrogenase 6PGDH-B, -C (EC 1.1.1.44) and shikimate dehydrogenase SKDH-A (EC 1.1.1.25). Protein extraction of bud samples, electrophoretic methods and genotyping were done according to Mueller-Starck (1998).

Quantification of intrapopulational genetic variation

Following genotyping, all calculations were performed by means of the GSED computer software version 2.0 (E. Gillet, University of Goettingen, Germany).

Results and discussion

cDNA sequencing and gene identification

Fifteen cDNA clones (summarized in Table 1) were randomly chosen for nucleotide sequence analysis. They originate from an elicitor-induced cDNA library and therefore represent expressed genes of the spruce genome. Their deduced amino acid sequences were fed to the network BLASTX algorithm in order to detect similarities with those polypeptides which have already been deposited in public databases. For 12 queries, all showing 54% to 100% sequence identity at the amino acid level with known genes from other organisms, putative biochemical functions were assigned. The spruce genes identified encode a broad spectrum of cellular components, including house-keeping proteins and defense proteins, which are in an evolutionary sense conserved among widely divergent taxa. Three non-identified ESTs, harbouring distinct repeat motifs, were included the subsequent analysis. Such dinucleotide and trinucleotide sequence repeats are known to exist in the genetic material of many plant species and tend towards length polymorphisms due to DNA slippage (Powell et al. 1996).

Development of PCR-based EST markers

PCR primer pairs (depicted in Table 2) were designed using the 15 cDNA sequences. For genomic PCR, we mainly focused on a target range that involves potential hypervariable 3′ untranslated gene regions, with the exception of primer pairs P3/P4, P7/P8 and P27/P28. All primer pairs generated bands from diploid genomic spruce DNA whose sizes matched exactly those predicted by the cDNA sequences (see underlined fragments in Table 2), but primer pair P1/P2 yielded an amplicon 190 bp longer than the size expected.

Analysing the fluorescently labeled amplicons from different trees by vertical denaturing polyacrylamide gel electrophoresis, we detected DNA length polymorphisms for the PCR products of four primer pairs (P11/P12, P15/P16, P23/P24, and P27/P28) in contrast to the monomorphic products obtained from the remaining primer pairs (detailed informations are given in Table 2). These size-polymorphic PCR products represent genomic markers which were named PA0034, PA0043, PA0066 and PA0076 according to the cDNA clones. The 1:1 segregation patterns of the marker bands which were found in PCR samples of haploid megagametophytes and the corresponding band combination achieved from the corresponding diploid bud extracts of heterozygous trees clearly demonstrated codominant inheritance for markers PA0034, PA0043, and PA0066 (Fig. 1). These markers indicate allelic polymorphisms at three single loci, each of which follows the Mendelian mode of inheritance. For marker PA0076, however, genomic PCR bands of 213 bp and 233 bp were repeatedly seen together in one and the

Table 1 Survey of Norway spruce cDNA clones used in this study, their EMBL accession numbers, each nucleotide (nc) length in base pairs (bp) and the length of the amino acid (aa) sequence deduced. After sequencing and similarity searches in public databases, pairwise maximum matching between the amino acid sequence deduced from Norway spruce and the protein database sequence harbouring the highest identity was locally computed (HIBIO DNASIS version 2.5 software, alignment algorithm according to Needleman and Wunsch 1970) in order to present most precise alignment results

cDNA clone	EMBL accession number	Length of nucleotide (amino acid) sequence	Description/putative identification
pPA0002	AJ271125	1,620 (380)	Partial cDNA clone, coding from nc 1 to nc 1141; 54% aa identity with 380 carboxy-terminal aa of an A-like cyclin from *A capillus-veneris*
pPA0006	AJ132531	1,090 (232)	Full-length cDNA clone, coding from nc 181 to nc 879; 63% aa identity with 40 S ribosomal protein S2 of *A. thaliana*
pPA0011	AJ271127	618 (82)	Partial cDNA clone, coding from nc 3 to 249; 67% aa identity with 78 carboxy-terminal aa of the 30 S ribosomal protein S9 from *Synechocystis* strain PCC6803
pPA0023	AJ271128	1,194	Non-identified cDNA clone, harbouring the compound sequence repeat $(GAX)_9GTC(GAX)_6$ X=A,G,T,C
pPA0031	AJ271129	1,020 (229)	Full-length cDNA clone, coding from nc 73 to nc 760; polyubiquitin with 3 repeat units identical to other plant ubiquitins
pPA0034	AJ132532	300	Non-identified cDNA clone, harbouring the imperfect sequence repeat $(CT)_6TT(CT)_5(CT)_6TT(CT)_5$
pPA0038	AJ271130	631 (120)	Partial cDNA clone, coding from nc 1 to nc 361; 72% aa identity with 122 carboxy-terminal aa of the halotolerance protein HAL3 homoloque from *A. thaliana*
pPA0043	AJ132533	1,658 (432)	Partial cDNA clone, coding from nc 3 to nc 1301; 86% aa identity with 430 carboxy-terminal aa from the 78 kDa glucose regulated homologue 4 precursor of *N. tabacum* (belongs to heat-shock protein 70 family)
pPA0052	AJ132534	1,747 (444)	Partial cDNA clone, coding from nc 1 to nc 1335; 93% aa identity with translation elongation factor-1 alpha of *O. sativa*
pPA0053	AJ132535	1,036 (262)	Partial cDNA clone, coding from nc 3 to nc 791; 90% aa identity with 262 carboxy-terminal aa of the ADP/ATP translocator from *O. sativa*
pPA0055	AJ132536	628 (76)	Partial cDNA clone, coding from nc 1 to nc 231; 88% aa identity with 73 carboxy-terminal aa of the ATP synthase beta-chain from *H. brasiliensis*
pPA0066	AJ132537	932 (207)	Full-length cDNA clone, coding from nc 66 to nc 689; 78% aa identity with 60 S ribosomal protein L13–2 (cold-induced protein 24B) of *B. napus*
pPA0067	AJ132538	1,078 (196)	Partial cDNA clone, coding from nc 1 to nc 591; 78% aa identity with 195 carboxy-terminal aa of the NADPH-cytochrome P450 reductase from *P. aureus*
pPA0076	AJ132539	899	Non-identified cDNA clone, harbouring the perfect sequence repeat $(TGG)_6$
pPA0078	AJ132540	1,139 (230)	Partial cDNA clone, coding from nc 3 to nc 695; 82% aa identity with 228 carboxy-terminal aa from the Glutamate-cysteine ligase of *L. esculentum*

same megagametophyte sample, strongly indicating that the present DNA length polymorphism is derived from at least two different genomic loci (Fig. 1). Using the 18 trees tested, we exclusively observed non-amplifying alleles in the case of the 233-bp band of marker PA0076 (labeled by symbol 0–1, Fig. 1). Based on heterozygous trees, this band was present in 50% of the megagametophyte reactions and was absent in the remaining 50% of megagametophyte samples, which is in agreement with a dominant inheritance (segregation data shown in Fig. 1).

Following incubation with single restriction endonucleases, as indicated in Table 2, some of the originally monomorphic PCR products were resolved into size-polymorphic bands. In this survey, polymorphic fluorescently labeled *Hin*fI fragments were discovered for marker PA0038. Moreover, unlabeled polymorphic frag-

Table 2 Primer sequences, PCR conditions and product sizes (target sizes predicted from cDNA sequences are underlined), genetic variation including the present number of alleles (named A-E, each fragment length indicated in base pairs) and inheritance analysis for EST-PCR markers which have been investigated in 18 spruce trees of the Kranzberg test population at both the diploid bud and the haploid megagametophyte level. Monomorphic and polymorphic markers were named according to the cDNA sequences given in Table 1

EST marker	Primer sequences	PCR programme	Sizes of PCR products	Genetic variation among individuals	Type of inheritance
PA0002	P1: 5′ TCT TGA CCC TAT GGT TCA CCC TT 3′ P2: 5′ CAC TCT CAT GGC TAA TAA TCC TCC 3′	JAN1	780	Polymorphic after cut with *Hae*III (A)480+300 (B)780	Codominant
PA0006	P3: 5′ GAG CAA GAG AAA GCG TTT CAG T 3′ P4: 5′ Cy5-ATT TGC TCG AGG GAA CGG ATT 3′	RR3	246	Monomorphic	nv
PA0011	P5: 5′ GCT CCA CAG TTC TCA AAA CGT TGA G 3′ P6: 5′ Cy5-CAG CAT GTT TCG CTG GCA CTAG GC 3′	JAN1	214	Monomorphic	nv[a]
PA0023	P7: 5′ GGA GTT TTT GAC CAA AGG TCG CGA 3′ P8: 5′ Cy5-CAC TTG GCA CAG AGG CCG CAT C 3′	JAN1	94 119	Monomorphic	Multigene family
PA0031	P9: 5′ ATA TTC CAG TTT GAA GAT GC 3′ P10: 5′ Cy5-CGA CAA ACA GAT ATG GAA C 3′	JAN1	850	Monomorphic	nv
PA0034	P11: 5′ AGG TCT GCT AAT GGT TCT G 3′ P12: 5′ Cy5-ATA TCA AAC AAA CTG TGT TAG CTC 3′	RR2	(A) 212 (B) 214 (C) 226 (D) 228 (E) 230	Polymorphic	Codominant
PA0038	P13: 5′ AAC GGC ATT TGG ACT CTA TCT C 3′ P14: 5′ Cy5-TTA GAT GAT ATG CAA TGT AGA TTG A 3′	JAN1	418	Polymorphic after cut with *Hin*fI (A)186 (B)192 (C)228	Codominant
PA0043	P15: 5′ AAA ACT GGA GGA CCT TCT GG 3′ P16: 5′ Cy5-GTG AAC CTC TAC AGA AAC ACA A 3′	RR3	(A) 368 (B) 372 (C) 389 (D) 420	Polymorphic	Codominant
PA0052	P17: 5′ GCG GTG GCA GAG TTT ACA TTA 3′ P18: 5′ Cy5-ACT GTG CTT TAA CAT AAA TTC CAA GTC 3′	RR3	246	Monomorphic	nv
PA0053	P19: 5′ AAC GAA GGT ACC AAG TCA CTC 3′ P20: 5′ Cy5-TTA TTT CCT TCC GCA ACA TCA TTA AG 3′	RR2	356	Monomorphic	nv
PA0055	P21: 5′ TTG GGT TTT AGG TCA TGA CTG C 3′ P22: 5′ TCC GAT TTA TTA TAT CAA AAC TGC CTC 3′	RR3	306	Polymorphic after cut with *Dra*I (A)267+39 (B)306	Codominant
PA0066	P23: 5′ CAA GCG GTT GGT TGG AGT TCG GTT G 3′ P24: 5′ Cy5-GCA GCA AAA ACT GCA CCT CTC TTC TG 3′	RR2	(A) 154 (B) 161 (C) 171 (D) 178	Polymorphic	Codominant
PA0067	P25: 5′ GGA AAG AAA TCT CAT GGC AAA C 3′ P26: 5′ Cy5-AAT CCA GTA TCG TTT CTT ATA TGC 3′	RR2	380	Monomorphic	nv
PA0076	P27: 5′ CAT AGA CGA ACA GAG ACC CAC A 3′ P28: 5′ Cy5-AGC TAG GCT TTT CCT CAT GCT 3′	RR3	213 233	Polymorphic with null amplification allele	Dominant multigene family
PA0078	P29: 5′ AAG ATC GAG GAT GGA GAA ACT G 3′ P30: 5′ Cy5-AGC CAT CCC TTC TAC CAA AAT 3′	RR3	313	Monomorphic	nv

[a] nv. Not verifiable

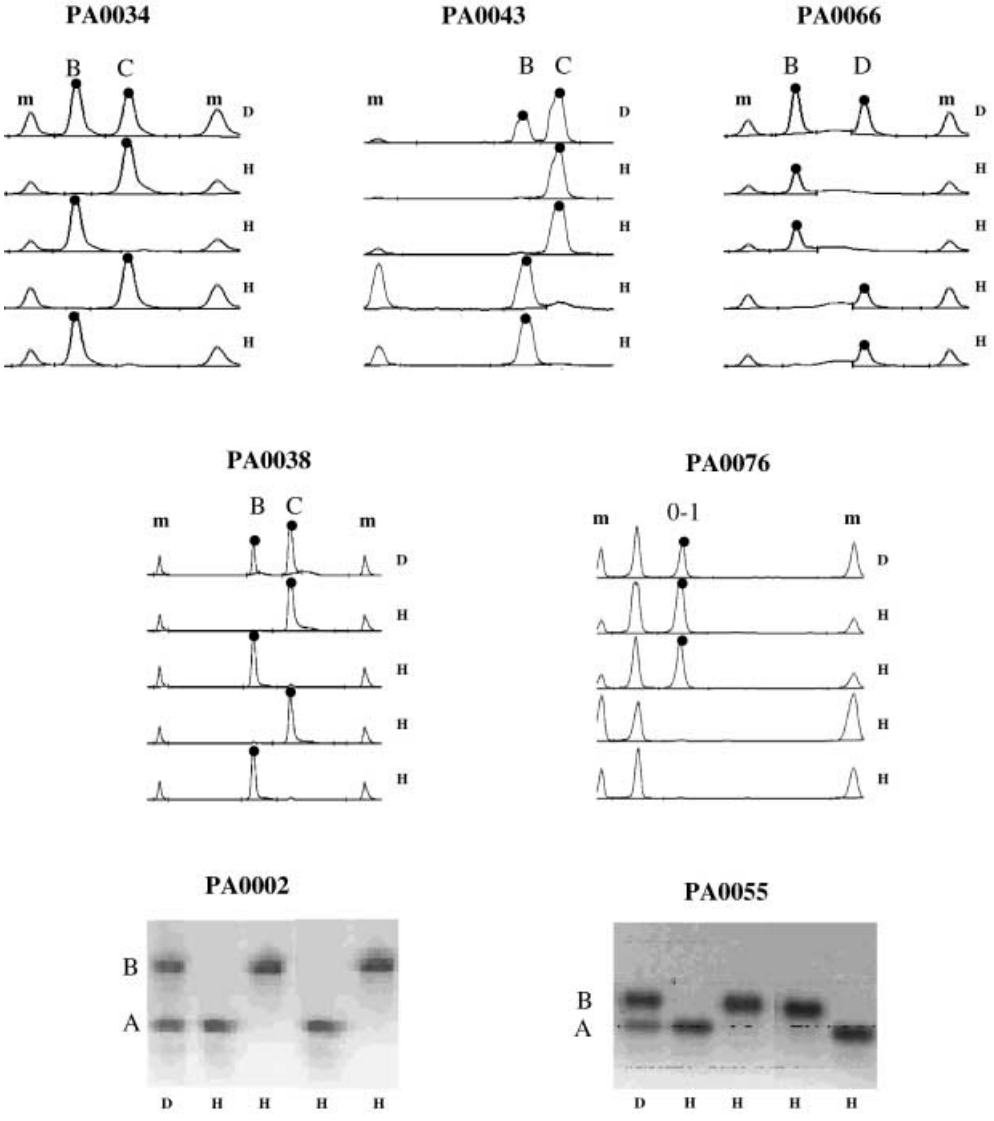

Fig. 1 Segregation patterns for six codominantly inherited EST-PCR markers (PA0002, PA0034, PA0038, PA0043, PA0055 and PA0066; alleles are indicated by *A*, *B*, *C*, and *D* according to Table 2) as well as one dominantly inherited band (indicated by *0–1*) of EST-PCR marker PA0076 in four megagametophyte samples (*H*) and the PCR banding patterns in the corresponding bud sample (*D*) of heterozygous trees. *M* Internal DNA size standards. For primer sequences and PCR conditions as well as the origin of polymorphisms and the separation techniques used, see Materials and methods and Table 2

ments were seen for markers PA0002 and PA0055 after a *Hae*III cut and a *Dra*I cut, respectively. Codominant inheritance of the marker bands was verified by analysing bud samples and the corresponding megagametophyte samples of heterozygous trees. Banding patterns were consistent with 1:1 segregation in megagametophyte PCR samples, and the appropriate band combination appeared in the PCR reaction of the bud sample (see Fig. 1). This Mendelian mode of inheritance demonstrates that the observed restriction fragment length polymorphisms represent different alleles at three single loci.

Altogether, we generated seven new EST markers for *P. abies* that were either polymorphic in length directly after genomic amplification or revealed restriction site variation. Among them, six markers showed codominant inheritance based on a comparison of both bud and megagametophyte PCR samples, whereas one marker band was dominantly inherited. Analysing the F_1 generation of controlled cross material (DNA extracts kindly provided by I. Scotti; University of Udine, Italy), we additionally investigated the inheritance of PCR bands for markers PA0002, PA0034, PA0038, PA0043, PA0055 and PA0066 with respect to both parental trees (data not shown). All loci behaved as strictly codominant Mendelian markers, doubtlessly demonstrating that the polymorphisms described are not a result of artificial PCR or partial enzymatic digestion.

Population analysis using EST markers versus isoenzyme markers

The six newly developed codominant EST markers were characterized in 110 trees of the Kranzberg test population. We observed neither additional alleles nor non-amplifying alleles in comparison with the 18 individuals that had been previously assayed in a neighbour population with the same markers. Altogether, we identified a total of 20 alleles for all EST markers surveyed (summa-

Table 3 Genetic variation parameters summarized for 110 trees (diploid bud samples) of the Norway spruce population Kranzberg using six codominant newly developed EST-PCR markers. Number of alleles (N_A), frequency distributions of alleles, number of genotypes (N_G) and the actual heterozygosity (H_a) are presented for each locus. Maximum number of genotypes (N_{Gmax}) is derived from N_A. Conditional heterozygosity (H_c) and diversity for alleles (V_A) were calculated according to Gregorius (1978, 1987) and Gregorius et al. (1986)

EST marker	Frequency distribution of alleles (%)	N_A	V_A	N_G	N_{Gmax}	H_a	H_c
PA0002	A 50 B 50	2	2.00	3	3	0.51	0.51
PA0034	A 5 B 7 C 71 D 5 E 12	5	1.90	10	15	0.46	0.80
PA0038	A 4 B 53 C 43	3	2.14	5	6	0.56	0.60
PA0043	A 55 B 16 C 24 D 5	4	2.58	8	10	0.58	0.65
PA0055	A 72 B 28	2	1.67	3	3	0.48	0.87
PA0066	A 2 B 74 C 1 D 23	4	1.66	6	10	0.41	0.79

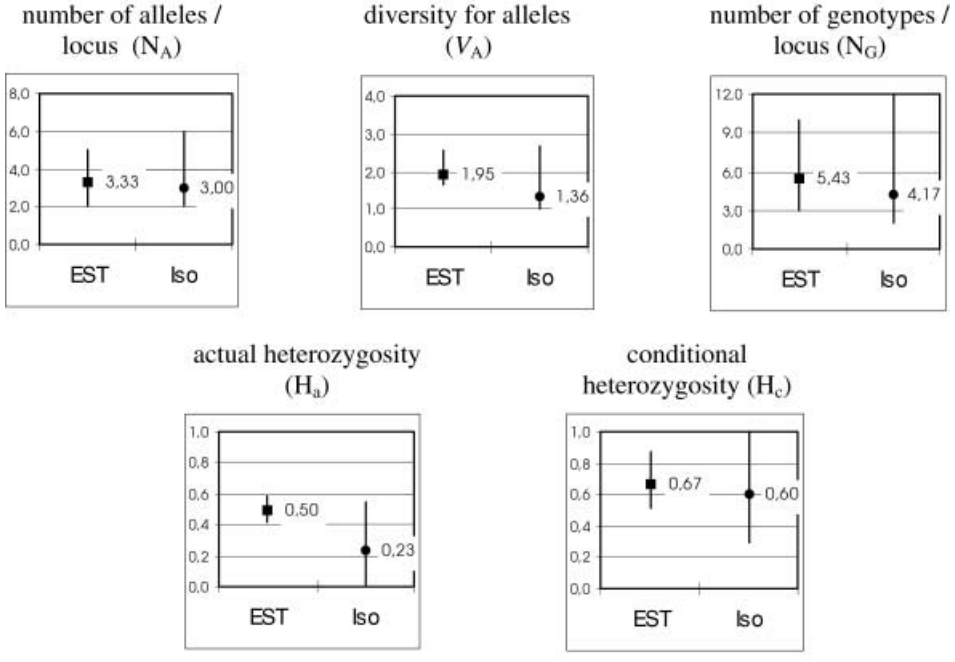

Fig. 2 Minimum, maximum and mean values of genetic variation parameters obtained for the Kranzberg spruce population (110 individual diploid bud samples) using six newly developed EST markers in comparison with data obtained from 18 isoenzyme gene markers

rized in Table 3 together with other genetic variation parameters). The number of alleles per marker (N_A) that was detected at each single locus varied between two (PA0002 and PA0055) and five (PA0034), indicating that the markers are not homogeneous in the estimation of genetic variation. The frequency distribution found among the different alleles of each EST marker deviates substantially. Marker PA0002 is the only one which shows a typical major polymorphisms, i.e. an even distribution of alleles. Its allelic diversity (V_A) is therefore equal to the number of alleles present. Markers PA0038 and PA0043 tend towards major polymorphisms. They are characterized by two or three main alleles that occur at high frequencies in combination with a single allele that is present at a low frequency. A remarkable consequence of major polymorphisms is the similarity in the levels of actual observed heterozygosity (H_a) and conditional heterozygosity (H_c) (Table 3), the latter of which takes into account the underlying allele frequencies. In contrast, markers PA0034, PA0055 and PA0066 tend towards minor polymorphisms. A typical minor polymorphism is defined by the existence of one predominant allele with at least 80% frequency and additional alleles which appear at very low frequencies. Consequently, a substantial increase in the H_c value as compared to the H_a value can be expected when the marker represents any type of minor polymorphism (see PA0034, PA0055 and PA0066 in Table 3). Heterogeneous data for single EST markers was also found with respect to the number of observed genotypes (N_G), which is equal to the maximum attainable one (N_{gmax}) only for the low polymorphic markers PA0002 and PA0055 (Table 3). For all the

other more polymorphic markers, deficiencies were observed in the test population which most likely resulted from small sample sizes or from the action of genetic selection against certain genotypes.

Finally, the EST marker-based results were compared with reference data obtained from 18 isoenzyme gene markers (shown in Fig. 2). The same 110 individuals of the Kranzberg test population were utilized for this purpose. A similar mean number of alleles was detected per locus using both marker types (3.33 compared to 3.00). However, the gene pool of six EST markers revealed larger intrapopulational genetic variation than the isoenzyme gene markers. This concerns the mean number of genotypes per locus as well as the mean diversity for alleles (Fig. 2). The remarkable heterogeneity is mainly the consequence of the different types of polymorphisms observed. Setting the threshold for the predominant alleles to a minimum of 80% and 70% frequency, the majority of isoenzyme markers, i.e. 77.77% (single data not presented), behaved as typical minor polymorphisms (AAP-B; AAT-A,-B; GDH-A; IDH-A; MDH-C; MNR-A,-B; PGM-A; SKDH-A) or tended towards minor polymorphisms (ACO-A; LAP-A, B; and PGI-B), respectively. In contrast, 50% of the EST markers revealed more even frequency distributions, as has been mentioned above. The gene pool of EST markers therefore shows a much higher amount of actual heterozygosity in contrast to isoenzyme gene markers (0.50 vs. 0.23). The average conditional heterozygosity values of both marker types are, however, similar (0.67 vs. 0.60), due to the fact that this parameter is not biased by deviations in the underlying allele frequencies, as is the case for the actual heterozygosities. This parameter indicates that, on average, 63.5% of the possible heterozygosity is realized in the random sample of 110 trees. The mean actual heterozygosity of 23% is nearly identical to those values which have been previously reported for other natural spruce populations using isoenzyme gene markers (for example, Konnert and Franke 1990; Bergmann 1991; Mueller-Starck 1995).

In conclusion, the new codominant multiallelic DNA markers presented are powerful tools for making genetic inventories of Norway spruce. They offer prospects in genetic mapping and monitoring of genetic differentiation by addressing genomic regions which were not yet characterized by markers. Based on their evenly distributed allele frequencies, the respective EST markers are expected to be highly sensitive in order to detect possible genetic loads of natural populations (i.e. inbreeding), even in the case of small sample sizes. They are therefore very useful for studies of genetic variation within and among populations and are warrant to replace such markers representing inappropriate minor polymorphisms. Furthermore, the genetic variation described here appears in some cases to be linked to genes that are known in other plant species to be expressed in response to environmental stress. For example, the ribosomal protein L13–2 represents a cold-regulated gene in *Brassica napus* (Saez-Vasquez et al. 1993). The homologous spruce gene, specified by marker PA0066, indicates natural genetic selection among stress-tolerant trees and sensitive genotypes in response to heavy metal pollution, as was observed in a pilot study (Riegel et al. 1999). In terms of marker-assisted breeding and investigating the adaptive part of the spruce genome, the successful design of markers for stress-defense genes is particularly important. Ecological genetics was limited in Norway spruce until now because the majority of routinously used isoenzyme gene markers detect variation in certain classes of water-soluble enzymes which have been affected in exceptional cases by stress-responding genetic selection (e.g. Hosius et al. 1996). Based on populations from the European geographic range, we are currently testing the new markers in a high-throughput screening of variation.

Acknowledgements The authors thank Ms. Eliane Roeschter for excellent technical assistance and Prof. Heinrich Sandermann as well as Dr. Dieter Ernst (Institute of Biochemical Plant Pathology, GSF Research Center for Environment and Health, Oberschleissheim, Germany) for utilizing a spruce cDNA library. The work was financially supported by the EU project BIO 4-CT96–0706 as well as by two grants obtained from the Deutsche Forschungsgemeinschaft (SFB 607 and Mu 774/4–1). The experiments described here comply with the current laws of the Federal Republic of Germany.

References

Allendorf W, Leary RF (1986) Heterozygosity and fitness in natural populations of animals. In: Soule ME (ed) Conservation biology: the science of scarcity and diversity. Sinauer Assoc, Sunderland, Mass, pp 57–76

Altschul SF, Gish W, Miller W, Myers EW, Lipman D (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Baradat P, Marpeau A, Walter J (1991) Terpene markers. In: Mueller-Starck G, Ziehe M (eds) Genetic variation in European populations of forest trees. Sauerlaender's, Frankfurt am Main, pp 40–66

Bergmann F (1991) Causes and consequences of species-specific genetic variation patterns in European forest tree species: examples with Norway spruce and Silver fir. In: Mueller-Starck G, Ziehe M (eds) Genetic variation in European populations of forest trees. Sauerlaender's, Frankfurt am Main, pp 192–204

Binelli G, Bucci G (1994) A genetic linkage map of *Picea abies* Karst., based on RAPD markers, as a tool in population genetics. Theor Appl Genet 88:283–288

Bucci G, Kubisiak TL, Nance WL, Menozzi P (1997) A population 'consensus', partial linkage map of *Picea abies* Karst. based on RAPD markers. Theor Appl Genet 95:643–654

Doyle J, Doyle J (1989) Isolation of plant DNA from fresh tissues. Focus 1:13–15

Galliano H, Cabane M, Eckerskorn C, Lottspeich F, Sandermann H, Ernst D (1993) Molecular cloning, sequence analysis and elicitor-/ozone-induced accumulation of cinnamyl alcohol dehydrogenase from Norway spruce (*Picea abies* L.). Plant Mol Biol 23:145–156

Geburek T, Wuehlisch G von (1989) Linkage analysis of isozyme gene loci in *Picea abies* (L.) Karst. Heredity 62:185–191

Govindaraju RD, Cullis CA (1991) Modulation of genome size in plants: the influence of breeding system and neighbourhood size. Evol Trends Plants 5:43–51

Gregorius HR (1978) The concept of genetic diversity and its formal relationship to heterozygosity and genetic distance. Math Biosci 41:253–271

Gregorius HR (1987) The relationship between the concepts of genetic diversity and differentiation. Theor Appl Genet 74:397–401
Gregorius HR, Krauhausen J, Mueller-Starck G (1986) Spatial and temporal genetic differentiation among the seeds in a stand of *Fagus sylvatica* L. Heredity 57:255–262
Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding and genetic resources. Sinauer Assoc, Sunderland, Mass., pp 77–104
Harwood J, Amos W (2000) Genetic diversity in natural populations. In: Magurran AE, May RM (eds) Evolution of biological diversity. University Press, Oxford, pp 56–69
Hosius B, Bergmann F, Hattemer HH (1996) Physiologische und genetische Anpassung von Fichtensaemlingen verschiedener Provenienz an schwermetallkontaminierte Boeden. Forstarchiv 67:108–114
Konnert M, Franke A (1990) Die Fichte (*Picea abies* (L.) Karst.) im Schwarzwald: Genetische Differenzierung von Beständen. Allg Forst Jagdztg 162:100–106
Lagercrantz U, Ryman N (1990) Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozymic variations. Evolution 44:38–53
Mitton JB (1983) Conifers. In: Tanksley SD, Orthon TJ (eds) Isoenzymes in plant genetics and breeding, part B. Elsevier, Amsterdam New York, pp 443–472
Mitton JB, Grant MC (1984) Associations among protein heterozygosity, growth rate, and developmental homeostasis. Annu Rev Ecol Syst 15:479–499
Mueller-Starck G (1991) Survey of genetic variation as inferred from enzyme gene markers. In: Mueller-Starck G, Ziehe M (eds) Genetic variation in European populations of forest trees. Sauerlaender's, Frankfurt am Main, pp 20–37
Mueller-Starck G (1995) Genetic variation in high elevated populations of Norway spruce (*Picea abies* (L.) Karst.) in Switzerland. Silva Genet 44:356–362
Mueller-Starck G (1998) Isozymes. In: Karp A, Isaac PG, Ingram DS (eds) Molecular tools for screening biodiversity. Plants and animals. Chapman and Hall, London, pp 75–81
Mueller-Starck G, Ziehe M (1991) Genetic variation in European populations of Forest trees, Sauerlaender's, Frankfurt am Main
Mueller-Starck G, Baradat P, Bergmann F (1992) Genetic variation within European tree species. New For 6:23–47
Nageisky B (1999) Identifizierung von Genen aus einer Elicitor-induzierten cDNS-Bibliothek der Gemeinen Fichte [*Picea abies* (L.) Karst.]. Diplomarbeit, Forstliche Fakultät der Technischen Universitaet in München-Weihenstephan
Neale DB, Harry DE (1994) Genetic mapping in forest trees: RFLPs, RAPDs and beyond. AgBiotech N Inf 6:107–114
Needleman SB, Wunsch CD (1970) Description of the method used in PCOMPARE and SCANISM. J Mol Biol 48:443–453
Paglia GP, Olivieri AM, Morgante M (1998) Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). Mol Gen Genet 258:466–478
Pfeiffer A, Olivieri AM, Morgante M (1997) Identification and characterization of microsatellites in Norway spruce (*Picea abies* K.). Genome 40:411–419
Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1:215–222
Riegel R, Schubert R, Mueller-Starck G, Karnosky DF, Paule L (1999) Characterization of genetic dynamics in a heavily polluted Norway spruce [*Picea abies* (L.) Karst.] population as indicated by new PCR-based DNA markers: a pilot study (Abstr). In: IUFRO Conf. (Genetic response of forest systems to changing environmental conditions – Analysis and management), University of Munich, Freising, Germany, p 32
Saez-Vasquez J, Raynal M, Meza-Basso L, Delseny M (1993) Two related, low-temperature-induced genes from *Brassica napus* are homologous to the human tumour *bbc 1* (Breast basic conserved) gene. Plant Mol Biol 23:1211–1221
Scheepers D, Eloy MC, Briquet M (1997) Use of RAPD patterns for clone verification and in studying provenance relationships in Norway spruce (*Picea abies*). Theor Appl Genet 94:480–485
Schmidt-Vogt H (1986) Die Fichte. Band II/1. Wachstum-Zuechtung-Boden-Umwelt-Holz. Paul Parey, Hamburg Berlin
Schmidt-Vogt H (1987) Die Fichte. Band I. Taxonomie-Verbreitung-Morphologie-Oekologie-Waldgesellschaften, 2nd edn. Paul Parey, Hamburg Berlin
Schmidt-Vogt H (1991) Die Fichte. Band II/3. Waldbau-Oekosysteme-Urwald-Wirtschaftswald-Ernaehrung-Duengung-Ausblick. Paul Parey, Hamburg Berlin
Schubert R, Ernst D, Sandermann H, Mueller-Starck G (1997) Gene discovery in Norway spruce based on cDNA sequencing (Abstr). In: Somatic Cell Genet Mol Genet Trees. Joint Meet IUFRO Working Parties 2.04–07 and 2.04–06. Quebec City, Canada, p 43
Scotti I, Troggio M, Soranz N, Vendramin GG, Bucci G (1998) A new set of PCR-based, locus-specific markers for *Picea abies* (L.) Karst. Mol Ecol 7:783–792
Scotti I, Magni F, Fink R, Powell W, Binelli G, Hedley PE (2000) Microsatellite repeats are not randomly distributed within Norway spruce (*Picea abies* K.) expressed sequences. Genome 43:41–46
Sunnucks P (2000) Efficient genetic markers for population biology. Tree 15:199–203