Theor Appl Genet (1997) 94:400–408

**M. B. Slabaugh · G. M. Huestis · J. Leonard
J. L. Holloway · C. Rosato · V. Hongtrakul
N. Martini · R. Toepfer · M. Voetz · J. Schell
S. J. Knapp**

Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea*

Received: 3 May 1996 / Accepted: 30 August 1996

Abstract Gene sequences are rapidly accumulating for many commercially and scientifically important plants. These resources create the basis for developing sequence-based markers for mapping and tracking known (candidate) genes, thereby increasing the utility of genetic maps. Members of most of the gene families underlying the synthesis of seed oil fatty acids have been cloned from the medium-chain oilseed *Cuphea*. Allele-specific-PCR (AS-PCR) and single-strand conformational polymorphism (SSCP) markers were developed for 22 fatty acid synthesis genes belonging to seven gene families of *Cuphea* using homologous and heterologous DNA sequences. Markers were developed for 4 fatty-acyl-acyl carrier protein thioesterase, 2 β-ketoacyl-acyl carrier protein synthase I, 4 β-ketoacyl-acyl carrier protein synthase II, 3 β-ketoacyl-acyl carrier protein synthase III, 3 acyl carrier protein, 2 β-ketoacyl-acyl carrier protein reductase, and 4 enoyl-acyl carrier protein reductase loci. Eighty-eight percent (14 of 16) of the SSCP loci were polymorphic, whereas only 9% (2 of 22) of the AS-PCR loci were polymorphic. These markers were mapped using a *Cuphea viscosissima* × *C. lanceolata* F_2 population and produced linkage groups of 10, 3, and 2 loci (3 loci segregated independently). The 10-locus linkage group had every gene but one necessary for the synthesis of 2- to 16-carbon fatty acids from acetyl-CoA and malonyl-ACP (the missing gene family was not mapped). SSCP analysis has broad utility for DNA fingerprinting and mapping genes and gene families.

Key words *Cuphea* · β-Ketoacyl-ACP synthase · Acyl-ACP thioesterase · Medium-chain fatty acids

Communicated by J. W. Snape

M. B. Slabaugh (✉) · G. M. Huestis · J. Leonard
J. L. Holloway[1] · C. Rosato · V. Hongtrakul · S. J. Knapp
Department of Crop and Soil Science, Oregon State University, Corvallis, Oregon, 97331 USA

N. Martini · R. Toepfer[2] · M. Voetz · J. Schell
Max Planck Institut fuer Zuechtungsforschung,
30 Carl-von-Linne Weg, D-50829 Cologne, Germany

Present addresses:
[1] Zymogenetics, 1201 Eastlake Ave. East, Seattle, WA 98102, USA
[2] Bundesanstalt fuer Zuechtungsforschung an Kulturpflanzen, Institut fuer Rebenzuechtung Geilweilerhof, D-76833 Siebeldingen, Germany

Abbreviations *FAS* Fatty acid synthesis · *ACP* acyl carrier protein · *FatB* fatty acyl-ACP thioesterase · *KasI* β-ketoacyl-ACP synthase I · *KasII* β-ketoacyl-ACP synthase II · *KasIII* β-ketoacyl-ACP synthase III · *Kar* β-ketoacyl-ACP reductase · *Ear* enoyl-ACP reductase · *PCR* polymerase chain reaction · *AS-PCR* allele-specific-PCR · *DGGE* denaturing gradient gel electrophoresis · *MDE* mutation detection enhancement · *RFLP* restriction fragment length polymorphism · *SSCP* single-strand conformational polymorphism · *SSR* simple sequence repeat

Introduction

Gene sequences are rapidly accumulating for many commercially and scientifically important plants, and random gene sequencing is producing catalogs of partially or completely sequenced genes for model organisms, e.g., *Arabidopsis thaliana* L. (Newman et al. 1994). Many genes underlying the synthesis of plant lipids, for example, have been cloned and sequenced over the last half decade (Toepfer and Martini 1994). These resources should promote the development and use of genetic markers for known genes, thereby augmenting random genetic markers and increasing the utility of genetic maps.

Methods for mapping sequenced genes differ in technical complexity, cost, DNA sample requirements, polymorphism rates, and phenotype resolution (Rafalski and Tingey 1993). AS-PCR markers can be

developed by using gene-specific primers and electrophoresing PCR-amplified genomic DNA fragments in agarose gels. Sequence variants (alleles) are produced when fragment lengths differ. When fragments are monomorphic, polymorphisms can sometimes be produced by using internal restriction site differences. AS-PCR markers spanning coding sequences tend to be significantly less polymorphic than non-coding sequences; however, the DNA fragments produced by this method can be subjected to a variety of electrophoretic methods for detecting sequence differences. SSCP (Orita et al. 1989a, b) and DGGE (Meyers et al. 1986) are two of the more widely used methods. These methods often detect a high frequency of sequence variants (Liu and Sommer 1994, 1995; Meyers et al. 1986).

SSCP analysis depends only on limited primary sequence variation. AS-PCR products are heat-denatured prior to electrophoresis under non-denaturing conditions (Orita et al. 1989a, b). The complementary strands from an AS-PCR product typically migrate at different rates, and alleles having small differences in nucleotide sequence often produce polymorphisms (Spinardi et al. 1991; Michaud et al. 1992; Lee et al. 1992). SSCP analysis is a sensitive, inexpensive, and comparatively fast method for detecting allelic (sequence) variation.

We describe sequence-based genetic markers for genes encoding proteins necessary for the synthesis of fatty acids of the oilseed *Cuphea* – a source of medium-chain (8 : 0–14 : 0) seed storage oils (Graham et al. 1981). This taxa has been widely used to study the biochemical and genetic mechanisms underlying the synthesis of medium-chain fatty acids. We assessed the utility of AS-PCR and SSCP analysis for mapping the FAS gene families of *Cuphea*. The methods presented should be useful for mapping and tracking candidate genes in other species.

Materials and methods

Marker development

Genetic markers were developed for members of the ACP, FatB, KasI, KasII, KasIII, Kar, and Ear gene families of *Cuphea viscosissima* and *C. lanceolata*. cDNA sequences and genetic marker loci were named using the nomenclature proposed by the Commission of Plant Gene Nomenclature (1994). The ACP gene nomenclature of von Wettstein-Knowles et al. (1994) was used. A and B suffixes were used to distinguish the Fat subclasses as proposed by Jones et al. (1995). I, II, and III suffixes were used to distinguish the Kas subclasses as proposed by Siggaard-Andersen (1993). The taxonomic origins of sequences and clones were specified by using the first letter of the genus and species; e.g., *Cw* was used for *Cuphea wrightii* and *Cl* was used for *Cuphea lanceolata*.

We used homologous and heterologous cDNA and genomic sequences to develop PCR primers for amplifying FAS gene families from *Cuphea*. DNA sequences were analyzed using Genetics Computing Group (GCG) (Madison, Wisconsin) software. The PILEUP function was used to align multiple DNA sequences, whereas the BESTFIT function was used to align pairs of DNA sequences and compute identity scores.

Primers were synthesized at the Orgon State University (OSU) on an Applied Biosystems DNA Synthesizer, Model 380B or 394 (Foster City, Calif.). *Taq* polymerase cycle-sequencing was done at OSU on an Applied Biosystems 373A Automated DNA Sequencer. *C. lanceolata* cDNA sequences for the FatB gene family (Martini et al. 1995) were aligned and used to select 2 pairs of conserved sequences for amplifying FatB genes: p-FatB15 and p-FatB13 and p-FatB20 and p-FatB21 (Table 1). *C. wrightii* KasIII cDNA sequences (Slabaugh et al. 1995) were used to design one pair of primers (p-KasIII37 and p-KasIII38) for amplifying KasIII genes based on highly conserved regions of the cDNAs (Table 1). *C. wrightii* KasII cDNA sequences (Slabaugh unpublished; GenBank #U67316, U67317) were used to design three pairs of primers for amplifying KasII genes: p-KasII44 and p-KasII45, p-KasII22 and p-KasII8, and p-KasII20 and p-KasII21 (Table 1). DNA sequences for a barley Kas12 gene (Kauppinen 1992), a castor (*Ricinus communis* L.) KasI cDNA (GenBank #L13241), and a *C. lanceolata* KasI gene (Bothmann et al. 1993) were aligned and used to design a pair of degenerate primers (p-KasI49 and p-KasI51) for amplifying KasI genes (Table 1). *C. lanceolata* ACP cDNA and genomic sequences (Voetz et al. 1994; Voetz unpublished) were used to design two pairs of primers for amplifying ACP genes: p-Acp11 and p-Acp214 for the *Acp1;1* gene and p-Acp104 and p-Acp105 for the *Acp1;2* gene (Table 1). Two *C. lanceolata* Kar cDNA sequences (Klein and Toepfer 1992) were used to design 2 pairs of gene-specific primers for amplifying Kar genes: p-Kar2 and p-Kar3 and p-Kar5 and p-Kar6 (Table 1). Three *C. lanceolata* Ear cDNA sequences (Walek et al. 1993) were used to design 4 primers (3 gene-specific primers and 1 conserved primer) for separately amplifying Ear genes. The gene-specific primers were p-Ear1, p-Ear8, and p-Ear20, while the common conserved primer was p-Ear21 (Table 1).

Marker assays

We assayed AS-PCR and SSCP markers for the ACP, FatB, KasI, KasII, and KasIII gene families and AS-PCR markers for the Kar and Ear gene families. DNA was PCR-amplified using 20 µl of reaction mixture with 10 m*M* TRIS-HCl (pH 9), 50 m*M* KCl, 2 m*M* $MgCl_2$, 0.1% Triton X-100, 100 µ*M* dNTPs, 0.1 µ*M* of each primer, 0.8 U *Taq* polymerase, and 20 ng of genomic DNA template. dATP was reduced to 50 µ*M* and 0.1 µl [^{33}P]dATP (NEN, 37-111 TB*q*/mmol, 10 mCi/ml) was added per 20 µl for radiolabeled PCR. PCR products and a 1 kb ladder (Gibco-BRL) were run in agarose gels and stained with ethidium bromide to estimate DNA yields and score AS-PCR bands.

PCR products were subjected to SSCP electrophoresis. Two µl of each PCR reaction was added to 9 µl denaturing solution (95% formamide, 0.01 *M* NaOH, 0.05% xylene cyanol, 0.05% bromophenol blue), heated to 94°C for 2 min, then chilled in an ice-water slurry. Samples of 3–5 µl were run on 0.5 × MDE gels (AT Biochem, Malvern, Pa.) using 0.6 × TBE running buffer. Gels for autoradiography (0.4 mm × 33 cm × 42 cm) were run at 8 watts constant power for 12–18 h at room temperature, transferred to 3M paper, dried, and exposed to film. Gels with unlabeled samples (0.7 mm × 16 cm × 14 cm) were run at 2 watts for 12–18 h at room temperature and stained with silver nitrate (Sanguinetti et al. 1994).

Single-stranded DNA was reamplified from silver-stained bands essentially as described by Sanguinetti et al. (1994). Single bands were excised from polyacrylamide gels, minced, and added to 100 µl of 10 m*M* TRIS-HCl (pH 9), 50 m*M* KCl, 1.5 m*M* $MgCl_2$, 0.1% Triton X-100. DNA was eluted for 60 min at 65°C with shaking. Two microliters of the eluate was added to a second PCR reaction (20 µl) and reamplified using the same primers. Secondary PCR products were purified and sequenced.

Table 1 Oligonucleotide primer names and sequences for FatB, KasI, KasII, KasIII, ACP, Kar, and Ear AS-PCR and SSCP markers

Name	Sequence	Loci
p-FatB15	5′-TGA CGG CTT TGG TCG TAC TCC TG-3′	*FatB1*,
p-FatB13	5′-TCG GAG TTA GAC CCT TGC G-3′	*FatB2*, and *FatB3*
p-FatB20	5′-GAC GGC TTT GGT CGT AC[C/T] CC-3′	*FatB1*,
p-FatB21	5′-GT[C/G] TTT TGA TTC ATC AT[A/C/G/T] [A/G]CC C-3′	*FatB2*, *FatB3*, and *FatB4*
p-KasIII37	5′-CCC AAC ACA AAT CCA CTG C-3′	*KasIII1*,
p-KasIII38	5′-ACC AAT GAT GAA TGG ATT TC-3′	*KasIII2*, and *KasIII3*
p-KasII44	5′-GCT [A/C]TC CAA GAG GAT GGA C-3′	*KasII1*,
p-KasII45	5′-T[C/T]G CAG CAT TCA GTA TAC AGA-3′	*KasII3*, and *KasII4*
p-KasII22	5′-TGA ATC CCT TTT GTG TAC C-3′	*KasII2*
p-KasII8	5′-GTG ATA GGC ATC GCA AGT G-3′	
p-KasII20	5′-TCC TCT AGG CCA TGA ACC-3′	*KasII3*
p-KasII21	5′-CAC CCA TCT GTG GAA AAA G-3′	
p-KasI49	5′-AAT GAT [C/A]G[T/A] [C/A]G[C/G] CT[C/T] GA[C/T] G-3′	*KasI1*,
p-KasI51	5′-[G/A]TA [A/T]GG [A/G]AT [G/A]AA GAA [T/A]GG [G/T]G-3′	*KasI2*
p-Acp11	5′-GTA GAG AAG ACT CAT TCT C-3′	*Acp1;1*
p-Acp214	5′-TAG TTC GAT TGC TGC TC-3′	
p-Acp104	5′-TTC TCT CGC AGG GGT TCG-3′	*Acp1;2*
p-Acp105	5′-GGA CAC TCC GGA AAC ATC-3′	
p-Kar2	5′-GCA GGG TGC AAG GTT TTA GTC-3′	*Kar27*
p-Kar3	5′-GAA ATC TCA GTG GCA GGA AG-3′	
p-Kar5	5′-AAA CAC AAA CGT CTG CTC AC-3′	*Kar10*
p-Kar6	5′-CCA GGA TTT TCT TCT CCG TT-3′	
p-Ear1	5′-TGG AAT GAC CGA GGA AGC T-3′	*Ear1*
p-Ear21	5′-ACA TAC ATT GCC TCT GA-3′	
p-Ear8	5′-AGA CAA TTA GTT CTG CCT C-3′	*Ear8*
p-Ear21	5′-ACA TAC ATT GCC TCT GA-3′	
p-Ear20	5′-GAA GAC AAC ATT CTA TAA C-3′	*Ear20*
p-Ear21	5′-ACA TAC ATT GCC TCT GA-3′	

Genetic analyses

The *C. viscosissima* line VS-6-CPR-1 (PI-574621) was crossed to the *C. lanceolata* line LN-43-1 (Crane et al. 1995) by manually emasculating VS-6-CPR-1 2 days preanthesis and applying freshly shed LN-43-1 pollen to the emasculated flower 1 day later. An F_2 population was developed by manually self-pollinating the F_1. F_1 and F_2 seed were produced in a greenhouse with 16 h of artificial light, minimum night temperatures of 16°C, and maximum day temperatures of 26°C. Leaves were harvested from 4- to 8-week-old F_2 plants and immediately lyophilized. DNA was extracted from the lyophilized leaf samples as described by Webb and Knapp (1990). AS-PCR and SSCP markers were assayed on a sample of 100 F_2 progeny.

The genetic map (Fig. 1) was built using the F_2 intercross mapping functions of MAPMAKER (Lander et al. 1987) and G-MENDEL (Holloway and Knapp 1993). Groups were estimated using a likelihood odds (LOD) threshold of 3.0. Locus orders were estimated using sums of adjacent recombination frequencies and simulated annealing (G-MENDEL) and likelihoods with a LOD threshold of 2.0 (MAPMAKER). The fit of observed-to-expected segregation ratios was tested using log-likelihood ratios estimated by G-MENDEL. The probability of the observed segregation ratio differing significantly from an expected segregation ratio, either codominant (1:2:1) or dominant (3:1), was estimated by $Pr[G \geqslant \chi_{df,\alpha}]$, where G is a log-likelihood ratio, $\chi_{df,\alpha}$ is a variable from the χ^2-distribution with df degrees of freedom ($df = 1$ for dominant and $df = 2$ for codominant markers), and α is the probability of a Type-I error.

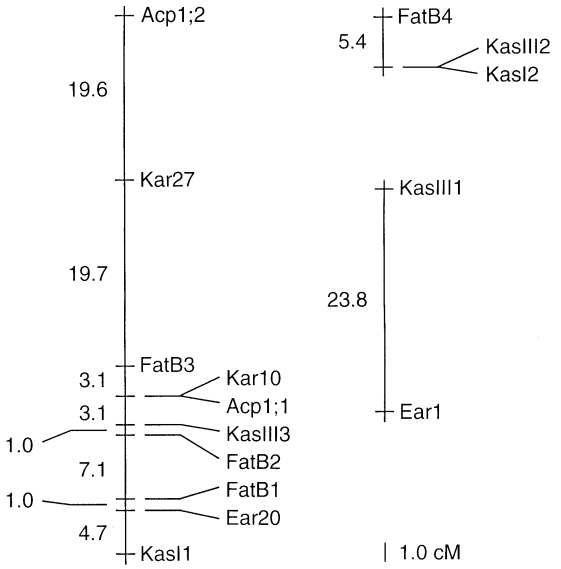


Fig. 1 Genetic map of the *Cuphea viscosissima* × *Cuphea lanceolata* (VS-6-CPR-1 × LN-43–1) F_2 population

Results

SSCP markers for the acyl-ACP thioesterase gene family

Primers p-FatB13 and p-FatB15 produced PCR products of approximately 600 and 750 bp from *C. lanceolata* and *C. viscosissima* genomic DNA (Fig. 2). A 383-bp product was predicted from the cDNA sequence; thus, these genomic products span one or more introns. The AS-PCR markers based on p-FatB13 and p-FatB15 were not polymorphic.

SSCP analysis separated the PCR products generated by p-FatB13 and p-FatB15 into three pairs of codominant polymorphic bands (Fig. 2). We extracted each of the six single-stranded bands from the *C. lanceolata* lane, reamplified DNA from each band, and analyzed the products on an SSCP gel. Adjacent bands were the complementary strands of the three PCR products. The sequences of these three products corresponded to the three *C. lanceolata* FatB (*Cl*FatB) cDNA sequences used to design p-FatB13 and p-FatB15.

Because a fourth medium-chain-specific acyl acyl carrier protein thioesterase (FatB) locus had been predicted by Martini et al. (1995), we selected and tested another set of FatB primers (p-FatB20 and p-FatB21) based on highly conserved regions of the *C. lanceolata* cDNAs. These primers produced monomorphic approximately 430- and 510-bp PCR products from *C. lanceolata* and *C. viscosissima* genomic DNA. SSCP analysis resolved the two AS-PCR markers into four SSCP markers (each AS-PCR band was a mixture of allelic DNA fragments from 2 loci). Two of the four SSCP markers produced by p-FatB20 and p-FatB21 were polymorphic. One of these was partially sequenced and found to correspond to the FatB4 cDNA (Martini et al. 1995).

SSCP markers for the β-ketoacyl-ACP synthase III gene family

SSCP markers were developed for KasIII genes using two *C. wrightii* KasIII cDNA sequences (Slabaugh et al. 1995). These sequences were aligned, and highly conserved regions were used to design PCR primers p-KasIII37 and p-KasIII38. A 271-bp product was predicted from *C. wrightii* cDNA sequences. When *C. lanceolata* and *C. viscosissima* genomic DNAs were used as templates, these primers produced 590-, 690-, and 750-bp PCR products; thus, these sequences presumably span introns. The *C. viscosissima* and *C. lanceolata* AS-PCR markers were monomorphic (Fig. 3). The three PCR products produced by p-KasIII37 and p-KasIII38 from *C. lanceolata* template DNA were sequenced and found to be homologous to *C. wrightii* KasIII cDNAs (85–96% identical within coding regions).

SSCP analysis produced three pairs of polymorphic bands (Fig. 3). The six *C. lanceolata* bands were reamplified to confirm that the members of each pair were complementary DNA strands and that the lower, middle, and upper sets corresponded to 590-, 690-, and 750-bp PCR products, respectively (data not shown).

SSCP markers for the *β*-ketoacyl-ACP synthase II gene family

Two unique KasII cDNAs were isolated from the *C. wrightii* immature embryo cDNA library and fully sequenced (Slabaugh unpublished). These sequences were aligned and used to design primers (p-KasII44 and p-KasII45) for amplifying KasII genes. Primers were selected to produce products of 333 bp from cDNA templates. The primers p-KasII44 and p-KasII45 were each twofold degenerate to accommodate nucleotide differences in *C. wrightii* KasII cDNA sequences. These primers produced PCR products of 550, 600, and 615 bp from *C. lanceolata* DNA and 570, 600, and

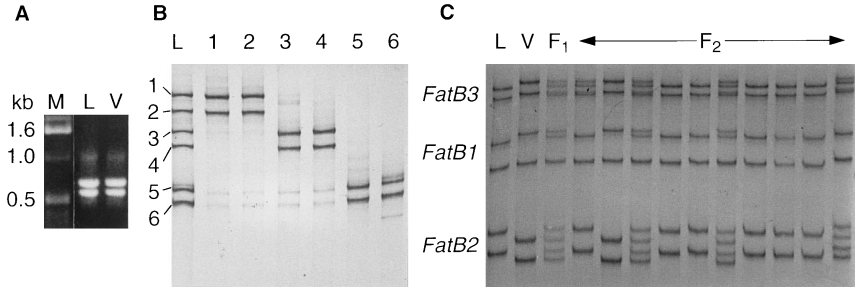


Fig. 2A–C FatB AS-PCR and SSCP markers. **Panel A** shows a length marker (*M*) and FatB PCR products (AS-PCR markers) produced from *C. lanceolata* (*L*) and *C. viscosissima* (*V*) DNA using primers p-FatB13 and p-FatB15. **Panel B** shows a silver-stained SSCP gel for *C. lanceolata* PCR products (*L*) and PCR products produced by reamplifying each of the six *C. lanceolata* PCR products (*Lanes 1–6*). **Panel C** shows an autoradiogram of an SSCP gel with PCR products produced from the parents (*L* and *V*), and ten progeny

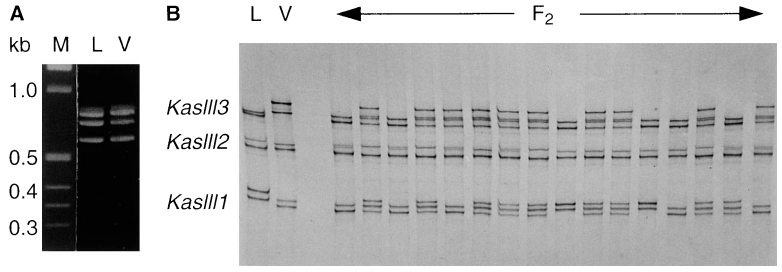


Fig. 3A, B KasIII AS-PCR and SSCP markers. **Panel A** shows a length marker (*M*) and KasIII PCR products (AS-PCR markers) produced from *C. lanceolata* (*L*) and *C. viscosissima* (*V*) DNA using primers p-KasIII37 and p-KasIII38. **Panel B** shows a silver-stained SSCP gel with PCR products produced from the parents (*L* and *V*) and 16 progeny

615 bp from *C. viscosissima* DNA; thus, one of the three AS-PCR markers was polymorphic (Fig. 4).

SSCP analysis produced six bands from each parental template. When individual SSCP bands were reamplified, we found the DNA strands from the 600- and 615-bp PCR products to be interleaved on the SSCP gel (adjacent pairs of bands were not complementary DNA strands) (Fig. 4). The reamplified *C. lanceolata* SSCP bands were completely sequenced and found to be homologous to *C. wrightii* KasII cDNA sequences (91–97% identical within coding regions). p-KasII44 and p-KasII45 amplified DNA sequences spanning two introns varying in length from 107 to 163 bp.

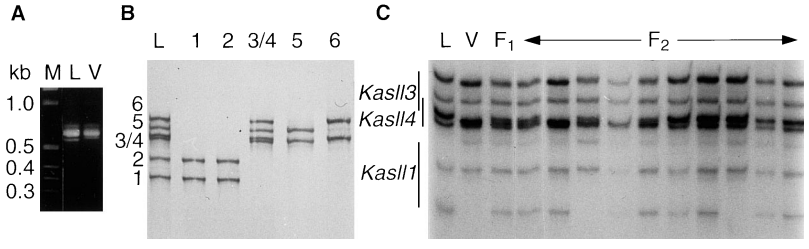
Two of the three SSCP markers (*KasII1* and *KasII3*) produced by primers p-KasII44 and p-KasII45 were polymorphic (Fig. 4). The *KasII1* SSCP marker corresponded to the polymorphic AS-PCR marker with a 550-bp allele in *C. lanceolata* and 570-bp allele in *C. viscosissima* (Fig. 4). The *KasII3* SSCP marker corresponded to the monomorphic 615-bp AS-PCR band.

The 600-bp AS-PCR marker was monomorphic (Fig. 4). The addition of glycerol to the SSCP gel and extended electrophoresis (Spinardi et al. 1991) failed to produce an SSCP polymorphism. We tested two additional primer pairs that had been used to sequence *C. wrightii* KasII cDNAs. Each pair amplified a single KasII locus. One primer pair (p-KasII20 and p-KasII21) produced monomorphic 450-bp PCR products from *C. lanceolata* and *C. viscosissima* DNA; however, the SSCP marker (KasII3) was polymorphic. Another primer pair (p-KasII22 and p-KasII8) produced monomorphic 1350-bp PCR products. The SSCP marker (*KasII2*) was polymorphic. p-KasII22 was used to sequence an intron contained within the 550-, 600-, 615-, and 1350-bp *C. lanceolata* PCR products. There were four unique intron sequences, thus demonstrating the presence of four KasII loci. AS-PCR and SSCP markers based on the 600-bp PCR product (*KasII4* locus) were monomorphic.

Fig. 4A–C KasII AS-PCR and SSCP markers. **Panel A** shows a length marker (*M*) and KasII PCR products (AS-PCR markers) produced from *C. lanceolata* (*L*) and *C. viscosissima* (*V*) DNA using primers p-KasII44 and p-KasII45. **Panel B** shows a silver-stained SSCP gel with the *C. lanceolata* PCR products (*L*) and PCR products produced by reamplifying *C. lanceolata* bands 1, 2, 5, and 6 (lanes 1, 2, 5, and 6) and a pooled sample of *C. lanceolata* bands 3 and 4 (Lane 3/4). **Panel C** shows an autoradiogram of an SSCP gel with PCR products produced from the parents (*L* and *V*), F_1, and ten F_2 progeny

SSCP markers for the β-ketoacyl-ACP synthase I gene family

PCR primers for amplifying KasI genes were selected by aligning castor (*Ricinus communis* L.) (GenBank #L13241) cDNA and barley (*Hordeum vulgare* L.) (Kauppinen 1992; GenBank #M60410) and *C. lanceolata* (Martini unpublished) genomic DNA sequences. Our aim was to produce primers with a high



probability of amplifying every member of the *C. lanceolata* and *C. viscosissima* KasI gene family. The selected sequences were highly conserved among KasI sequences and divergent from KasII sequences to preclude amplifying KasII loci. Primers p-KasI49 and p-KasI51 were 64-fold degenerate to accommodate castor, barley, and *Cuphea* KasI sequence differences. These primers flank an intron in the barley and *Cuphea* genes. Although the intron site is the same in both genes, the barley intron is tenfold larger than the 101-bp *Cuphea* intron.

Primers p-KasI49 and p-KasI51 were predicted to amplify a 222-bp PCR product from KasI cDNA and a 323-bp PCR product from *C. lanceolata* genomic DNA. This primer pair, under a variety of conditions, amplified six to seven PCR products ranging from 320 to 1350 bp from *C. lanceolata* and *C. viscosissima* DNA (Fig. 5). SSCP analysis of the PCR products produced a complex electrophorogram consisting of a series of distinct band pairs (not shown) and a cluster of six to eight very closely spaced bands (Fig. 5). We reamplified and sequenced DNA from the *C. lanceolata* cluster. This sequence was 323 bp long, spanned a 101-bp intron, and, apart from a few ambiguities, was identical to the *C. lanceolata* KasI genomic DNA sequence.

We reamplified and sequenced a pair of SSCP bands which migrated more slowly than the 323-bp *C. lanceolata* cluster. This SSCP marker (*KasI2*) corresponded to a 334-bp KasI PCR product with a 112-bp intron. Because the *C. viscosissima* allele was partially obscured by the 323-bp cluster of bands, this marker was dominant and was the only dominant SSCP marker among the 16 we tested. The dominant phenotype, however, was an artifact of the close migration

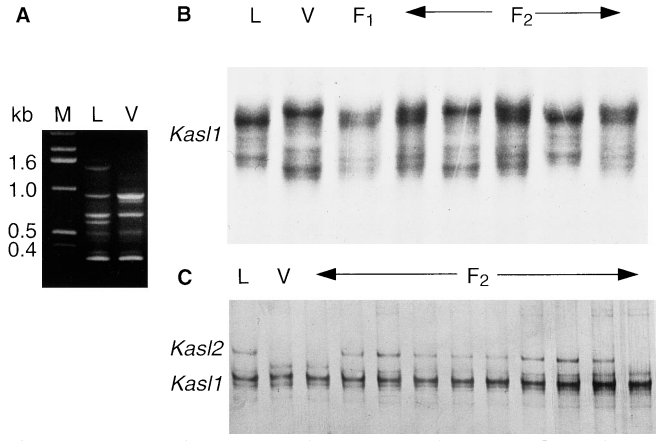


Fig. 5A–C KasI AS-PCR and SSCP markers. **Panel A** shows a length marker (*M*) and PCR products (AS-PCR markers) produced from *C. lanceolata* (*L*) and *C. viscosissima* (*V*) DNA using primers p-KasI49 and p-KasI51. **Panel B** shows an autoradiogram of a portion of an SSCP gel displaying *KasI1* locus PCR products produced from the parents (*L* and *V*), F_1, and five F_2 progeny. **Panel C** shows an autoradiogram of an SSCP gel with *KasI1* locus and *KasI2* locus PCR products produced from the parents (*L* and *V*) and ten F_2 progeny

distances for the two alleles. The 540-, 630-, 860- and 1350-bp PCR products produced by p-KasI49 and p-KasI51 were extracted from the *C. lanceolata* SSCP gel and partially sequenced. None were homologous to KasI DNA sequences.

SSCP markers for the ACP gene family

cDNA and genomic clones encoding ACP from *C. lanceolata* (Voetz et al. 1994; Voetz unpublished) were used to select primers for separately amplifying members of the ACP gene family. Primers p-Acp11 and p-Acp214 were selected from intron sequences in genomic clone Acl1 : *Cl*-g1 (GenBank Acc. No. X95253) to amplify the *Acp1;1* gene. These primers produced a 355-bp PCR product from *C. lanceolata* DNA and 340- and 355-bp PCR products from *C. viscosissima* DNA. The 355-bp PCR product was polymorphic on SSCP gels.

Primers p-Acp104 and p-Acp105 were selected from the cDNA clone Acl1 : *Cl*-1b (GenBank #X77621) to amplify the *Acp1;2* gene. These primers produced a 255-bp PCR product from *C. lanceolata* and *C. viscosissima* DNA. The 255-bp product was polymorphic on SSCP gels. We tested several different primer sets for a third ACP sequence from *C. lanceolata* (Acl1;3), but none produced polymorphic AS-PCR or SSCP bands.

AS-PCR markers for the β-ketoacyl-ACP reductase gene family

Primers were chosen to separately amplify the two Kar genes (*Cl*Kar27 and *Cl*Kar10) described by Klein and Toepfer (1992). Primers p-Kar3 and p-Kar5 were developed using the *Cl*Kar27 cDNA sequence. These primers produced a 1200-bp PCR product from *C. lanceolata* DNA and a 1450-bp PCR product from *C. viscosissima* DNA. The length predicted from the *Cl*Kar27 cDNA sequence was 643 bp. This was one of the few polymorphic AS-PCR markers. Primers p-Kar2 and p-Kar6 were developed from the *Cl*Kar10 cDNA sequence. These primers produced a 1350-bp PCR product from *C. lanceolata* and *C. viscosissima* DNA. The length predicted from the *Cl*Kar10 cDNA sequence was 534 bp. A polymorphic AS-PCR marker was developed by digesting these products with *Rsa*I – the *C. viscosissima* product was not digested, whereas the *C. lanceolata* product was digested into 580-and 780-bp fragments.

AS-PCR markers for the enoyl-ACP reductase gene family

Walek et al. (1993) described three *C. lanceolata* Ear cDNAs (*Cl*Ear1, *Cl*Ear8, and *Cl*Ear20). We chose three

specific upstream primers (p-Ear1, p-Ear8, and p-Ear20) and a highly conserved downstream primer (p-Ear21) to separately amplify these genes. Primers p-Ear1 and p-Ear21 produced a 1000-bp PCR product from *C. lanceolata* and *C. viscosissima* DNA. These products were digested with *Alu*I, which cut the *C. lanceolata* allele, to produce a polymorphic AS-PCR marker for the *Ear1* locus. Primers p-Ear8 and p-Ear21 produced a 1000-bp PCR product from *C. lanceolata* and *C. viscosissima* DNA. We found no restriction site differences between the two species and thus no AS-PCR polymorphisms. Primers p-Ear20 and p-Ear21 produced a 1050-bp PCR product from *C. lanceolata* and *C. viscosissima* DNA. This product was digested with *Xho*I, thereby producing a polymorphic marker for the *Ear20* locus (the *C. viscosissima* allele was cut, whereas the *C. lanceolata* was not cut). A fraction of the 1050-bp *C. viscosissima* band remained undigested irrespective of the restriction enzyme concentration used (data not shown), thus demonstrating that the original *C. viscosissima* PCR product was amplified from at least two Ear genes (one or more genes in the mixture lack the *Xho*I site).

Genetic analysis of the FAS gene markers

Ten FAS gene loci (*Acp1;1*, *Acp1;2*, *Ear20*, *FatB1*, *FatB2*, *FatB3*, *Kar10*, *Kar27*, *KasI1*, and *KasIII3*) mapped to one linkage group (Fig. 1). Every marker within this linkage group had severely distorted segregation ratios with strong selection against *C. viscossisima* (V) alleles (strong gamete selection) (Table 2). Selection was most severe for the *Acp1;1* and *Kar10* loci – these loci completely lacked *C. viscossisima* homozygotes (Table 2). Selection weakened as the distance from these loci increased in either direction, but was nevertheless still strong and significant over this chromosome segment.

Three FAS gene loci (*FatB4*, *KasI2*, and *KasIII2*) mapped to a second linkage group (Fig. 1). There was significant gamete selection in this linkage group, but selection was milder and in the opposite direction (against *C. lanceolata* alleles) (Table 2).

Gamete selection has been reported for *C. viscosissima* × *C. lanceolata* hybrids (Brandt and Knapp 1993). This interspecific hybrid is fertile, but typically produces half the number of chiasmata of the species *per se*. Because recombination is reduced, the genetic distances for the interspecific map (Fig. 1) should be less than for either intraspecific map. Reduced recombination and distorted segregation, however, should not affect the groupings or locus orders (Bailey 1961), but could decrease the resolution of the map because fewer recombinants are produced by the interspecific F_1.

Discussion

SSCP markers were developed for 16 FAS gene loci belonging to the FatB, KasI, KasII, KasIII, and ACP gene families. Of the 16 SSCP markers 14 (88%) were polymorphic, whereas only 9% of the undigested AS-PCR markers (2 of 22) were polymorphic. We produced polymorphic markers for 3 additional AS-PCR markers by digesting the AS-PCR products with restriction enzymes for known restriction sites. This strategy was tested with Kar and Ear sequences only.

Table 2 Observed phenotypic ratios, log likelihood ratios (G), and test probabilities ($Pr[G > \chi^2]$) for AS-PCR or SSCP markers for fatty acid synthesis genes segregating in a *C. viscosissima* × *C. lanceolata* F_2 population

Gene family	Locus	Phenotypic ratio[a]			*G*	$Pr[G > \chi^2]$
		V/V	V/L	L/L		
Acyl carrier protein	*Acp1;2*	11	42	49	29.2	< 0.0001
β-Ketoacyl-ACP reductase	*Kar27*	4	45	53	51.5	< 0.0001
Fatty-acyl-ACP thioesterase	*FatB3*	2	46	54	61.4	< 0.0001
β-Ketoacyl-ACP reductase	*Kar10*	0	48	54	75.2	< 0.0001
Acyl carrier protein	*Acp1;1*	0	48	54	75.2	< 0.0001
β-Ketoacyl-ACP synthase III	KasIII3	3	48	51	52.0	< 0.0001
Fatty-acyl-ACP thioesterase	*FatB2*	3	50	49	49.2	< 0.0001
Fatty-acyl-ACP thioesterase	*FatB1*	3	49	50	50.6	< 0.0001
Enoyl-ACP reductase	*Ear20*	3	47	48	48.1	< 0.0001
β-Ketoacyl-ACP synthase I	*KasI1*	4	48	50	46.7	< 0.0001
Fatty-acyl-ACP thioesterase	*FatB4*	33	48	15	6.9	0.03
β-Ketoacyl-ACP synthase III	*KasIII2*	40	49	13	14.6	0.0007
β-Ketoacyl-ACP synthase I	*KasI2*	40	49	13	14.6	0.0007
β-Ketoacyl-ACP synthase III	*KasIII1*	26	58	15	52.0	< 0.0001
Enoyl-ACP reductase	*Ear1*	16	54	32	5.8	0.06
β-Ketoacyl-ACP synthase II	*KasII1*	27	44	28	1.2	0.55
β-Ketoacyl-ACP synthase II	*KasII2*	17	50	33	5.2	0.07
β-Ketoacyl-ACP synthase II	*KasII3*	14	55	33	8.5	0.01

[a] V/V are *C. viscosissima* homozygotes, V/L are heterozygotes, and L/L are*C. lanceolata* homozygotes

Additional polymorphisms might have been produced by digesting monomorphic FatB, KasI, KasII, KasIII, and ACP AS-PCR markers, but this still would not have eliminated the problem of multilocus monomorphic AS-PCR alleles (AS-PCR alleles of the same length from different loci). There is no way to know the number of loci represented by an AS-PCR band without subjecting the PCR products to SSCP analysis.

SSCPs were detected using DNA fragments ranging from 200 to 800 bp. Fragments smaller than 300 bp are typically needed to detect single nucleotide differences via SSCPs (Liu and Sommer 1995). The SSCPs we found for fragments larger than this were undoubtedly enhanced using an interspecific cross and selecting sequences spanning introns. The probability of detecting sequence variants is enhanced with SSCP analysis because each locus produces two SSCP bands (complementary DNA strands), either or both of which can be polymorphic. One method for increasing polymorphism rates is to digest large (approx. 1000 bp) PCR products with restriction enzymes before SSCP analysis (Liu and Sommer 1994, 1995). This strategy can be used to scan 1 kb or more of DNA in a single gel lane and greatly increases the probability of finding polymorphisms without the trial and error testing of multiple pairs of primers spanning shorter (200–300 bp) arbitrarily chosen segments.

When gene families are amplified using highly conserved primer pairs, the AS-PCR bands can be mixtures of allelic and non-allelic DNA fragments. This problem can be avoided by separately amplifying members of a gene family using gene-specific primers when the sequences of different members of the gene family are known. Gene-specific primers and stringent PCR conditions were used to amplify individual members of the ACP and Ear gene families by aligning sequences for different genes and carefully selecting primers to amplify one member of the family.

We chose to use highly conserved primers to amplify multiple members of the FatB, KasI, KasII, and KasIII gene families. Many of the PCR products we produced for these genes were mixtures of allelic and non-allelic fragments. These fragments were separated by SSCP electrophoresis and consistently produced easily scored multilocus SSCP markers. There was only one case (KasII) where the bands for two or more loci were interleaved on the SSCP electrophorogram, but this still did not prevent accurate phenotyping because the two bands (complementary strands) of each allele were isolated and sequenced to ascertain their allelic origin (Fig. 4).

The utility of using highly conserved primers for SSCP analysis of gene families is illustrated by the FatB gene family (Fig. 2). Primers pairs p-FatB13 and p-FatB15 produced two AS-PCR bands and three pairs of SSCP bands, while primer pairs p-FatB20 and p-FatB21 produced two AS-PCR bands and four pairs of SSCP bands. AS-PCR analysis often obscures members of a gene family, while SSCP analysis permits careful simultaneous tracking of each member of a gene family.

SSCP analysis is a powerful method for candidate gene analysis – a specific SSCP marker can be unequivocally assigned to a specific cDNA or genomic sequence by isolating the allele from the SSCP gel and reamplifying and sequencing the allele as we did for several SSCP markers. This is useful for mapping or studying gene families within and between populations where RFLP probes (cDNA or genomic clones) cross-react with multiple members of the family or are not polymorphic.

Several SSCP markers can be simultaneously assayed by multiplexing when standard PCR conditions are used for different primer pairs or by pooling PCR products produced by different primers pairs when variable PCR conditions are used for each primer pair. As with any multiplexed or pooled marker assay, the multiplexed or pooled markers must be selected so that the bands (phenotypes) of different markers do not obscure one another.

Knowledge of the gene family is important, but not essential, for designing primers and developing SSCP assays. We originally developed FatB primers to simultaneously amplify three sequenced members of the FatB gene family (Fig. 2). The presence of a fourth gene came to light later and permitted us to develop a new set of highly conserved primers to simultaneously amplify all four members of the FatB gene family. The two SSCP marker assays shared 3 of the 4 FatB loci and produced polymorphisms for all four FatB loci.

Previously unsequenced members of a gene family are often amplified by using highly conserved primers. The PCR products for these genes can be the same length and thus hidden when assayed using AS-PCR. Complex monomorphic PCR products can be separated by SSCP electrophoresis, isolated from a silver-stained SSCP gel, sequenced, and compared to the original sequences to determine their origin. New members of a gene family can be found and specific loci can be assigned to specific sequences using these methods. Pseudogenes may also be amplified using conserved primers. Pseudogenes can often be identified by sequence characteristics (e.g., nonsense or stop codons in coding regions). Such loci may nonetheless be useful highly polymorphic markers.

While SSRs fill the need for random hyperpolymorphic markers, there is still a need for hyperpolymorphic markers for mapping known (candidate) genes. SSCPs typically detect a high percentage of sequence variants (Fan et al. 1993; Liu and Sommer 1995) and should be one of the more useful methods for mapping known genes (as opposed to random loci). SSCP polymorphism rates, however, are not known for most species. The polymorphism rates we report for *Cuphea*

interspecific FAS gene SSCP variants are almost certainly not representative of those within most intraspecific gene pools, e.g., the elite gene pools of many domesticated species. The utility of SSCPs (and of other sequence-based markers) partly hinges on the sequence diversity within these gene pools.

Acknowledgements Ore Agric Exp Stat Tech Paper No 10,967. This work was funded by grants from the USDA NRICGP (#91-37300-6569), USDA-ARS (#58-5114-9-1002), and Procter and Gamble Company, and an Alexander von Humbolt Fellowship to SJK.

References

Bailey NTJ (1961) Mathematical theory of genetic linkage. Clarendon, London

Bothmann H, Klein B, Voetz M, Pawlowski K, Bautor J, Schell J, Toepfer R (1993) Isolation and characterization of a gene from *Cuphea lanceolata* encoding a β-ketoacyl-[ACP] synthase I. Biol Chem Hoppe-Seyler 374: 521

Brandt T, Knapp S (1993) Gametic selection at fatty acid synthesis and allozyme marker loci and meiosis of *Cuphea lanceolata* × *Cuphea viscosissima* populations. Crop Sci 33: 1138–1143

Commission on Plant Gene Nomenclature (1994) A nomenclature for sequenced plant genes: a brief description. Plant Mol Biol Rep 12 [CPGN Suppl]: S4–S9

Crane JM, Tagliani LA, Knapp SJ (1995) Registration of four fatty acid mutant lines (VS-6-CPR-1, VS-6-CPR-4, VS-6-CPY-1, and VS-6-MYR-1) of *Cuphea viscosissima* Jacq. Crop Sci 35: 1517

Fan, E, Levin DB, Glickman BW, Logan DM (1993) Limitations in the use of SSCP analysis. Mutat Res 288: 85–92

Graham SA, Hirsinger F, Robbelen G (1981) Fatty acids of *Cuphea* (Lythraceae) seed lipids and their systematic significance. Am J Bot 68: 908–917

Holloway J, Knapp SJ (1993) Manual for G-MENDEL 3.0 genetic mapping software. Oregon State University, Corvallis, Ore.

Jones A, Davies M, Voelker TA (1995) Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell 7: 359–371

Kauppinen S (1992) Structure and expression of the *Kas12* gene encoding a β-ketoacyl-acyl carrier protein synthase I isozyme from barley. J Biol Chem 267: 23999–24006

Klein B, Toepfer R (1992) Characterization of cDNAs encoding β-ketoacyl-ACP reductases of *Cuphea lanceolata*. In: Cherif A, Miled-Daoud DB, Marzouk B, Smaoui A, Zarouk M (eds) Metabolism, structure, and utilization of plant lipids. Centre National Pedagogique, Tunis, Tunisia, pp 156–159

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

Lee H-H, Loo WX-J, Choo K-B (1992) Mutational analysis by a combined application of the multiple restriction fragment-single strand conformation polymorphism and the direct linear amplification DNA sequencing protocols. Anal Biochem 205: 289–293

Liu Q, Sommer SS (1994) Parameters affecting the sensitivities of dideoxy fingerprinting and SSCP. PCR Methods Applic 4: 97–108

Liu Q, Sommer SS (1995) Restriction endonuclease fingerprinting (REF): a sensitive method for screening mutations in long, contiguous segments of DNA. BioTechniques 18: 470–477

Martini N, Schell J, Toepfer R (1995) Expression of acyl-[ACP] thioesterase in *Cuphea lanceolata* and in transgenic rapeseed. *In*: Kader J-C, Mazliak P (eds) Plant lipid metabolism. Kluwer, Dordrecht, The Netherlands, pp 495–498

Meyers RM, Maniatis T, Lerman LS (1986) Detection and localization of single base changes by denaturing gradient gel electrophoresis. Methods Enzymol 155: 501–527

Michaud J, Brody L, Steel G, Fontaine G, Martin L, Valle D, Mitchell G (1992) Strand-separating conformational polymorphism analysis: efficacy of detection of point mutations in the human ornithine aminotransferase gene. Genomics 13: 389–394

Newman T, de Bruijn F, Green P, Keegstra K, Kende H, McIntosh L, Ohlrogge J, Raikhel N, Somerville S, Thomashow M, Retzel E, Somerville C (1994) Genes galore: summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. Plant Physiol 106: 1241–1255

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

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

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites, and machines. Trends Genetics 9: 275–280

Sanguinetti C, Neto E, Simpson A (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. Biotechniques 17: 915–919

Siggaard-Andersen M (1993) Conserved residues in condensing enzyme domains of fatty acid synthases and related sequences. Protein Seq Data Anal 5: 325–335

Slabaugh MB, Tai H, Jaworski J, Knapp SJ (1995) Two cDNAs encoding β-ketoacyl-ACP synthase III from *C. wrightii*. Plant Physiol 108: 443–444

Spinardi L, Mazars R, Theillet C (1991) Protocols for improved detection of point mutations by SSCP. Nucleic Acids Res 19: 4009

Toepfer R, Martini N (1994) Molecular cloning of cDNAs or genes encoding proteins involved in *de novo* fatty acid biosynthesis in plants. J Plant Physiol 143: 416–425

Voetz M, Klein B, Schell J, Toepfer R (1994) Three different cDNAs encoding acyl carrier proteins from *Cuphea lanceolata*. Plant Physiol 106: 785–786

von Wettstein-Knowles P, Knauf V, Ohlrogge JB, Lamppa G, Safford R, Couciet G (1994) Nomenclature for acyl carrier protein (ACP) genes. Plant Mol Biol Rep 12 [CPGN Suppl]: S31-S34

Walek J, Voetz M, Filsak E, Schell J, Toepfer R (1993) Analysis of isoforms of the enoyl-[ACP] reductase from *Cuphea lanceolata*. Biol Chem Hoppe-Seyler 374: 541

Webb DM, Knapp SJ (1990) DNA extraction from a previously recalcitrant plant genus. Plant Mol Biol Rep 8: 180–185