Development and characterization of microsatellite markers in *Cynara cardunculus* L.

Alberto Acquadro, Ezio Portis, David Lee, Paolo Donini, and Sergio Lanteri

Abstract: *Cynara cardunculus* L. is a species native to the Mediterranean basin that comprises 2 crops, globe artichoke (var. *scolymus* L.) and cultivated cardoon (var. *altilis* DC), as well as wild cardoon (var. *sylvestris* (Lamk) Fiori). Globe artichoke represents an important component of the South European agricultural economy but is also cultivated in North Africa, the Near East, South America, the United States, and China. Breeding activities and molecular marker studies have been, to date, extremely limited. Better knowledge of the genome of the species might be gained by developing a range of molecular markers. Here, we report on the development of 14 microsatellites (simple sequence repeats (SSRs)) through a novel approach that we have defined as the microsatellite amplified library (MAL). The approach represents a combination of amplified fragment length polymorphism and a primer extension based enriched library, is rapid, and requires no hybridization enrichment steps. The technique provided a ~40-fold increase in the efficiency of SSR identification compared with conventional library procedures. The developed SSRs were applied for genotyping 36 accessions of *C. cardunculus*, including a core of 27 varietal types of globe artichoke, 3 accessions of cultivated cardoon, and 6 Sicilian accessions of wild cardoon. Principal coordinates analysis made it possible to differentiate both cultivated and wild forms from each other.

Key words: globe artichoke, wild and cultivated cardoon, molecular markers, AFLP, MAL (microsatellite amplified library).

Résumé : Le *Cynara cardunculus* L. est une espèce originaire du bassin méditerranéen et comprend deux espèces cultivées : l'artichaut (var. *scolymus* L.) et le cardon (var. *altilis* DC), ainsi que le cardon sauvage (var. *sylvestris* (Lamk.) Fiori). L'artichaut constitue une composante importante de l'économie agricole du Sud de l'Europe, mais il est aussi cultivé en Afrique du Nord, au Moyen-Orient, en Amérique du Sud, aux É-U et en Chine. Les travaux d'amélioration génétique et l'emploi de marqueurs moléculaires ont été très limités jusqu'à ce jour. Une meilleure connaissance du génome de ces espèces pourrait découler du développement de marqueurs moléculaires. Les auteurs rapportent ici le développement de 14 microsatellites (SSR) via une approche originale : MAL (« microsatellite amplified library »). Cette approche fait appel à une combinaison de la technique AFLP et de la production de banques enrichies par extension d'amorces. Cette approche est rapide et ne nécessite aucune étape d'hybridation pour l'enrichissement. La technique permet d'accroître d'environ 40 fois l'efficacité d'identification de microsatellites par rapport aux approches conventionnelles fondées sur la production de banques. Les microsatellites ainsi obtenus ont été employés pour génotyper 36 accessions du *C. cardunculus* incluant 27 types variétaux d'artichauts, 3 accessions du cardon cultivé et 6 accessions siciliennes du cardon sauvage. Une analyse en coordonnées principales a permis de distinguer les deux formes cultivées et la forme sauvage les unes des autres.

Mots clés : artichaut, cardon cultivé et sauvage, marqueurs moléculaires, AFLP, MAL (« microsatellite amplified library »).

[Traduit par la Rédaction]

Introduction

Cynara cardunculus L. is a diploid ($2n = 2x = 34$), predominantly cross-pollinated species native to the Mediterranean basin. It includes 2 crops, globe artichoke (var. *scolymus* L.) and cultivated cardoon (var. *altilis* DC), as well wild cardoon (var. *sylvestris* (Lamk) Fiori]) which is a robust thistle with a characteristic rosette of large spiny leaves and branched flowering stems (Rottenberg and Zohary 1996).

Globe artichoke represents an important component of the European agricultural economy, crop production being in ex-

Received 23 May 2004. Accepted 23 November 2004. Published on the NRC Research Press Web site at http://genome.nrc.ca on 2 March 2005.

Corresponding Editor: T. Schwarzacher.

A. Acquadro, E. Portis, and S. Lanteri.[1] Di.Va.P.R.A. Plant Genetics and Breeding, University of Turin, via L. da Vinci 44, Grugliasco, Turin I-10095, Italy.
D. Lee and P. Donini. Molecular Research Group, NIAB, Huntingdon Road, Cambridge CB3 0LE, UK.

[1]Corresponding author (e-mail: sergio.lanteri@unito.it).

 doi: 10.1139/G04-111

cess of 1 million metric tons and about 120 000 ha in cultivation, while Italy is the leading producer. Although to a lesser extent, globe artichoke is also cultivated in North Africa (Egypt, Morocco, and Tunisia), the Near East (Israel, Lebanon, Syrian Arab Republic, and Turkey), South America (Argentina, Brazil, and Chile), the United States (California), and China (FAO 2004).

Italy has the richest globe artichoke primary cultivated "gene pool" and harbours many distinct clonal varietal groups. Recently, studies based on amplified fragment length polymorphism (AFLP) (Lanteri et al. 2004) and random amplified polymorphic DNA (RAPD) (Sonnante et al. 2002) have proven a good consistency between globe artichoke germplasm classification based on molecular and morphological data, suggesting that traits selected by humans play an important role in understanding variation and differentiation within the species.

Cynara cardunculus breeding activities as well as molecular marker studies have been, to date, rather limited. To gain a better knowledge of the species genome, for future applications of more appropriate and effective breeding strategies, it is of pivotal importance to develop a range of molecular markers. Microsatellite markers in plants have proven to be valuable tools for DNA genotyping, genetic mapping studies, population genetics, and conservation and management of genetic resources (Thomas and Scott 1993; Bowcock et al. 1994; Jarne and Lagoda 1996; Zane et al. 2002). Their major drawback is that they need to be isolated de novo whenever a species is investigated for the first time (Zane et al. 2002).

Conventionally, microsatellites have been isolated from genomic libraries. A number of enrichment methods have been developed to increase the rate of recovery (Zane et al. 2002) based on RAPD (Wu et al. 1994; Cifarelli et al. 1995; Richardson et al. 1995), primer extension (Fisher et al. 1996; Cooper et al. 1997; Yang et al. 2001), or selective hybridization (reviewed in Zane et al. 2002). More recently, an AFLP-based enrichment procedure, incorporating biotin–streptavidin bead capture, has been proposed (Hakki and Akkaya 2000; Hayden and Sharp 2001*a*; Hayden et al. 2002).

In a previous work (Acquadro et al. 2003), we reported on the development of 9 simple sequence repeats (SSRs) in globe artichoke of which 4 were obtained from a search in a sequence database and 5 from construction of an enriched library. Here, we report on further development of SSRs in *C. cardunculus* using an alternative approach that we have termed MAL (microsatellite amplified library). The approach is based on a two-step "primer extension" procedure and requires basic polymerase chain reaction (PCR) technology for the production of highly enriched SSR libraries; it avoids the need for both hybridization enrichment steps and conversion of the SSRs into sequence tagged microsatellite site (STMS) markers.

The developed SSRs have been applied for genotyping 36 accessions of *C. cardunculus* including cultivated and wild forms.

Materials and methods

Plant materials and genomic DNA isolation

Genomic DNA was extracted from 150 mg of frozen young leaves following the procedure of Lanteri et al. (2001). For microsatellite development, DNA from globe artichoke (varietal type 'Spinoso sardo') was used.

Thirty-six *C. cardunculus* accessions, including a panel of 27 clones of globe artichoke (var. *scolymus*) belonging to as many varietal types, 3 accessions of cultivated cardoon (var. *altilis*), and 6 Sicilian accessions of wild cardoon (var. *sylvestris*) were used for microsatellite genotyping. The globe artichoke accessions included in the study (Table 1) were found to be representative of the genetic variation within the germplasm at present in cultivation in the Mediterranean Basin (Lanteri et al. 2004). The 3 accessions of cultivated cardoon included 2 provenances of the varietal type 'Bianco avorio' and 1 of Gobbo di Nizza. The 6 accessions of wild cardoon were sampled from the wild in 3 locations in Sicily (Roccella, Bronte, and Palazzolo) and in 3 locations in Sardinia (Sassari, Nuoro, and Oristano).

Forward primer design (primary library)

Restriction ligation reactions were performed as described by Vos et al. (1995) with some modifications. Genomic DNA (0.5 µg) was restricted for 2 h at 37 °C in the presence of 2.5 U of *Mse*I (New England Biolabs, Beverly, Mass.) in 40 µL of 10 mmol Tris–HCl/L, pH 7.5, 10 mmol magnesium acetate/L, 50 mmol potassium acetate/L, and 50 ng BSA/µL. Ligation of the adapter was performed by the addition of 10 µL of ligation solution containing 1 U of T4 DNA-ligase and 1 mmol ATP/L in 10 mmol Tris–HCl/L, pH 7.5, 10 mmol magnesium acetate/L, 50 mmol potassium acetate/L, 50 ng BSA/µL, and 70 pmol of *Mse*I adapters (5′-CGTCTCGAG-GATGAGTCCTGAG-3′ and 3′-TACTCAGGACTCATCCTC-GAGACG-5′) and incubated overnight at 15 °C. Hot-start PCR was performed by adding 2 µL of the restriction–ligation reaction to produce a 20-µL PCR consisting of 15 mmol Tris–HCl/L, pH 8.0, 50 mmol KCl/L, 1.5 mmol $MgCl_2$/L, 0.2 mmol/L of each dNTP, 1 U of AmpliTaqGold® (Applied Biosystem, Warrington), 0.5 µmol/L *Mse*I adapter-directed primer (5′-CGTCTCGAGGATGAGTCCTGAG-3′), and 0.5 µmol/L of a universal SSR primer ((PCT10 = 5′-NNCTCGAG$(CT)_{10}$ or PCA10 = 5′-NNCTCGAG$(CA)_{10}$ or PGATA6 = 5′-NNCTCGAG$(GATA)_6$). Both the universal primers and the *Mse*I adapter contain an *Xho*I restriction site (CTCGAG) to facilitate subsequent cloning of the PCR products into *Sal*I-digested pUC19 vector. Because the hybrid sites formed upon ligation are not cleavable by either enzyme, the restriction–ligation steps can be carried out without any need to abolish restriction enzyme activity. PCR amplifications consisted of 45 cycles of 60 s at 94 °C, 60 s at 55 °C, and 60 s at 72 °C completed by a 10-min extension at 72 °C; 3 sets of amplification products were generated: *Mse*I-$(CT)_{10}$, *Mse*I-$(CA)_{10}$, and *Mse*I-$(GATA)_6$. The presence of a product smear (ranging from 100 to 400 bp in length) following agarose electrophoresis was indicative of a successful reaction. PCR products were purified using a Qiaquick kit (Qiagen) and restricted for 2 h at 37 °C in a 20-µL reaction containing 10 mmol Tris–HCl/L, pH 7.5, 10 mmol magnesium acetate/L, 50 mmol potassium acetate/L, and 1 U of *Xho*I. Restricted PCR fragments were size selected (>75 bp) as described elsewhere (Frohlich and Parker 2001). *Xho*I-restricted fragments were ligated in *Sal*I-digested pUC19 vector and used to transform TOP10 competent cells (Invitrogen).

Table 1. List of the 27 globe artichoke (*C. cardunculus* var. *scolymus*) accessions included in this study.

Varietal type	Origin
Bayrampasa	Turkey
Bianco tarantino	Italy
Blanco	Argentina
Camard	France
Caribou	France
Catanese	Italy
CB 642	France
EB 9	Spain
Gagliardo Sgrò	Italy
Green globe	USA
Gross Camus	France
Locale di Cuneo	Italy
Locale di Ostuni	Italy
Locale di Strancona	Italy
Mazzaferrata di Termoli	Italy
Pasquaiolo	Italy
R 35	France
Romanesco	Italy
Sakiz	Turkey
Selezione 67	Italy
Spinoso di Licata (Sp. di Gela)	Italy
Spinoso sardo	Italy
Spinoso violetto di Liguria	Italy
Terom	Italy
Testa di ferro	Italy
Tonda di Paestum	Italy
Violetto di Toscana	Italy

Three libraries, one for each SSR motif, were generated. White colonies were screened with PCR-based isolation of microsatellite arrays (Lunt et al. 1999) and positive clones were sequenced by standard methods on an ABI3100 platform (Applied Biosystem). From the sequence of each positive clone, a primer directed towards the microsatellite motif was derived (the forward primer) using BioOligo (Biogene). Each forward primer, used in combination with the corresponding universal SSR primer, generated an STM marker (Hayden et al. 2002). A nested primer (forward nested) was designed for most of the clones containing a repetitive motif (Table 2) as was used for the reverse primer design. A schematic representation of this first step is shown in Fig. 1A.

Reverse primer design

The opposite microsatellite flanking sequence was amplified using the forward primer in a 15-µL reaction using 1 µL of the restriction–ligation reaction (see above) as template. Reaction conditions were 0.1 µmol *Mse*I primer/L, 0.1 µmol forward primer/L, 15 mmol Tris–HCl/L, pH 8.0, 50 mmol KCl/L, 1.5 mmol $MgCl_2$/L, 0.2 mmol/L of each dNTP, and 1 U of FastStart *Taq*® (Roche Applied Science). Hot-Start PCR was performed with the following profile: 94 °C for 5 min followed by 35 cycles of 30 s at 94 °C, 30 s at 55 °C, and 60 s at 72 °C and ending with a 10-min extension at 72 °C. When applied, the nested PCR used, as amplification template, 1 µL of 100× dilution of the first reaction, employing the same PCR conditions. The presence of amplification products was checked by agarose electrophoresis, and where a single band was detected, it was directly sequenced. A second primer (reverse) was designed for each positive sequence. The sets of primers designed in this way are listed in Table 2. A schematic representation of the second step is shown in Fig. 1B.

STMS genotyping and statistical analysis

The SSRs developed were tested for their informativeness on 36 *C. cardunculus* accessions. PCR amplifications were carried out in a volume of 20 µL in the presence of 50 ng of template DNA, 0.2 mmol/L of each dNTP, 0.5 µmol/L of forward and reverse primers, and 1 U of *Taq* polymerase (Promega) in the manufacturer-supplied buffer. A touchdown PCR protocol was used: 94 °C for 5 min and then 11 cycles at 94 °C for 30 s, 60 °C for 30 s decreasing by 0.5 °C every cycle, and 72 °C for 60 s followed by 24 cycles at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 60 s. Amplification products were mixed with 40–50 µL of formamide dye (98% formamide, 10 mmol EDTA/L, 0.01% *w*/*v* bromophenol blue, and 0.01% *w*/*v* xylene cyanol), denatured at 95 °C for 5 min and quenched at 0 °C, separated on 6% denaturing polyacrylamide gels, and visualized with silver staining (Bassam et al. 1991).

To increase the efficiency of the SSR assay for mass genotyping of *C. cardunculus* germplasm, we evaluated the potential of multiloading PCRs; the simultaneous analysis of 2 target loci in a single run was tested by combining, in different ways, the 14 primer pairs reported in Table 2 with the 9 that we previously designed (Acquadro et al. 2003).

Identity 1.0 software (Wagner and Sefc 1999) was used to calculate expected and observed heterozygosity, null allele frequency, and the probability of finding 2 identical genotypes (probability of identity). The expected heterozygosity (H_E) was calculated as $H_E = 1 - \sum p_i^2$ (Nei 1973), where p_i represents the frequency of allele i among the varietal set. The observed heterozygosity (H_O) was obtained by direct calculation. The null allele frequency was calculated as $r = (H_E - H_O)/(1 + H_E)$ (Brookfield 1996). The probability of identity was calculated as $PI = 1 - \sum p_i^4 + \sum\sum (2p_ip_j)^2$ (Paetkau et al. 1995), where p_i and p_j represent the frequency of alleles i and j, respectively.

A binary matrix was imported into NTSYS-pc (numerical taxonomy and multivariate analysis system) version 1.80 package (Rohlf 1993) for cluster analysis. Genetic similarity among all accessions was calculated according to Dice's similarity index (Dice 1945) using the SIMQUAL (similarity of qualitative data) routine. Principal coordinates analysis was carried out to show genetic differentiation of accessions in a multidimensional space.

Results

STMS developed by MAL

Three primary libraries with a high content of repetitive sequences were produced using the *Mse*I adapter-directed primer in combination with 1 of the 3 SSR universal primers (Fig. 1A). Two of them were constructed using the dinucleotide primers PCT10 and PCA10 and a third using the primer PGATA6, which extends the application of the technique to tetranucleotide motifs. From the 3 primary li-

Table 2. Primer sequences and characteristics of the 14 conventional SSR markers generated from MAL in 27 accessions of globe artichoke (*C. cardunculus* var. *scolymus*).

Locus	Repeat motif	Primer pair sequence (5′–3′)	GenBank accession No.	T_m (°C)	Allele size range (bp)	No. of alleles	H_E	H_O	*r*	PI
CMAL-02	$(CT)_{12}$ imperfect	F: TAACCAAGAATATATCATGGAC	AY601783	62	104–134	2	0.500	1.000	–0.333	0.625
		R: AGAACCAGTTTTCATCAGAT	AY601784	60						
		Fn: ATGGACAATTCGTCTGAC	AY601783	61						
CMAL-06	$(CT)_{13}$	F: CTGCAAGGAAAAGTAATCAT	AY601785	60	144–164	3	0.427	0.593	–0.116	0.558
		R: CCGATATAGATGAACCTAGC	AY601786	60						
		Fn: AAGTAATCATTTTCACAAGGTC	AY601785	61						
CMAL-07	$(AG)_7$	F: AACCTGGTCGATGAATAATG	AY601787	55	160–184	2	0.500	1.000	–0.333	0.625
		R: GGTCTTGGTCTTCATGTCTC	AY601788	54						
CMAL-08	$(CT)_{18}$	F: GAAGCAGATAATTGAAGTGTTG	AY601791	62	150–174	2	0.500	1.000	–0.333	0.625
		R: AATCCAATTCAACAAAATCTAC	AY601792	61						
		Fn: AGTGTTGAGCAACGATATGG	AY601791	65						
CMAL-11	$(TCA)_4(GA)_5$ imperfect	F: GAAGGAGAAGCTTGATATCTG	AY601789	54	270–290	5	0.545	0.667	–0.079	0.393
		R: CATCCTCACGAGGACATC	AY601790	54						
CMAL-21	$(CA)_9(AT)_7$	F: TAAATAGTTAGTGTTCTCGTTTG	AY601793	60	130–160	4	0.581	0.519	0.039	0.407
		R: TGGGGTTGTATTGGTTG	AY601794	61						
		Fn: GTTCTCGTTTGAACTWTGATGT	AY601793	62						
CMAL-24	$(AT)_7(AC)_{10}$ imperfect	F: GCCCGTTCACACACAACA	AY601795	60	186–220	5	0.709	0.296	0.242	0.254
		R: CAGGTTCTTTTTATACAGCAG	AY601796	60						
		Fn: GTCAGCAAATTGTTAGTGAG	AY601795	59						
CMAL-25	$(AC)_{12}(AG)_3$	F: GAATCTAGTATACACAAGGTCA	AY601797	57	105–115	2	0.500	1.000	–0.333	0.625
		R: GGGTGCATAATGGTGATG	AY601798	64						
		Fn: AGGTCATACAACAACATCTAAC	AY601797	59						
CMAL-38	$(GATA)_4$ imperfect	F: TCGCATGACACCCAACTC	AY601799	55	75–95	3	0.263	0.222	0.032	0.595
		R: ACTAGAATACATACATACA	AY601800	55						
		Fn: CACCCAACTCTATCTATCTATCT	AY601799	59						
CMAL-108	$(CT)_7$	F: TAACTAACCTCTAACATTGCCA	AY601801	63	86–92	3	0.600	0.731	–0.082	0.386
		R: AGAGAGCAATCTCTAACAAAG	AY601802	59						
		Fn: TAACATTGCCACAAGAGAG	AY601801	60						
CMAL-110	$(CT)_{11}(CA)_7(CT)_{12}$	F: AGTGGGTAAGTGGGGATG	AY601803	64	80–100	5	0.479	0.111	0.249	0.348
		R: ATCTCCACATTTTCTCCTCC	AY601804	62						
		Fn: TGAGAGCGAGAGCGAGA	AY601803	65						
CMAL-117	$(CA)_8$ imperfect	F: AAGTTATGATTCATCTTATTTCA	AY601805	59	124–140	4	0.497	0.222	0.184	0.411
		R: TCCCGGTATTCTCGTAAA	AY601806	62						
		Fn: TTCATATATTTTGTTCTTGTATG	AY601805	58						
CMAL-124	$(AC)_5$ imperfect	F: AATCAATTGGTTTGTATGG	AY601807	59	100–115	2	0.388	0.526	–0.100	0.600

Table 2 (*concluded*).

Locus	Repeat motif	Primer pair sequence (5′–3′)	GenBank accession No.	T_m (°C)	Allele size range (bp)	No. of alleles	H_E	H_O	r	PI
		R: CACCGGTTACTTCCCTT	AY601808	62						
		Fn: TACTGATGTCGGGCTTGTA	AY601807	63						
CMAL-144	$(GATA)_4$ imperfect	F: CGCACACCTTTGTATTGAT	AY601809	63	165–181	2	0.500	1.000	–0.333	0.625
		R: GAACCAACCCCTAAATGTG	AY601810	64						
		Fn: TTGATCGTTAGTGTAAAATCAT	AY601809	60						

Note: F, forward primer; R, reverse primer; Fn, forward nested primers; T_m, melting temperature; H_E, expected heterozygosity; H_O, observed heterozigosity; r, frequency of null alleles; PI, probability of identity.

braries, 192 clones were screened, of which 162 (84%) were found to be positive at PCR-based isolation of microsatellite arrays screening. Among these, 30 were chosen (10 from each library) for sequencing, and forward primers were designed to generate STM markers (Fig. 1A). The second MAL technique step made it possible to convert 14 (47%) of the 30 STM markers into conventional SSRs, which consisted of 7 $(CT)_n$, 5 $(CA)_n$, and 2 $(GATA)_n$ motifs. In our experiments, the forward primer in combination with the *Mse*I primer often originated multiple amplicons, presumably because of the *Mse*I–*Mse*I residual amplification (data not shown) or as a consequence of the complexity of the *C. cardunculus* genome (Bennett and Leitch 1997). To reduce the *Mse*I–*Mse*I-amplified fragments and increase the selectivity of the reaction, we combined a hot-start with a nested PCR approach.

The presence of extra SSR motifs suggests that there are many compound SSRs as found in other plant genomes, implying that there may be clustering of microsatellites in some genomic regions (Fisher et al. 1996). The final efficiency of the procedure was estimated to be 39% (0.84 × 0.47), which is comparable with other approaches described in the literature (Zane et al. 2002).

STMS genotyping

The DNA samples obtained from the 36 *C. cardunculus* accessions were amplified at the 14 SSR loci and displayed at least 2 alleles per locus.

In globe artichoke, 44 alleles in total were detected from the 14 SSRs, which ranged from 2 to 5 per locus (Table 2) and with an average of 3.1; the number of informative alleles was 29 (65.9%). The loci with the highest allele number were CMAL-11, CMAL-24, and CMAL-110 with 5 alleles. The largest size range between alleles was 34 bp in CMAL-24. The mean expected heterozygosity was 0.499 (range 0.263–0.709), while the mean observed heterozygosity was 0.635 (range 0.111–1.000). The estimate for the frequency of null alleles was nonzero for 5 of the 14 loci. The probabilities of identity ranged from 0.254 to 0.625 and the mean value was 0.506. Thus, we estimated that the combined probability of finding 2 identical genotypes was 4.26×10^{-5}.

Within this set of microsatellites, 33 alleles were identified in cultivated cardoon (range 2–3 for each locus) and 64 in wild cardoon (range 2–10) (Table 3). The number of informative alleles was 16 (45.7%) and 56 (87.5%), respectively.

Six exclusive/distinctive alleles were amplified in globe artichoke and 26 in wild cardoon. On the other hand, no exclusive alleles were identified in cultivated cardoon, as they were shared by at least 1 other globe artichoke accession.

The scatterplot obtained from principal coordinates analysis (Fig. 2) showed genetic clustering of the globe artichoke accessions, with cultivated and wild cardoon accessions well separated. Principal coordinates 1, 2, and 3 accounted for about 56% of the total genetic variation, with each coordinate contributing 27.8%, 16.4%, and 11.8%, respectively. The first coordinate distinguished the cultivated forms (var. *scolymus* and var. *altilis*) from the wild cardoon (var. *sylvestris*) accessions, while the second co-ordinate made it

Fig. 1. Diagram of the MAL procedure. (A) First step: forward primer design; (B) second step: reverse primer design.

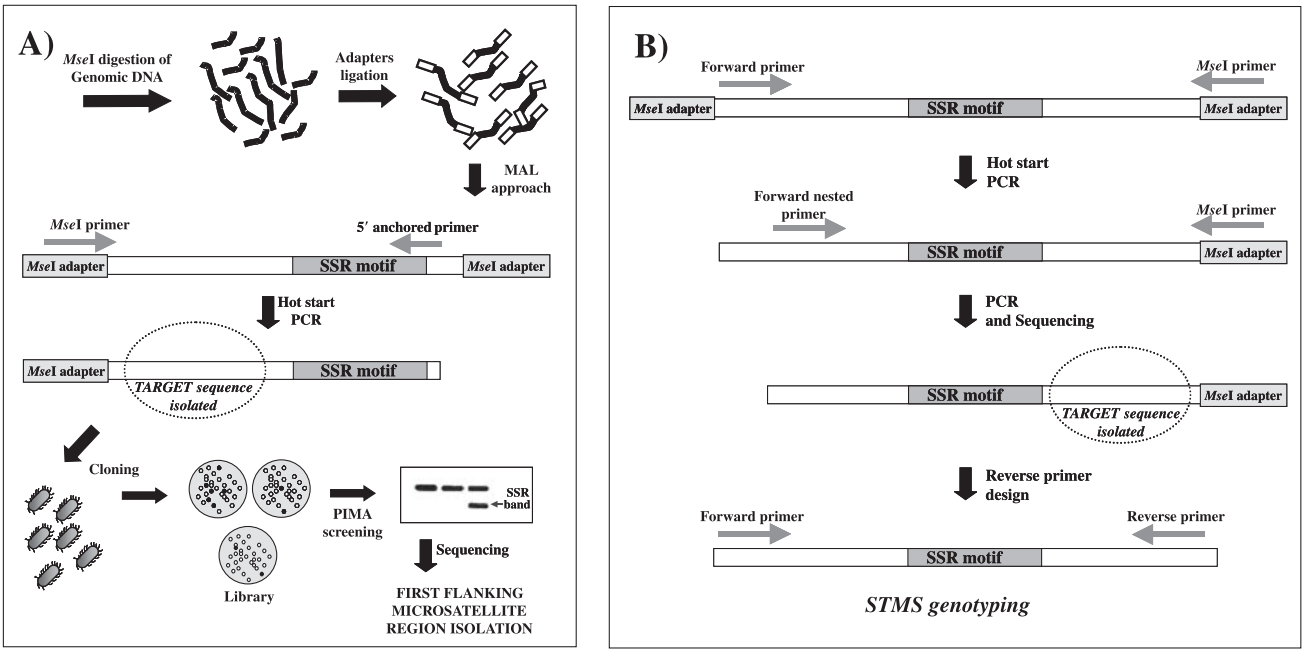


Table 3. Characteristics of the 14 SSRs when applied for genotyping 3 cultivated cardoon (*C. cardunculus* var. *altilis*) and 6 wild cardoon (*C. cardunculus* var. *sylvestris*) accessions.

	C. cardunculus var. *altilis*		*C. cardunculus* var. *sylvestris*	
Locus	Allele size range (bp)	No. of alleles	Allele size range (bp)	No. of alleles
CMAL-02	110–120	2	110–130	3
CMAL-06	144–164	3	140–180	10
CMAL-07	160–184	2	154–184	3
CMAL-08	150–174	2	150–174	6
CMAL-11	270–290	3	270–293	3
CMAL-21	150–160	2	144–170	5
CMAL-24	190–206	3	180–210	9
CMAL-25	105–115	2	105–115	2
CMAL-38	79–91	3	83–103	4
CMAL-108	86–92	3	86–104	6
CMAL-110	90–100	2	80–100	5
CMAL-117	124–130	2	124–140	4
CMAL-124	100–115	2	100–115	2
CMAL-144	165–181	2	165–181	2

possible to separate the cultivated cardoon accessions from the others.

The wild cardoon accessions clustered together, showing an average genetic similarity of 58% (Dice's similarity index range 0.51–0.66) and an average genetic differentiation of 60% from the *C. cardunculus* cultivated forms. The 3 accessions of cultivated cardoon also formed a separate cluster but showed an higher average genetic similarity (85%) and a lower genetic differentiation (44%) from globe artichoke accessions. The third cluster, formed by the 27 globe artichoke accessions, showed an average genetic similarity of 70%, and Dice's similarity index values of 1.00 (identity) were found among the 3 accessions.

Examples of multiple loading of SSR are reported in Fig. 3; of course, a fundamental requirement was that no overlap occur in product size between the individual components of the pair of SSRs.

Discussion

Here, we report on a two-step process that we have defined as MAL, which makes it possible to produce locus-specific sets of primers for the conventional SSR assay. The first advantage of the methodology lies in the fact that no hybridization enrichment step is required for designing the forward primer. The use of 5′ anchoring has been reported to target SSR length heterogeneity in PCR (Fisher et al. 1996), but we have found that unless the PCR is optimized for each primer, slippage in primer annealing results in products with the same repeat length as the primer.

Hot-start PCR was used in the initial PCR to prevent *Mse*I–*Mse*I amplification and thereby favouring products with SSRs. The *Mse*I primer was nonselective, but despite this, none of the 30 STM markers generated were redundant. Thus, the first enrichment step in MAL is both time and cost saving and produced a highly enriched (84%) primary library. The second step, performed to isolate the second flanking microsatellite region, was different from the one reported by Hayden et al. (2002), as it was based on a nested PCR approach and did not require troublesome hybridization enrichment steps. The reverse primer designing permitted the conversion of the markers, based only on the first flanking microsatellite region (i.e., MFLP (Yang et al. 2001), M-AFLP (Acquadro et al. 2002; Albertini et al. 2003), STMP (Hayden and Sharp 2001*a*), and SAM (Hayden and Sharp 2001*b*)) into STMS assays.

Fig. 2. Principal coordinates analysis based on microsatellite data depicting the genetic relationship among the 36 *C. cardunculus* accessions.

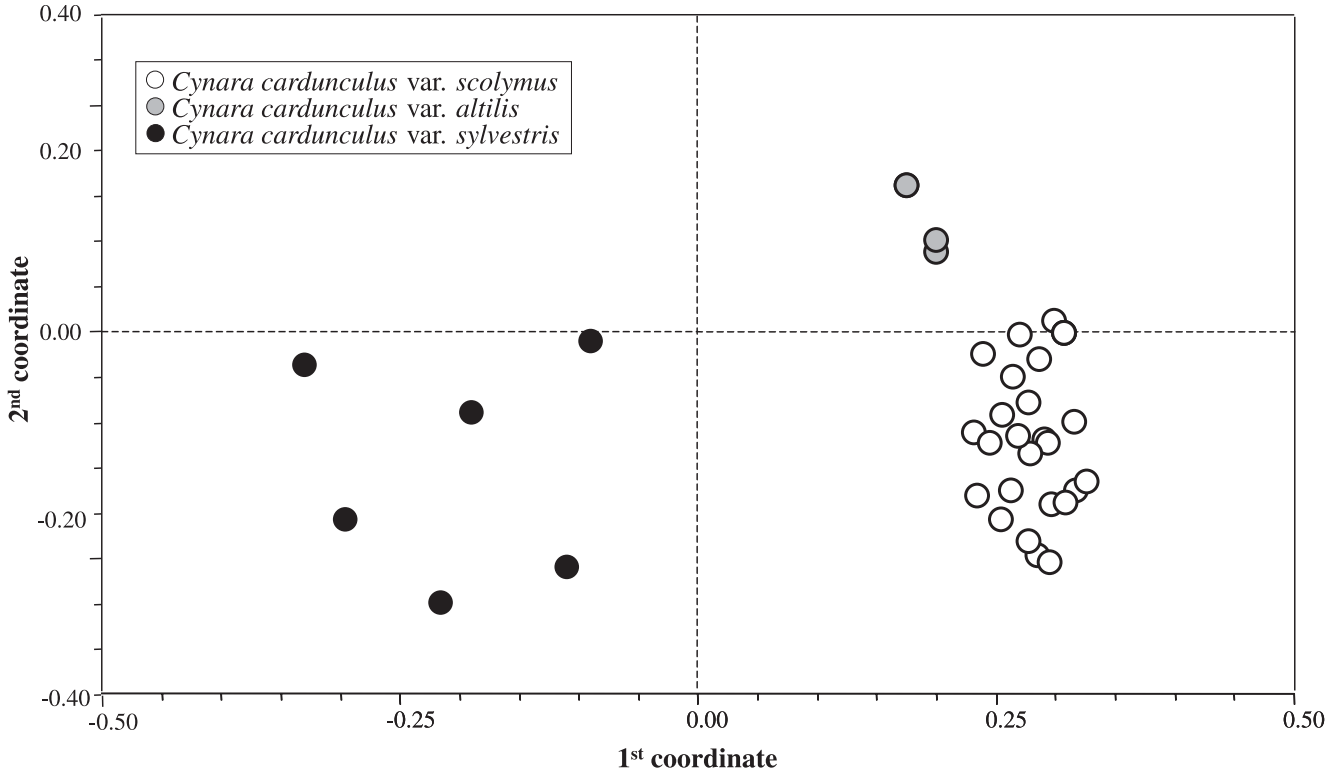
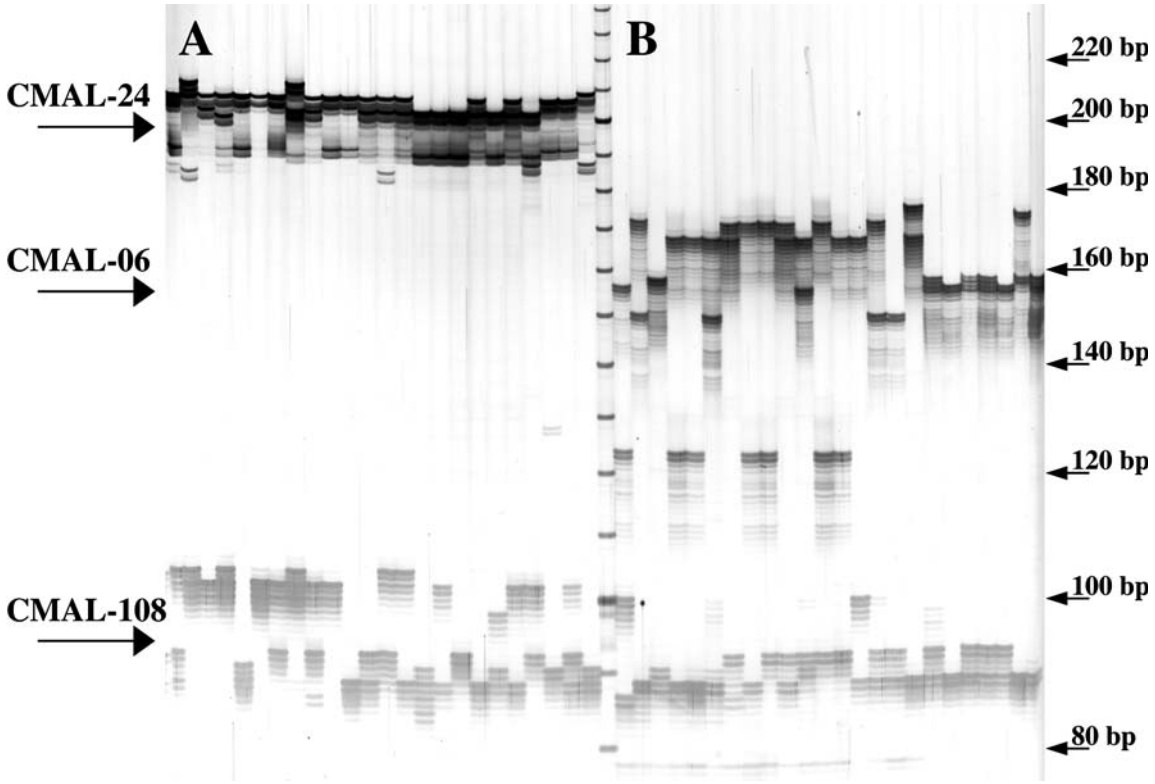


Fig. 3. Examples of SSR polymorphism in 24 accessions of wild cardoon (*C. cardunculus* var. *sylvestris*) detected by multiple loading of 2 microsatellite amplification products. (A) Multiple loading of CMAL-24 and CMAL-108; (B) multiple loading of CMAL-06 and CMAL-108.



From 15 sequences, we were able to design reverse primers, and from 14 of them, we observed amplification of the SSRs of the expected size, demonstrating their successful conversion into STMS.

The STMS markers identified were applied for genotyping wild (var. *sylvestris*) and cultivated (var. *altilis* and var. *scolymus*) forms of *C. cardunculus*. In spite of the limited number of accessions considered, the highest number of total alleles (i.e., 64) was detected in wild cardoon, which also included the highest number (i.e., 26) of exclusive alleles. A total of 44 alleles, of which only 6 were exclusive, were found in globe artichoke. Interestingly, no exclusive alleles were found in cultivated cardoon, presumably as only 3 accessions were included in the study.

The occurrence of null alleles has been pointed out as a possible problem associated with the use of microsatellite markers (Callen et al. 1993), which might result in an individual being scored as homozygote instead of heterozygous and thus resulting in loss of information. From our data, 3 loci (i.e., CMAL-24, CMAL-110, and CMAL-117) showed high positive values for the frequency of null alleles and might be regarded as loci with a possible occurrence of null alleles (O'Reilly and Wright 1995). On the other hand, the negative value for the frequency of null alleles (*r*) observed in most loci might associate the homozygotes with negative characters that have been eliminated by humans in cultivated germplasm (Marinoni et al. 2003).

Principal coordinates analysis was carried out on the mean pairwise genetic similarities and showed that wild cardoon accessions were highly differentiated both between each other and from the *C. cardunculus* cultivated forms. Yet, although to a lesser extent, the cultivated forms (var. *scolymus* and var. *altilis*) were also genetically differentiated. In a previous work based on AFLP analysis of genetic relatedness among globe artichoke, cultivated cardoon, and wild cardoon accessions, we also found that the latter clustered together and showed an average genetic differentiation of 64% from the *C. cardunculus* cultivated forms. Similarly, the cultivated cardoon accessions were differentiated from globe artichoke accessions but with a lower JSI value (Lanteri et al. 2004).

Our previous (Lanteri et al. 2004) and present molecular data, together with results obtained from cytogenetic and isozyme studies (Rottenberg and Zohary 1996), support the hypothesis that wild cardoon was the progenitor of both cultivated forms, with which it is cross-compatible with the fully fertile F_1 hybrids. Indeed, globe artichoke and wild cardoon were also found to possess ribosomal genes of the same length (Tucci and Maggini 1986; Maggini et al. 1988). *Cynara cardunculus* cultivated forms, which were probably taken into cultivation in the central or western sections of the Mediterranean Basin (Basnizki and Zohary 1994), were likely to include only a share of the huge genetic variation that is present in the wild form. Presumably, globe artichoke and cultivated cardoon evolved alongside each other as a result of different agricultural selection criteria: the cultivated cardoons being selected for the production of inner stalks and globe artichoke for the production of heads.

As previously detected by AFLP (Lanteri et al. 2004) and RAPD (Sonnante at al. 2002) markers, our microsatellites confirm the presence of genetic variation within globe artichoke cultivated germplasm as a consequence of the long process of selection for adaptation to different environments and local tastes. Indeed, the earliest clear documentation on *C. cardunculus* cultivation comes from the middle of the first century (as reported by Columella's "De Rei Rusticae", book X; Basnizky and Zohary 1994), and at present, many distinct clonal varietal types are in cultivation in the Mediterranean area.

The STMS markers developed in the present and previous works (Acquadro et al. 2003) have been shown to be highly polymorphic and suitable for multiplex analyses. They might find applications in phylogenetic and population genetics studies as well as in pedigree analysis or variety fingerprinting. Furthermore, they will contribute to genome coverage which is "conditio sine qua non" for genome mapping. Recent analyses have shown that the frequency of microsatellites is significantly higher in transcribed regions, especially in the untranslated portions, than in non-transcribed ones (Morgante et al. 2002). Indeed various studies based on SSR identification in expressed sequence tags have been reported for a wide range of species (Smulders et al. 1997; Cho et al. 2000; Huang et al. 2000; Lee et al. 2004). Our STMS markers might thus be highly efficient in tagging genes of interest for future applications of breeding programmes based on marker-assisted selection.

Acknowledgements

We thank Dr. Eva Saba and Dr. Cinzia Comino for technical assistance, Dr. Ben Vosman for technical advice on the nested PCR approach, and Dr. Emidio Albertini and Prof. Gianni Barcaccia for technical advice on setting up the MAL procedure. Special thanks to Gian Mario Mallica and the staff of the Centro Regionale Agrario Sperimentale for providing part of the material used for SSR genotyping. This work was supported by a Marie Curie Host Fellowship under the FPV of E.U. contract No. QLK5-1999-50753 and the Italian Project "Carciofo" funded by the Ministry of Forest and Agricultural Policies.

References

Acquadro, A., Portis, E., and Lanteri, S. 2002. Development of microsatellite markers in pepper (*Capsicum annuum* L.) using microsatellite-amplified fragment length polymorphism (M-AFLP). Capsicum Eggplant Newsl. **21**: 25–28.

Acquadro, A., Portis, E., and Lanteri, S. 2003. Isolation of microsatellite loci in artichoke (*Cynara cardunculus* L. var. *scolymus*). Mol. Ecol. Notes, **3**: 37–39.

Albertini, E., Porceddu, A., Marconi, G., Barcaccia, G., Pallottini, L., and Falcinelli, M. 2003. Microsatellite-AFLP for genetic mapping of complex polyploids. Genome, **46**: 824–832.

Basnizki, J., and Zohary, D. 1994. Breeding of seed planted artichoke. Plant Breed. Rev. **12**: 253–269.

Bassam, B.J., Caetano-Anolles, G., and Gresshoff, P.M. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal. Biochem. **19**: 680–683.

Bennett, M.D., and Leitch, I.J. 1997. Nuclear DNA amounts in angiosperms — 583 new estimates. Ann. Bot. (Lond.), **80**: 169–196.

Bowcock, A.M., Ruizlinares, A., Tomfohrde, J., Minch, E., Kidd, J.R., and Cavallisforza, L.L. 1994. High-resolution of human

evolutionary trees with polymorphic microsatellites. Nature (Lond.), **368**: 455–457.

Brookfield, J.F.Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Mol. Ecol. **5**: 453–455.

Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C., and Sutherland, G.R. 1993. Incidence and origin of null alleles in the $(AC)_n$ microsatellite markers. Am. J. Hum. Genet. **52**: 922–927.

Cho, Y.G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCouch, S.R., Park, W.D., Ayres, N., and Cartinhour, S. 2000. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor. Appl. Genet. **100**: 713–722.

Cifarelli, R.A., Gallitelli, M., and Cellini, F. 1995. Random amplified hybridization microsatellites (RAHM): isolation of a new class of microsatellite-containing DNA clones. Nucleic Acids Res. **23**: 3802–3803.

Cooper, S., Bull, C.M., and Gardner, M.G. 1997. Characterization of microsatellite loci from the socially monogamous lizard *Tiliqua rugosa* using a PCR-based isolation technique. Mol. Ecol. **6**: 793–795.

Dice, L.R. 1945. Measures of the amount of ecological association between species. Ecology, **26**: 297–302.

FAO. 2004. [Online]. Available from http://faostat.fao.org.

Fisher, P.J., Gardner, R.C., and Richardson, T.E. 1996. Single locus microsatellites isolated using 5′-anchored PCR. Nucleic Acids Res. **24**: 4369–4371.

Frohlich, M.W., and Parker, D.S. 2001. Benchmarks — running gels backwards to select DNA molecules larger than a minimum size. BioTechniques, **30**: 264–265.

Hakki, E.E., and Akkaya, M.S. 2000. Microsatellite isolation using amplified fragment length polymorphism markers: no cloning, no screening. Mol. Ecol. **9**: 2152–2154.

Hayden, M.J., and Sharp, P.J. 2001*a*. Sequence-tagged microsatellite profiling (STMP): a rapid technique for developing SSR markers. Nucleic Acids Res. **29**: e43.

Hayden, M.J., and Sharp, P.J. 2001*b*. Targeted development of informative microsatellite (SSR) markers. Nucleic Acids Res. **29**: e44.

Hayden, M.J., Good, G., and Sharp, P.J. 2002. Sequence tagged microsatellite profiling (STMP): improved isolation of DNA sequence flanking target SSRs. Nucleic Acids Res. **30**: e129.

Huang, S., Zhang, B., Milbourne, D., Cardle, L., Yang, G., and Guo, J. 2000. Development of pepper SSR markers from sequence databases. Euphytica, **117**: 163–167.

Jarne, P., and Lagoda, J.L. 1996. Microsatellites, from molecules to populations and back. Trends Ecol. Evol. **11**: 424 429.

Lanteri, S., Di Leo, I., Ledda, L., Mameli, M.G., and Portis, E. 2001. RAPD variation within and among populations of globe artichoke (*Cynara scolymus* L.), cv 'Spinoso sardo'. Plant Breed. **120**: 243–247.

Lanteri, S., Saba, E., Cadinu, M., Mallica, G.M., Baghino, L., and Portis, E. 2004. Amplified fragment length polymorphism for genetic diversity assessment in globe artichoke. Theor. Appl. Genet. **108**: 1534–1544.

Lee, J.M., Nahm, S.H., Kim, Y.M., and Kim, B.D. 2004. Characterization and molecular genetic mapping of microsatellite loci in pepper. Theor. Appl. Genet. **108**: 619–627.

Lunt, D.H., Hutchinson, W.F., and Carvalho, G.R. 1999. An efficient method for PCR-based isolation of microsatellite arrays (PIMA). Mol. Ecol. **8**: 891–894.

Maggini, F., Tucci, G.F., and Gelati, M.T. 1988. Ribosomal RNA genes in species of Cynareae tribe (Compositae) II. Protoplasma, **144**: 125–131.

Marinoni, D., Akkak, A., Bounous, G., Edwards, K.J., and Botta, R. 2003. Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). Mol. Breed. **11**: 127–136.

Morgante, M., Hanafey, M., and Powell, W. 2002. Microsatellites are preferentially associated with non-repetitive DNA in plant genomes. Nat. Genet. **30**: 194–200.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA, **70**: 3321–3323.

O'Reilly, P., and Wright, J.M. 1995. The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. J. Fish Biol. **47**(Suppl. A): 29–55.

Paetkau, D., Calvert, W., Stirling, I., and Strobeck, C. 1995. Microsatellite analysis of population-structure in Canadian polar bears. Mol. Ecol. **4**: 347–354.

Richardson, T., Cato, S., Ramser, J., Kahl, G., and Weising, K. 1995. Hybridization of microsatellites to RAPD: a new source of polymorphic markers. Nucleic Acids Res. **23**: 3798–3799.

Rohlf, F.J. 1993. NTSYS-pc numerical taxonomy and multivariate analysis system, version 1.80. Owner's manual. Exter Software, New York

Rottenberg, A., and Zohary, D. 1996. The wild ancestry of the cultivated artichoke. Genet. Resour. Crop Evol. **43**: 53–58.

Smulders, M.J.M., Bredemeijer, G., Rus Kortekaas, W., Arens, P., and Vosman, B. 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. Theor. Appl. Genet. **94**: 264–272.

Sonnante, G., De Polis, A., Lattanzio, V., and Perrino, P. 2002. Genetic variation in wild and cultivated artichoke revealed by RAPD markers. Genet. Resour. Crop Evol. **49**: 247–252.

Thomas, M.R., and Scott, N.S. 1993. Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). Theor. Appl. Genet. **86**: 985–990.

Tucci, G.F., and Maggini, F. 1986. Ribosomal RNA genes in species of Cynareae tribe (Compositae) I. Protoplasma, **132**: 76–84.

Vos, P., Hogers, R., Bleeker, M., Van de Lee, T., Hormes, M., Fritjers, A., Pot, J., Paleman, J., Kuiper, M., Zabeau, M. 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. **23**: 4407–4414.

Wagner, H.W., and Sefc, K.M. 1999. Identity 1.0. Centre for Applied Genetics, University of Agricultural Sciences, Vienna.

Wu, K., Jones, R., Danneberger, L., and Scolnik, P.A. 1994. Detection of microsatellite polymorphisms without cloning. Nucleic Acids Res. **22**: 3257–3258.

Yang, H., Sweetingham, M.W., Cowling, W.A., and Smith, P.M.C. 2001. DNA fingerprinting based on microsatellite-anchored fragment length polymorphisms, and isolation of sequence-specific PCR markers in lupin (*Lupinus angustifolius* L.). Mol. Breed. **7**: 203–209.

Zane, L., Bargelloni, L., and Patarnello, T. 2002. Strategies for microsatellite isolation: a review. Mol. Ecol. **11**: 1–16.