Theor Appl Genet (2002) 104:301–307

F. Carriero · G. Fontanazza · F. Cellini · G. Giorio

Identification of simple sequence repeats (SSRs) in olive (*Olea europaea* L.)

Received: 2 February 2001 / Accepted: 1 June 2001

Abstract A small insert genomic library of *Olea europaea* L., highly enriched in (GA/CT)*n* repeats, was obtained using the procedure of Kandpal et al. (1994). The sequencing of 103 clones randomly extracted from this library allowed the identification of 56 unique genomic inserts containing simple sequence repeat regions made by at least three single repeats. A sample of 20 primer pairs out of the 42 available were tested for functionality using the six olive varieties whose DNA served for library construction. All primer pairs succeeded in amplifying at least one product from the six DNA samples, and ten pairs detecting more than one allele were used for the genetic characterisation of a panel of 20 olive accessions belonging to 16 distinct varieties. A total of 57 alleles were detected among the 20 genotypes at the ten polymorphic SSR loci. The remaining primer pair allowed the amplification of a single SSR allele for all accessions plus a longer fragment for some genotypes. Considering the simple sequence repeat polymorphism, 5.7 alleles were scored on average for each of the ten SSR loci. A genetic dissimilarity matrix, based on the proportion of shared alleles among all the pair-wise combinations of genotypes, was constructed and used to disentangle the genetic relationships among varieties by means of the UPGMA clustering algorithm. Graphical representation of the results showed the presence of two distinct clusters of varieties. The first cluster grouped the varieties cultivated on the Ionian Sea coasts. The second cluster showed two subdivisions: the first sub-cluster agglomerated the varieties from some inland areas of Calabria; the second grouped the remaining varieties from Basilicata and Apulia cultivated in nearby areas. Results of cluster analysis showed a significant relationship between the multilocus genetic similarities and the geographic origin of the cultivars.

Keywords *Olea europaea* · SSR · Microsatellites · Variety genotyping

Communicated by H.C. Becker

F. Carriero · F. Cellini · G. Giorio (✉)
Metapontum Agrobios, SS Jonica Km 448.2,
75010 Metaponto (Matera), Italy
e-mail: ggiorio@agrobios.it
Fax: +39-0835-740204

F. Fontanazza
Istituto di Ricerca sull'Olivicoltura, C.N.R.,
Via della Madonna Alta, 128, 06128 Perugia, Italy

Introduction

Olea europaea L. (n=23) is the sole species of the genus *Olea* (Oleaceae) having an edible drupe: the olive. Although the taxonomy of this genus is still questioned there is a general agreement on the subdivision of *O. europaea* into two distinct subspecies: *sativa* and *oleaster*. The former includes all the cultivated olive varieties, whereas the latter are wild shrubs bearing small fruits with low pulp-oil content. Feral forms also occur near cultivated areas and are considered to be the results of crosses between cultivated and wild plants (Angiolillo et al. 1999). The cultivated olive is an evergreen long-lived tree adapted to Mediterranean climates. Since the beginning of its domestication, olive has been propagated vegetatively in order to exploit the best combination of genes which arose by random crosses or mutations. As a result, a great number of varieties are present in all the countries where this species is cultivated, raising several problems for germplam management and preservation. Until recently the identification of propagated plants has been based solely on morphological traits and the expertise of nurserymen. Moreover, very few studies have been conducted until now on the genome organisation and variation of olive and on the genetic basis of the traits affecting table olives and oil production. It has now become possible to examine genetic variation in much greater detail through the use of the large array of DNA molecular-marker types (RFLPs, RAPDs, AFLPs and SSRs; for a review see Rafalski et al. 1996; Mueller and Wolfenbarger 1999). Molecular markers can be used for the determination of the genetic relationships among va-

rieties or for the identification and mapping of loci affecting simple or complex traits (QTLs) (Soller and Beckmann, 1983). The choice of the type of molecular marker to be used is, primarily, conditioned by the objective of the study, since any type bears positive and negative aspects. In the last decade, almost all types of DNA-derived markers have been applied to olive. RAPDs have been mainly used for cultivar discrimination (Gallitelli et al. 1992; Bogani et al. 1994; Fabbri et al. 1995; Vergari et al. 1996; Wiesman et al. 1998). Genetic distance estimation between wild and cultivated olive genotypes sampled across the Mediterranean Basin has been carried out using AFLPs (Angiolillo et al. 1999) as well as using RAPDs and mitochondrial RFLPs (Besnard and Bervillé 2000). The colonization history of *O. europaea* in Macaronesia was investigated using RAPDs, ISSRs (inter-simple sequence repeats) and by sequence analysis of nuclear ribosomal internal transcribed spacer 1 (ITS-1) (Hess et al. 2000). In the present paper we report on the development of simple sequence repeat (SSR) markers in olive obtained through the screening of a (GA/CT)*n* enriched genomic library. Recently two papers have been published on the development of SSR markers in olive (Rallo et al. 2000; Sefc et al. 2000), thus raising the number of PCR-based markers available for the genetic characterisation of olive germplasm and for mapping projects.

Materials and methods

Plant material and DNA extraction

A panel of 20 accessions belonging to 16 varieties of *O. europaea* subsp. *sativa*, was used (Table 1). The accessions were sampled from the Metapontum Agrobios collection with the aim of being representative of the olive varieties with a longstanding cultivation through the Basilicata Region and along the Ionian Coasts. The Leccino and Frantoio varieties, whose traditional area of cultivation is Central Italy, were also included in the panel because their cultivation is rapidly rising in the Ionian area. Total genomic DNA was extracted from fresh leaves using the CTAB method described by Saghai-Maroof et al. (1984) with minor modifications.

Construction of the size-fractionated genomic library enriched in GA repeats

The enrichment of the genomic library for GA repeat regions was achieved by applying the protocol of Hammond et al. (1998) which is a modification of the original method proposed by Kandpal et al. (1994). Equal amounts of genomic DNA from six genotypes (Maiatica A1, Cima di Melfi A1, Ogliarola del Bradano A1, Ogliarola del Vulture, Leccino PG and Kalamata PG) were mixed to obtain 30 µg of total DNA which was digested with *Mbo*I. Digested DNA was separated on a 2% agarose gel and the region of the gel with fragments in the range of 150–700 bp was excised. DNA fragments were purified from the agarose using the Geneclean II kit (Bio101 Inc., Calif., USA), and ligated to a linker obtained by the annealing of equimolar amounts of the primers SAULA and SAULB (SAULA: 5′-GCG GTA CCC GGG AAG CTT GG-3′; SAULB: 5′-GAT CCC AAG CTT CCC GGG TAC CGC). Using 2.0 µl of the ligation reaction, the first amplification of the fragment solution was carried out using SAULA as the PCR primer. The whole PCR solution was boiled to denature DNA fragments, immediately chilled on ice and then supplemented with 5 µg of the biotinylated $(GA)_{20}$ oligonucleotide. The solution was held at 50°C for 16 h to allow hydridization between the oligonucleotide and the DNA fragments containing complementary microsatellite regions. After hybridisation the solution was mixed with Vectrex-Avidin D matrix (Vector Laboratories Inc., Calif., USA) and incubated for 40 min at room temperature on a shaking platform. During this incubation, the biotinylated oligonucleotide-genomic fragment hybrids became covalently bound to the matrix thus allowing the separation of non-hybridised DNA fragments by means of two washing steps. After that, the DNA fragments containing the microsatellite regions complementary to the probe were eluted by carrying out two washes at increasingly stringent conditions, the first at 55°C, the second at 65°C. The solution of single-stranded DNA fragments eluted at 65°C was concentrated with a Centricon-100 spin column (Millipore), and then used as a template in the second PCR which was also required to generate double-stranded DNA fragments before cloning. In order to increase the probability of cloning fragments containing simple sequence repeat regions, the enrichment procedure was reiterated using the solution obtained after the second PCR. Following the second enrichment, DNA fragments were digested with *Mbo*I to re-

Table 1 Olive accessions characterised in the study

Variety	Native area of cultivation	Abbreviation
Carolea	Calabria Region, South Italy	Carolea
Cellina Di Nardo’	Apulia Region, South Italy	CellinaN
Cima Di Melfi	Basilicata Region, South Italy	CimaM-A1
Cima Di Melfi	Basilicata Region, South Italy	CimaM-A2
Coratina	Apulia Region, South Italy	Coratina
Dolce Di Rossano	Calabria Region, South Italy	DolceRos
Frantoio	Central Italy	Frantoio
Grossa Di Cassano	Calabria Region, South Italy	GrossaCa
Kalamata	Kalámata, Greece	KalamPG
Leccino	Central Italy	LeccinPG
Leccino	Central Italy	LeccinA2
Maiatica Di Ferrandina	Basilicata Region, South Italy	Maiat-A1
Maiatica Di Ferrandina	Basilicata Region, South Italy	Maiat-A2
Ogliarola Del Bradano	Basilicata Region, South Italy	OBrad-A1
Ogliarola Del Bradano	Basilicata Region, South Italy	OBrad-A2
Ogliarola Del Vulture	Calabria Region, South Italy	OVulture
Ogliarola Leccese	Apulia Region, South Italy	OLeccese
Rotondella	Basilicata Region, South Italy	Rotondel
Sinopolese	Calabria Region, South Italy	Sinopole
Tonda	Calabria Region, South Italy	Tonda

move the linker sequences and inserted into the *Bam*HI site of the Ready-To-Go pUC18 *Bam*H I/BAP vector (Amersham Pharmacia Biotech AB, Sweden). Ligation solution was used to transform the bacterial host *Escherichia coli* DH5α. Transformed bacterial colonies carrying recombinant plasmids were selected in Petri dishes on LB medium containing 50 µg/ml of Ampicillin and 40 µg/ml of X-Gal (5-bromo-4cloro-indolyl-β-D-galactoside).

Sequencing of clones

Recombinant plasmids were prepared using the Quantum Prep Plasmid Miniprep kit (Bio-Rad Laboratories Inc. Calif., USA). The nucleotide sequences of the DNA inserts were established by using the Dye Terminator Cycle Sequencing Kit (PE Applied Biosystems, Calif., USA). DNA amplifications were carried out starting from the M13 priming sites of the vector plasmid with the PCR conditions suggested by the supplier. Polymerisation products were electrophoresed on a PE ABI 373 DNA Sequencer (PE Applied Biosystems, Calif., USA) and the sequences of the clones were obtained after the elaboration of raw data by the companion Sequencing Analysis Software program.

Sequence analysis and primer design

During the sequencing process each new sequence was compared against the local database of sequences using the Stand-alone BLAST program (BLASTN ver. 2.0.10, ftp://ncbi.nlm.nih.gov/blast/executables) in order to find duplications or redundancy. The design of primer pairs for the putative SSR loci was carried out with the Primer 3.0 program (Rozen and Skaletsky 1998) using default parameters.

DNA amplification and fragment electrophoresis

The two DNA amplifications of the enrichment phase were carried out in a final volume of 25 µl with 3 µl of the ligation reaction, 25 pmol of SAULA primer, 3 mM of each dNTP, 1 unit of *Taq* DNA polymerase (Roche Diagnostics), 2.5 µl of buffer provided with the enzyme, and H_2O to the final volume of reaction.

Functionality of the primer pairs for the selected putative SSR loci was tested on the plasmid DNA and on six genomic DNA samples of the subset of olive cultivars. PCR solutions in this case contained 20 pmol of each primer. Samples were overlaid with mineral oil and amplified in a MJR-PT100 thermal cycler programmed as follows: a step of 5 min at 72°C, followed by 30 cycles of 45 s at 94°C, 45 s at the annealing temperature of the primer (67°C for the SAULA primer; 57 or 60°C for the SSR primer pairs) and 1 min at 72°C, and a final extension at 72°C for 10 min. The amplification products were resolved by electrophoresis in a 1.8% agarose gel in TAE buffers and revealed under UV illumination by ethidium bromide staining.

Allele detection and polymorphism evaluation

Eleven primer pairs were assayed for their ability to detect polymorphism among the 20 accessions of *O. europaea*. PCR reactions were carried out as described above, except that the nucleotide mix contained 0.5 µl of [^{35}S] dCTPαS (Ci/mmol >1000, Amersham Pharmacia Biotech AB, Sweden). The amplification products were separated on 6% denaturing polyacrylamide gels (Long Ranger, FMC BioProducts). The gels, after drying, were used to expose BIOMAX MR film (Kodak) using intensifying screens at –80°C until development.

Data analysis

The genetic similarity between the multiple-locus genotype of two accessions was estimated by means of the *Ps* index (Bowcock et al. 1994), which calculates the proportion of shared alleles between two accessions averaged over *n* loci [*Ps*=(shared alleles/2*n*)]. The complement to one, (1-*Ps*), provided the cognate dissimilarity measure. The genetic dissimilarity matrix was estimated using the program MICROSAT 2.0 (http://hpgl.stanford.edu/microsat/) with exhaustive bootstrapping. To facilitate the interpretation of the dissimilarity matrix a tree matrix was estimated using the sequential, agglomerative, hierarchical and non-overlapping clustering algorithm UPGMA (unweighted pair-group method using an arithmetic average) (Sneath and Sokal 1973) with the NTSYS 1.8 package.

Results

Simple sequence repeat clone recovery and characterisation

The enrichment procedure enabled the characterisation of 56 unique SSR loci out of the 103 genomic clones sequenced.

The enrichment of the library was verified in a preliminary step by comparing three different screening methods. Ten randomly chosen clones were sequenced and screened for the presence of microsatellite repeat regions by means of hybridisation with a synthetic repeat $(GA)_{20}$ as well as the amplification protocol of Morgante et al. (1997). This control showed that the hybridisation method give rise to a very high rate of false positives while the PCR selection method produce too high a rate of false negatives. By comparing the costs and time involved in the three procedures it was decided to proceed to the direct sequencing of the clones extracted at random from the enriched library. From a total of 103 sequenced clones, 79 (nearly 77%) contained microsatellite motif regions constituted by at least three single repeats. The 24 residual clones all contained stretches of G(C) and A(T) with irregular fashions, such that they could not be classified as microsatellite regions. During the sequencing process any new obtained sequence was compared against the local sequence database using the Stand-alone BLAST program in order to find duplications or redundancy. As a result of this constant control, it was possible to detect 23 clones, out of the 79 containing microsatellites, that were identical though often cloned in opposite direction to previously sequenced clones. In two of these cases, two different alleles from the same SSR loci were sequenced.

Therefore, out of the 79 selected clones, 56 were unique thus giving rise to 54% as the frequency of recovered SSR-containing clones out of 100 randomly chosen for sequencing. Among the 24 sequences not containing simple sequence repeats, only two were duplicated, 15 were unique and the last seven constituted a cluster of sequences with strong homology to the *Sau*3AI 81-bp repeat family (Katsiotis et al. 1998; Bitonti et al. 1999). The frequency distribution of the size of the inserts proved to be quite normal with the sample mean equal to 256 and the mean standard error equal to 4.83 base pairs (bp). The very short length of the inserts, realised after the screening control, prompted us to perform a single

sequencing run for each clone using the M13Forward primer. Complementary strand sequencing reactions were carried out only in the case of questionable sequence records.

As concerning the type of sequence repeat motifs (Morgante and Vogel 1999), our microsatellite database contained simple perfect and simple imperfect microsatellites constituted by di- and tri-nucleotides, and a single clone classified as a perfect compound. Although the enriched library was constructed using a synthetic $(GA)_{20}$ oligonucleotide, several trinucleotide-based repeats were also found with (CCT)*n* and (TTC)*n* motifs. The most numerous class was the *simple perfect dinucleotide* with 22 (39%) occurrences, followed by the *simple imperfect dinucleotide* with 15 (27%) clones, and the simple perfect trinucleotide with five (9%) SSRs. The remaining 12 (21%) clones were assigned to the class *multiple* which included the sequences that contained more than one microsatellite repeat region at a distance longer than three nucleotides. Among the simple microsatellites (perfect and imperfect), the number of repeats varied between 3 and 31. Only three SSRs had more than 20 repeats, five contained between 10 and 20 repeats, while 22 sequences had between 4 and 10 repeats. The last 12 clones, three of them being of the trinucleotide type, had only three repeats.

Simple sequence repeat length polymorphism

The nucleotide sequences of the 44 clones containing SSRs with more than four repeats were submitted to the Primer 3.0 program (Rozen and Skaletsky 1998) for the design of the primer pairs. In only two cases, the microsatellite was too close to one of the two ends thus preventing the design of a potentially functional primer. Twenty primer pairs out of the 42 available were tested for functionality using a panel of six *O. europaea* varieties. In all cases, an amplified fragment of the expected size was obtained for all six genotypes as well as the control clone. Microsatellite motifs, primer pair nucleotide sequences and the number of the alleles detected for the 20 SSR loci are summarised in Table 2. Following this pre-screening, 11 primer pairs were selected and tested for their ability to disclose simple sequence length polymorphism at the corresponding SSR loci with the panel of the 20 accessions by PAGE analysis of radioactively labelled PCR products.

Ten primer pairs, out of the 11 tested, were able to show polymorphism among the 20 olive genotypes. The number of alleles detected ranged between nine for locus GAPU101 and three for locus GAPU11. The primer pair specific for fragment GAPU113 allowed the amplification of the expected fragment for all genotypes. Moreover, an additional fragment of greater size, probably the result of a duplication, was visualised for same genotypes. The locus GAPU12 exhibited two different alleles, the clone allele and a longer one specific for the Ogliarola Leccese variety. However, in eight out of the

Table 2 Olive SSR loci characterised in the study

Locus	Repeat sequence	Alleles detected	Size of sequenced allele (bp)	Primer sequences (5′ to 3′)	
				Forward	Reverse
GAPU101	$(GA)_8(G)_3(AG)_3$	9	264	CATGAAAGGAGGGGGACATA	GGCACTTGTTGTGCAGATTG
GAPU103 A	$(TC)_{26}$	8	245	TGAATTTAACTTTAAACCCACACA	GCATCGCTCGATTTTTATCC
GAPU47	$(CT)_{16}$	8	193	GATCAGCTTAGTCTCATATTCTCTCTC	CCTCGACTGATTTACACACCA
GAPU89	$(AG)_{16}(G)_3(GA)_9$	8	215	GATCATTCCACACACGAGAG	AACACATGCCCACAAACTGA
GAPU71B	$GA(AG)_6(AAG)_8$	6	285	GATCAAAGGAAGAAGGGGATAAA	ACAACAAATCCGTACGCTTG
GAPU71 A	$(AG)_{10}$	5	271	GATCATTTAAAATATTAGAGAGAGAGA	TCCATCCATGCTGAACTT
GAPU45	$(AG)_7$	4	292	ATCGGGAGGGATGTGATGTA	CATCGCATCGCCTGTAAATA
GAPU59	$(CT)_9$	4	227	CCCTGCTTTGGTCTTGCTAA	CAAAGGTGCACTTTCTCTCG
GAPU11e17	$(CT)_2(TT)(CT)_7$	3	230	CGCGTTACCATACCTTAGCC	TTGAATCTGACGTGGATGGA
GAPU12	$(AG)_{10}$	2	203	TTTTGCAAAGCATTAGAGCA	AGCAGCAGCAGAAGCTTGAT
GAPU113	$CTT(CT)_{14}$	1	300	TGAATTCCTTAGGCATCTTGC	GGGAAGCTTGGGACATTAAA
GAPU14	$(CT)_2C(CT)_5$	1[a]	220	CACGCCAAGTCACTTTTCAA	CCCAGTAGCATGTTGTGAGC
GAPU19	$(CCT)_9$	1[a]	247	GATCAGTGTACTACGGTTC	TCTGTCACAACTGCGGTA
GAPU72	$(TC)_5C(CT)_4$	>1[a]	267	GAGGCTTTTTAATCCGAGCA	AAAAAGAGGGGAGGAGAGAG
GAPU62	$(CT)_4$	1[a]	252	GATCACGAATCCCCAAATAA	TGCGTTCCTGTATAATTGCATC
GAPU82	$(AG)_5TC(AG)_3$	1[a]	292	TGAATCAACCCGTCAATAAGG	TGCTATTTGCACATCATTGTTT
GAPU90	$(CT)_8GTCT$	1[a]	257	GCTGAGCAGCGAAAAATGAT	GCGACATATCTCTATGAGCAAGAA
GAPU92	$(CT)_4ATCTG(TC)_6$	1[a]	269	ATTGAGCGGCTCCTCAGTTA	TGCAACAAGCTATAACGCAAA
GAPU108	$(GT)_5$-$(GA)_5$-$(GA)_4$	>1[a]	247	GATCCTTAGAGGATTCAATGAGAA	GCAAGTCCACCATCTTCAGAC
GAPU95	$(CTT)_4$-$(GA)_3(AG)_5$	>1[a]	253	TCTCTTCTCTTCTCCAGTTTCC	ATTTGAAAGATACACATTAACTCAGG

[a] Primer pair tested on six genotypes; PCR products resolved on ethidium bromide-stained agarose gel

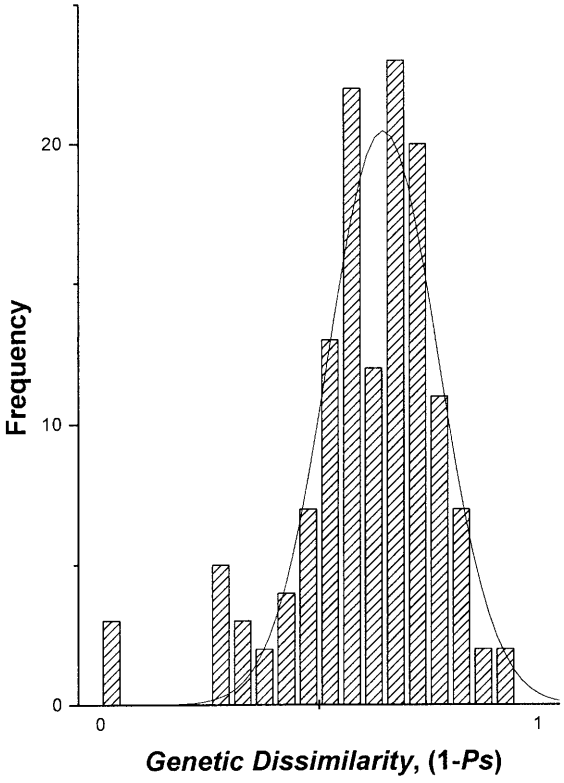

Fig. 1 Frequency distribution of genetic dissimilarity relative to all pair-wise combinations among the 20 olive accessions

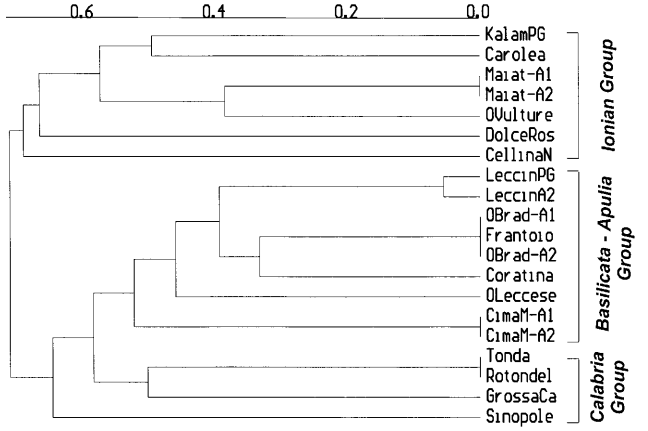

Fig. 2 Phenogram of the 20 olive accessions based on the UPGMA clustering carried out on the (1-*Ps*) dissimilarity matrix

20 genotypes no amplification fragment was produced, indicating a possible sequence divergence in the complementary nucleotide sequence of one of the two primers. By taking into account the ten polymorphic SSR loci, the characterisation of the variety panel allowed the detection of 57 alleles, an average of 5.7 alleles per locus. However, further data analysis was conducted excluding locus GAPU12 due to the inability to resolve the allelic state for genotypes which showed no amplification.

Variety genetic relationships

Allelic characterisation of the nine SSR loci allowed the analysis of the genetic relationships among the 20 olive genotypes. The complement to one of the proportion of shared fragments *Ps* (Bowcock et al. 1994) among two accessions was used as a measure of their genetic dissimilarity.

The frequency distribution of the genetic dissimilarities is presented in Fig. 1. Since the two distinct accessions each of the Cima di Melfi, Ogliarola del Bradano and Maiatica varieties were scored identically, they were not included in the estimation of sample statistics. Genetic dissimilarity values ranged between 0.914 for the pair Carolea–Cima di Melfi, to 0 for the pairs Frantoio–Ogliarola del Bradano and Tonda–Rotondella, while the sample mean was 0.61.

The genetic dissimilarity matrix thus estimated was used to obtain a graph matrix by means of the UPGMA agglomerative algorithm of clustering. The graphical visualisation of the results of this analysis is shown in Fig. 2. Two distinct clusters of olive varieties were clearly recognisable with the exception of Sinopolese, which can be considered a true outlier. In particular, the first cluster included the Greek variety Kalamata and the varieties cultivated along the Coasts of the Ionian Sea (Carolea and Dolce di Rossano from Calabria; Maiatica and Ogliarola del Vulture from Basilicata; Cellina di Nardò from the Apulian shore of Taranto Gulf). From the analysis of the second cluster it is possible to distinguish two sub-groups. The first included the varieties originating from inland areas of Calabria (Grossa di Cassano, Tonda, Rotondella) while the second grouped together the genotypes cultivated in the North of Basilicata (Ogliarola del Bradano, Cima di Melfi,) and inland areas of Apulia (Ogliarola leccese and Coratina) as well as the Sinopolese. Unexpectedly, the cultivars Frantoio and Leccino, that were alleged to be outsiders since their cultivation area is the Central of Italy, were also clustered in this group. Moreover, Ogliarola del Bradano and Frantoio showed the same multilocus allelic profile. This latter result demonstrates that, though the two genotypes have a longstanding tradition of cultivation in two distant area of Italy, they are probably two clones of the same variety. This result is also supported by the fact that Frantoio and Ogliarola del Bradano are almost indistinguishable at the morphological level.

Discussion and conclusions

A small insert genomic library of *O. europaea* highly enriched in (GA/CT)*n* repeats has been obtained using the procedure of Kandpal et al. (1994). A total of 56 unique genomic inserts, containing simple sequence repeat regions with at least three single repeats, were identified after the sequencing of 103 clones extracted at random

from the library. A sample of 20 primer pairs, out of the 42 available, was tested for functionality on the DNA of six olive varieties. All primer pairs succeeded in amplifying all DNA samples and the ten pairs detecting more than one allele were used for the genetic characterisation of 20 accessions belonging to 16 distinct olive varieties. A total of 57 distinct alleles at the ten SSR loci were detected with that panel, which means than 5.7 alleles were scored on average for each locus. A genetic dissimilarity matrix based on the proportion of shared alleles among all the pair-wise combinations of genotypes was constructed and used to clarify the genetic relationships among varieties by means of the UPGMA clustering algorithm.

The enrichment procedure proved to work very efficiently since 54% of the clones sequenced contained simple sequence repeat regions made up of at least three repeating units. This percentage was calculated excluding duplications that accounted for 22% of all sequenced clones. These results are quite similar to those obtained by Rallo et al. (2000) in olive and by others who applied similar enrichment procedures (Brondani et al. 1998; Cipriani et al. 1999; Gianfranceschi et al. 1998). However, one notable difference is that we did not screen the library by means of hybridisation with (GA)*n* (Rallo et al. 2000) or by the anchored PCR assay because we found this step to be time-consuming and to suffer too high a rate of false positives. With respect to the type of sequence-repeat motifs, the simple perfect dinucleotide class was the most frequent (39%) followed by the imperfect cognate (27%). This is in contrast with what was obtained by Soranzo et al. (1998) in *Pinus sylvestris*, Cipriani et al. (1999) in Peach and Huang et al. (1998) in *Actinidia*, who found that all the (GA)*n* microsatellites were perfect. This discrepancy is probably due to the different stringency conditions applied at the elution step during the enrichment procedure. This factor can also explain why we obtained quite a high number of clones classified as multiple type, presenting more than one SSR region, often of different type, within the nucleotide sequence.

Phenetic analysis of the genetic polymorphism carried out by means of the UPGMA clustering algorithm helped to clarify the genetic relationships among 20 accessions belonging to 16 olive varieties. The phenogram showed the presence of two main clusters, the second of which seemed to be subdivided according to the native areas of differentiation for the genotypes.

The first cluster included the Greek variety Kalamata and the varieties Carolea, Dolce di Rossano, Maiatica and Cellina di Nardò, all presumably which originated along the Coasts of the Ionian Sea. The only exception seemed, at first glance, to be Ogliarola del Vulture which is cultivated in the North-East of Basilicata. However, a more-accurate investigation of its origin showed that this variety was imported from Calabria during the 18th century.

The second cluster included the varieties originating from the North of Basilicata (Ogliarola del Bradano and Cima di Melfi), from inland places of Calabria (Grossa di Cassano, Tonda and Rotondella), Apulia (Ogliarola leccese and Coratina), and the two outsiders Frantoio and Leccino and the Sinopolese variety.

The placement of cv Frantoio and Leccino among the varieties from Basilicata and Apulia stands in contrast to historical beliefs on the origin of these cultivars. This evidence, together with the demonstration that Frantoio and Ogliarola del Bradano are probably two clones of the same variety, casts a shadow on the hypothesis that Benedictine monks, establishing their Monastery in the Montescaglioso City in the early 1300s, had brought the Frantoio variety from Central Italy. Our results do not ruled out the hypothesis that Frantoio could be a derivative of Ogliarola del Bradano. Since the demonstration of this hypothesis could be of major interest to the olive community, additional analyses are in progress to confirm our preliminary results.

Acknowledgements The authors thank Prof. Mike Kearsey (Birmingham University, UK) and Dr. Stephan Summerer (Metapontum Agrobios) for reviewing the manuscript. This work was partially financed by a joint MURST/CNR Italian program, through Istituto Tossine e Micotossine da Parassiti Vegetali, within the FERS Project "Promozione di interventi per la migliore utilizzazione delle risorse nel settore della olivicoltura nelle Regioni Meridionali." All the experiments reported in this paper were conducted in Italy and comply with all current national laws.

References

Angiolillo A, Mencuccini M, Baldoni L (1999) Olive genetic diversity assessed using amplified fragment length polymorphisms. Theor Appl Genet 98:411–421

Besnard G, Bervillé A (2000) Multiple origins for Mediterranean olive (*Olea europaea* L. ssp. *europaea*) based upon mitochondrial DNA polymorphisms. Life Sci 323:173–181

Bogani P, Cavalieri D, Petruccelli R, Polsinelli L, Roselli G (1994) Identification of olive tree cultivars by using random amplified polymorphic DNA. Acta Hort 356:89–101

Bitonti MB, Cozza R, Chiappetta A, Contento A, Minelli S, Ceccarelli M, Gelati MT, Maggini F, Baldoni L, Cionini PG (1999) Amount and organization of the heterochromatin in *Olea europaea* and related species. Heredity 83:188–195

Bowcock A. M, Ruíz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution human evolutionary trees with polymorphic microsatellites. Nature 368:455–457

Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla.* Theor Appl Genet 97:816–827

Cipriani G, Lot G, Huang W–G, Marrazzo MT, Peterlunger E, Testolin R (1999) AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterisation and cross-species amplification in *Prunus*. Theor Appl Genet 99:65–72

Fabbri A, Hormaza JI, Polito VS (1995) Random amplified polymorphic DNA analysis of olive (*Olea europaea* L.) cultivars. J Am Soc Hort Sci 120:538–542

Gallitelli M, Semeraro L, Giorio G (1992) Valutazione del grado di variabilità genetica tra 15 cultivar di olivo (*Olea europaea* L.) con l'uso di marcatori molecolari (RAPDs). XXXVI Italian Society for Agricultural Genetics National Congress, Metaponto, Italy, 5–8 October, Abstract 84

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998) Simple sequence repeats for the genetic analysis of apple. Theor Appl Genet 96:1069–1076

Hammond RL, Saccheri IJ, Ciofi C, Coote T, Funk SM, Mcmillan WO, Bayes M K, Taylor E, Bruford MW (1998) Isolation of

microsatellite markers in animals. In: Karp A, Isaac PG, Ingram DS (eds) Molecular tools for screening biodiversity. Chapman and Hall, London, pp 279–285

Hess J, Kadereit JW, Vargas P (2000) The colonization history of *Olea europaea* L. in Macaronesia based on internal transcribed spacer 1 (ITS-1) sequences, randomly amplified polymorphic DNAs (RAPD), and intersimple sequence repeats (ISSR). Mol Ecol 9:857–868

Huang W–G, Cipriani G, Morgante M, Testolin R (1998) Microsatellite DNA in *Actinidia chinensis*: isolation characterisation, and homology in related species. Theor Appl Genet 97:1269–1278

Kandpal RP, Kandpal G, Weissman SM (1994) Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. Proc Natl Acad Sci USA 91:88–92

Katsiotis A, Hagidimitriou M, Douka A, Hatzopoulos P (1998) Genomic organization, sequence interrelationship, and physical localization using in situ hybridization of two tandemly repeated DNA sequences in the genus *Olea*. Genome 41:527–534

Morgante M, Vogel M (1999) Compound microsatellite primers for the detection of genetic polymorphisms. U.S. Patent and Trademark Office, PN: 5,955,276

Morgante M, Pfeiffer A, Jurman I, Paglia G, Olivieri AM (1997) Isolation of microsatellite markers in plants. In: Karp A, Isaac PG, Ingram DS (eds) Molecular tools for screening biodiversity. Chapman and Hall, London, pp 288–295

Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. TREE 14:389–394

Rafalski JA, Vogel JM, Morgante M, Powell W, Andre C, Tingey SV (1996) Generating and using DNA markers in plants. In: Birren B, Lai E (eds) Nonmammalian genomic analysis. A practical guide. Academic Press, San Diego, pp 75–134

Rallo P, Dorado G, Martin A (2000) Development of simple sequence repeats (SSRs) in olive tree (*Olea europaea* L.). Theor Appl Genet 101:984–989

Rozen S, Skaletsky HJ (1998) Primer3. Code available at http://www–genome.wi.mit.edu/genome_software/other/primer3.html.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81:8014–8019

Sefc KM, Lopes MS, Mendonca D, Rodrigues dos Santos M, Laimer da Camara Machado M, da Camara Machado A (2000) Identification of microsatellite loci in olive (*Olea europaea*) and their characterization in Italian and Iberian olive trees. Mol Ecol 9:1171–1173

Sneath PHA, Sokal RR (1973) Numerical taxonomy. Freeman, San Francisco

Soller M, Beckmann JS (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor Appl Genet 67:25–33

Soranzo N, Provan J, Powell W (1998) Characterization of microsatellite loci in *Pinus sylvestris* L. Mol Ecol 7:1247–1263

Vergari G, Patumi M, Fontanazza G (1996) Use of RAPDs markers in the characterisation of olive germplasm. Olivae 60:19–22

Wiesman Z, Avidan N, Lavee S, Quebedeaux B (1998) Molecular characterization of common olive varieties in Israel and the West Bank using randomly amplified polymorphic DNA (RAPD) markers. J Am Soc Hort Sci 123:837–841 reflist>