

Microsatellite markers to identify specific alleles in DNA extracted from monovarietal virgin olive oils

Vittorio Alba · Wilma Sabetta · Antonio Blanco · Antonella Pasqualone · Cinzia Montemurro

Received: 17 October 2008 / Revised: 2 March 2009 / Accepted: 15 March 2009 / Published online: 11 April 2009
© Springer-Verlag 2009

Abstract The suitability of DNA present in olive oil for PCR analysis has been reported by several authors. However, low concentration, degradation, and the possible presence of additional alleles due to paternal contribution in oils extracted from entire drupes, should be taken into consideration when comparing the amplification profiles of leaves with the corresponding oils for varietal traceability purposes. The aim of this work was to assess, by capillary electrophoresis of microsatellite markers, the phenetic relationships among seven certified *Olea europaea* L. cultivars, and to verify the genomic equality between leaf DNA and the corresponding monovarietal oil DNA. Moreover, the aim was to establish an identification key to distinguish all the types of oil among them with the minimum number of markers. The results referred to oil DNA, obtained in 70 PCR experiments, confirmed the possibility, in 85.7% of cases, of extracting DNA suitable for the analysis of microsatellites. Ninety percent of the successful amplifications led to identical patterns for leaves and oil DNA. Cima di Melfi shared only 20% of the alleles with the other cultivars. Toscanina and Leccino, showed the highest similarity (about 60%). A single microsatellite was able to distinguish the seven types of oil.

Keywords SSR markers · Monovarietal olive oil · Traceability

Introduction

Olive (*Olea europaea* L.) is one of the most ancient crops typical of the Mediterranean basin. Olive oil consumption is increasing throughout the entire world, especially due to its beneficial health effects [1]. Depending on the acidity level, virgin olive oils may be classified as extra-virgin (≤ 0.8), virgin (≤ 2.0), or “lampante” (≥ 2.0) [2]. Among them, extra-virgin category shows the highest quality from the chemical and sensory point of view, with a total absence of defects. In particular, many extra-virgin olive oils, because of their healthiness and typicality, have obtained marks of protected designation of origin (PDO) at a European level according to EC Regulation 2081/92. Thus, the concept of “typicality” implies excellent and peculiar sensory features which determine a higher economic value to PDO products with respect to non-PDO foodstuffs. Consequently, certification of the origin of raw materials and industrial processes to obtain the final product is of primary importance for the protection of consumers, in particular for fraud prevention [3].

The fundamental goal is to accurately identify and characterize the numerous commercial cultivars existing in Italy, and more generally, all over the Mediterranean basin. Because morphological traits and chemical analyses are not sufficient to verify olive oil varietal authenticity, DNA markers, which have already been used to identify olive cultivars, are increasingly being applied to solve traceability and provenance issues [4]. Microsatellites, or simple sequence repeats (SSR), are among the most suitable

V. Alba · W. Sabetta · A. Blanco · C. Montemurro (✉)
Section of Genetics and Breeding,
Department of Agro-forestry and Environmental
Biology and Chemistry, University of Bari,
via Amendola 165/a, 70126 Bari, Italy
e-mail: c.montemurro@agr.uniba.it

A. Pasqualone
Section of Agro-food Industries, Department of Management
of Agro-zootechnic and Forestry Systems, University of Bari,
via Amendola 165/a, 70126 Bari, Italy

markers since they are characterized by a high polymorphism level, due to variations of the number of repeats [5].

Significant amounts of DNA are present in olive oil obtained solely by mechanical treatments [1], and its suitability for PCR analysis has been reported by several authors [6–9]. However, low oil DNA concentrations and nuclease degradation, could cause the lack of amplification of some alleles, and in some cases additional alleles due to paternal contribution in oils extracted from entire drupes [4]. These issues need to be taken into account when comparing the profiles of olive leaves with the corresponding oil for varietal traceability purposes.

The recent use of fluorescent labeled primer pairs in the amplification protocols let the PCR products to be analyzed on automatic sequencers, which are known to be less time-consuming than classical electrophoretic methods, i.e. by agarose or polyacrylamide gels, and have minor health implications for operators. Besides, the use of automatic sequencer is more effective, because capillary electrophoresis allows to distinguish alleles with very small differences in molecular weight, and has a higher resolutive power than classical methods. Moreover, its sensibility allows to evidence also weak signals, such as those that could come from olive oil DNA, known to be partially degraded [3, 10].

Finally, another problem in the characterization and identification of olive cultivars is the lack of official standard varieties, as well as the presence of different clones within the same cultivar (multiclonal populations). On the other hand, olive is submitted to certain procedures for plant material marketing and has to satisfy the requirements of the Community Directives n. 92/34/EEC of 28 April 1992 and n. 93/48/EEC of 23 June 1993 about the conditions to be met by fruit plant propagating material and fruit plants intended for fruit production [11, 12]. The use of certified material should be taken in account to validate the results when setting up varietal traceability methods.

The goal of this research was to verify, by capillary electrophoresis of microsatellite markers, the phenetic relationships among seven certified *O. europaea* cultivars, and to verify the genomic equality between leaf DNA and the corresponding monovarietal oil DNA. All the cultivars chosen were included in the list of starting materials to obtain some Italian PDO oils. Another aim was to establish an identification key with the minimum number of markers to distinguish all the cultivars.

Materials and methods

Plant material and DNA extraction

Fresh leaves of seven cultivars (Coratina, Picholine, Toscanina, Cima di Melfi, Frantoio, Leccino, Cellina di

Nardò)—diffused in Italy and included in the process to obtain PDO recognition—were sampled from the Olive Pre-multiplication Centre field “Concadoro” at Palagiano (Taranto, Italy). The genetic correspondence and sanitary state of the olive trees were certified, according to the requirements of the Dir. 92/34/EEC of 28 April 1992 [11] and Dir. 93/48/EEC of 23 June 1993 [12]. The field was spatially isolated respect to other olive cultivation, in order to limit contamination caused by pollen flow and fly vectors only to cultivars present in the same field. For each clone, a total amount of 3 kg of drupes were collected to obtain monovarietal oils. Oil extraction was performed with an experimental oil mill almost similar to those employed for industrial purposes, the details of which are reported in previous papers [3, 8].

Genomic DNA from fresh leaves was extracted according to Doyle and Doyle protocol [13]. DNA was also extracted by means of Gene Elute Plant Kit (Sigma, St. Louis, MO) from cellular residuals obtained by centrifuging 250 ml of each monovarietal oil at 10,000 rpm for 5 min. The DNA extracted was checked in terms of quality and concentration by means both of 0.8% agarose gel electrophoresis and of a spectrophotometer (Ultrospec 2100 pro, Biochrom, Cambridge) at 260 nm. In particular, DNA from fresh leaves resulted to have optimal quality and a concentration of 100 ng/μl, while DNA extracted from oil had lower concentration (5 ng/ml) and was partially degraded.

Molecular analysis

Ten microsatellite primer pairs (Table 1), labeled either with fluorochromes FAM or HEX [14], were used. These primer pairs are available in current literature [15–18] and were chosen, after a screening of 18 microsatellite markers carried out in previous works [8], for their high polymorphism and low complexity in pattern profiles. Amplification reactions were performed in an I-Cycler programmable thermal cycler (Bio-Rad Laboratories, Hercules, CA) in a reaction mix with 50 ng of DNA, 1 × PCR buffer, 2.4 μM MgCl₂, 2 μM dNTP, 2.5 μM primer forward and reverse each, and 1 U of *Taq* DNA polymerase (Sigma, St. Louis, MO), in a total volume of 25 μL. The amplification conditions were: 5 min at 94 °C; 35 cycles composed of 30 s at 94 °C, 30 s at the appropriate annealing temperature as reported in Table 1, and 30 s at 72 °C; final elongation at 72 °C for 60 min. The amplification products were first denatured at 94 °C for 5 min and then separated by capillary electrophoresis on an ABI PRISM[®] 3100 Avant Genetic Analyzer (Applied Biosystems). The obtained electrophoregrams were acquired and analyzed by the software GeneMapper 3.7 (Applied Biosystems). Each PCR procedure was reproduced at least twice for each DNA sample and the electrophoregrams were scored by two different operators.

Table 1 Primer sequences, repeat motif, size range, number of alleles and power of discrimination (PD), observed (H_o) and expected (H_e) heterozygosity of ten microsatellites used in the present analysis

SSR denomination	Repeat motif	Primer sequences (5 ^a → 3 ^a)	T _a (°C)	Size range recorded	No. alleles	PD	H_o	H_e
DCA 03	(GA) ₁₉	cccaagcggaggtgtatattgttac tgcttttgcgtgtttgagatggtg	50	231–253	5	0.82	0.86	0.73
DCA 04 ^a	(GA) ₁₆	cttaacttgcgtctctccataacc agtgacaaaagcaaaagactaaagc	55	128–188	6	0.82	–	–
DCA 07 ^a	(AG) ₁₉	ggacataaaacatagagtgcgggg agggtagccaactgctaataagacg	60	121–169	10	0.82	–	–
DCA 14	(CA) ₁₈ A ₆ (TAA) ₇	aatttttaatgcactataatttac ttgaggtctctatatctcccagggg	50	145–190	5	0.78	0.57	0.62
DCA 15	(CA) ₃ G(AC) ₁₄	gatcttgcgtgtatattccacac tatacctttccatcttgacgc	50	244–267	4	0.69	0.71	0.68
DCA 18	(CA) ₄ CT(CA) ₃ (GA) ₁₉	aagaagaaaaaggcagaattaagc gttttcgtctctacataagtgc	50	167–195	4	0.74	0.71	0.68
GAPU 103 ^a	(TC) ₂₆	tgaatttaactttaaccacacaca gcatcgcctcattttatcc	52	132–196	10	0.86	–	–
EMO 90	(CA) ₁₀	catccggatttcttgccttt agccgaatgtagctttgcatgt	55	183–193	4	0.69	0.86	0.68
EMO L	(GA) ₁₂	ctttccaatatgggctctcg atggcactttacgggaaaaa	55	192–212	4	0.69	0.71	0.58
UDO 43 ^a	(GT) ₁₂	tgcgcttacaaccatttc tgccaattatggggtaact	50	141–218	12	0.86	–	–

^a H_o and H_e were not determined since some cultivars yielded three or more different amplified products. The attribution of alleles to any locus has not been attempted, in the absence of segregation data

Statistical analysis

For studying the informative potential of the microsatellites, the observed (H_o) and expected (H_e) heterozygosities were calculated using the software POPGENE ver. 1.31 [19]. H_e values were estimated using the formula proposed by Nei [20]:

$$H_e = 1 - \sum p_i^2$$

where p_i is the frequency of the i th allele.

The power of discrimination (PD) [21] of microsatellite primer pairs was calculated as reported by Cipriani et al. [16], where the allele frequency of the H_e formula is replaced by the genotype frequency.

Cultivars showing only one fragment amplified by a primer pair were conventionally considered to be homozygous at that locus. As a consequence the genetic diversity could be underestimated, if null alleles occurred.

Microsatellite polymorphic profiles were used to create ten matrices, one for each primer pair, where fragments were scored as either present (1) or absent (0). Finally the data were grouped in one single matrix. NTSYS-PC ver. 2.0 [22] was used to calculate genetic similarity among cultivars using Jaccard index [23]. The Unweighted Pair Group Method using Arithmetic Averages (UPGMA)

procedure was used for cluster analysis and to generate both a dendrogram and an identification key.

Results and discussion

Ten SSR primer pairs were analyzed on the set of leaves and oil samples from seven cultivars. They revealed a total of 64 alleles ranging from 4 to 12 alleles per primer pair (Table 1). We considered as reliable those amplifications that resulted repeatable and unambiguous in their interpretation. In general, the primer pairs employed gave acceptable amplifications (Fig. 1) both on leaf and oil DNA, apart from few cases, reported in Table 2, when the oils missed to amplify. This could be due to the fact that the target DNA of SSR markers is represented by short fragments, that are easily detectable, on not degraded DNA, but could be undetectable on high degraded DNA as the case of DNA from oil [3, 4, 9].

For these reasons SSRs are suitable tools for the analysis of DNA extracted from processed food, but sometimes is very difficult to obtain a reliable amplification, even if the DNA extraction and amplification experiments are repeated several times. This can be caused by the low concentration of DNA and by the nature of SSR markers,

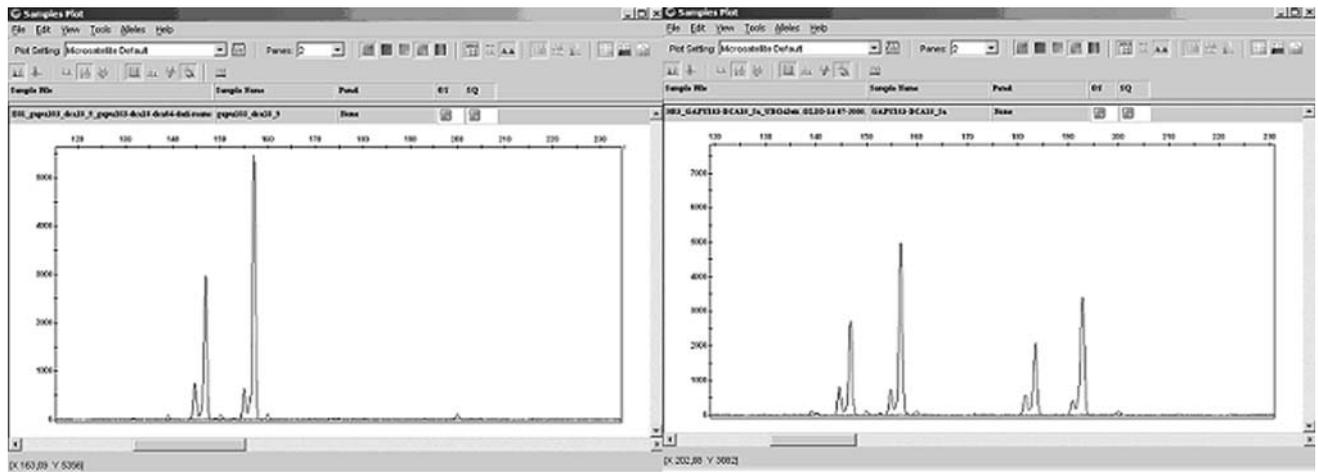


Fig. 1 Two different capillary electropherograms showing two additional alleles detected in oil DNA of Toscanina (*right*), with GAPU103, respect to leaf DNA (*left*) (internal size standard—ROX

500). X-axis indicates fragment length (bp), while Y-axis indicates relative fluorescent units (RFU) that are proportional to the amount of PCR product

which do not anneal when the flanking regions result extremely degraded. In particular, for what concerns oil samples, on a total of 70 PCR reactions (10 primer pairs \times 7 oil samples) we failed in ten cases, with a percentage of failure of 14.3%. This result is similar to that obtained by Breton [9] (20%). Despite this “physiological lost of data”, on the whole, the technique is efficient and feasible for the traceability of olive oil.

Observing the data on Table 2, it is possible to evidence that, discarding the ten failed PCR, 90% of the experiments showed identical patterns between leaves and oil DNA. In particular, in many cases, also complex electrophoretic patterns composed of multiple alleles resulted reproducible. This is the case of Toscanina cultivar which, when amplified with DCA07, showed five alleles (121, 131, 139, 146 and 163 bp) both in leaf and oil DNA. In particular, the monovarietal oil extracted from Cima di Melfi showed several mismatches with leaves, such as the lost of the 186-bp allele with DCA04, the lost of the 121 and 169-bp alleles, and the appearance of an additional 163-bp allele with DCA07, and that of the 184 and 204-bp alleles with GAPU103.

Other mismatches were encountered for the Leccino oil (lost of the 254-bp allele with DCA15), Toscanina oil (additional 184 and 192-bp alleles with GAPU103), and Cellina di Nardò (additional 192-bp allele with EMO L). In our previous works [3, 8] these non-complete concordances between leaves and oil DNA did not emerged, probably due to the methods employed to screen amplification products (i.e. agarose electrophoresis, PAGE), which are known to have lower resolving power and sensitivity than capillary electrophoresis. The loss of some peaks in oil profiles could be due to a degraded DNA for the process of olive crushing for mechanic oil extraction, to a low signal

of the amplicons, or to troublesome working conditions for the polymerase. The disappearance of some peaks is particularly evident in SSR markers with a complex profile (4–5 peaks), where the simultaneous analysis of different regions of the genome could be compromised by the degradation of DNA template.

On the other hand, the appearance of additional alleles in oil profiles with respect to leaf DNA could not be due to wrong denominations of the cultivars or to the occurrence of somaclonal mutations, since the cultivars examined have a certified origin and the leaves and drupe samples belong to the same clone. But they could derive from a paternal contribution of embryo in our oil samples, extracted from entire drupes, although at this regard there is not a complete concordance in literature. Muzzalupo et al. [24] did not found differences comparing destoned and conventional olive oil from the cultivar “Ogliarola Salentina”. On the contrary, Doveri et al. [4] observed diverse profiles in oil and plant tissues of the cultivar “Leccino”, suggesting that additional alleles in oil profiles can have a non-maternal origin that must have arisen from out-crossing. Anyway, both researches focused their investigation on single putative self-sterile cultivars. This discordance in terms of results could be ascribed to: different field location and composition, low number of cultivars investigated, different set of SSR markers employed in the two researches, different PCR and electrophoretic conditions and tools.

A confirmation of the hypothesis that a paternal contribution could be occurred, derives from the molecular weight of all the scored additional alleles, which, even if mismatching from leaf DNA, can be found in the electrophoretic profiles of the other cultivars. An example is given by Cima di Melfi oil profile obtained with DCA07, where

Table 2 Fragment size of ten microsatellites on leaf DNA and monovarietal oil DNA of seven cultivars diffused in Italy

	Amplified fragment size (bp)			Amplified fragment size (bp)	
	Leaf DNA	Monovarietal oil DNA		Leaf DNA	Monovarietal oil DNA
<i>DCA 03</i>			<i>DCA 18</i>		
Coratina	239, 243	239, 243	Coratina	173, 176	n.a.
Picholine	231, 253	231, 253	Picholine	167, 176	167, 176
Toscanina	243	243	Toscanina	167, 176	167, 176
Cima di Melfi	243, 253	243, 253	Cima di Melfi	173, 176	n.a.
Frantoio	237, 243	237, 243	Frantoio	173	173
Leccino	243, 253	243, 253	Leccino	173	173
Cellina di Nardò	231, 239	231, 239	Cellina di Nardò	167, 195	167, 195
<i>DCA 04</i>			<i>GAPU 103</i>		
Coratina	128, 130, 162	128, 130, 162	Coratina	132, 160	n.a.
Picholine	128, 130, 160	128, 130, 160	Picholine	147, 184, 192	147, 184, 192
Toscanina	128, 130, 162, 188	128, 130, 162, 188	Toscanina	147, 157	147, 157, 184, 192
Cima di Melfi	130, 186	130	Cima di Melfi	132, 160, 170, 196	132, 160, 170, 184, 196, 204
Frantoio	128, 130	128, 130	Frantoio	160, 172	n.a.
Leccino	128, 130	128, 130	Leccino	172, 184	172, 184
Cellina di Nardò	130	130	Cellina di Nardò	172	172
<i>DCA 07</i>			<i>EMO 90</i>		
Coratina	121, 148	121, 148	Coratina	183, 189	183, 189
Picholine	121, 133, 148, 163	121, 133, 148, 163	Picholine	183, 185	183, 185
Toscanina	121, 131, 139, 146, 163	121, 131, 139, 146, 163	Toscanina	183, 193	183, 193
Cima di Melfi	121, 133, 148, 169	133, 148, 163	Cima di Melfi	185	185
Frantoio	121, 131, 139, 146, 163	121, 131, 139, 146, 163	Frantoio	183, 189	183, 189
Leccino	121, 131, 139, 146, 163	121, 131, 139, 146, 163	Leccino	183, 189	183, 189
Cellina di Nardò	121, 129, 135, 148, 163	121, 129, 135, 148, 163	Cellina di Nardò	183, 185	183, 185
<i>DCA14</i>			<i>EMO L</i>		
Coratina	179, 186	n.a.	Coratina	192, 198	192, 198
Picholine	186	n.a.	Picholine	192, 198	192, 198
Toscanina	186	186	Toscanina	198	198
Cima di Melfi	186, 190	n.a.	Cima di Melfi	192, 200	192, 200
Frantoio	179, 186	179, 186	Frantoio	192, 198	192, 198
Leccino	177	n.a.	Leccino	198	198
Cellina di Nardò	145, 186	145, 186	Cellina di Nardò	198, 212	192, 198, 212
<i>DCA15</i>			<i>UDO 43</i>		
Coratina	244	n.a.	Coratina	175	175
Picholine	244, 264	244, 264	Picholine	209, 218	209, 218
Toscanina	244, 267	n.a.	Toscanina	171, 215	171, 215
Cima di Melfi	244	244	Cima di Melfi	141, 157, 171, 175, 198	141, 157, 171, 175, 198
Frantoio	244, 264	244, 264	Frantoio	175, 213	175, 213
Leccino	244, 254, 264	244, 264	Leccino	157, 211, 215	157, 211, 215
Cellina di Nardò	244, 264	244, 264	Cellina di Nardò	173, 179	173, 179

Values in italics amplicons different between Leaf DNA and monovarietal oil DNA samples

n.a. not amplified

the additional allele 163 bp is present also in all the other cultivars, except for Coratina. This is true for all the scored additional alleles except for the allele 204 bp scored on Cima di Melfi with GAPU103, which is absent in all the

other profiles. Probably, the origin of this allele derives from a paternal contribution of cultivars not included in this report but present in the same pre-multiplication field. In addition, the high rate of additional alleles scored in

Cima di Melfi can be due to the central position of the tree in the pre-multiplication field and, probably, to a longer flowering duration which overlaps the flowering period of the other cultivars. Another example of additional alleles is given in Fig. 1, where the cultivar Toscanina shows the presence of two paternal alleles in the oil profile. In particular, the 184 bp allele was also scored in Picholine, Cima di Melfi and Leccino, while the allele 192 bp was detected in Picholine, whose pollen probably out-crossed bringing the two alleles. Anyway, the confirmation of this hypothesis should be obtained by setting up a cross-breeding experiment in all the pre-multiplication field, in order to have a precise identification of the origin of the additional alleles in oil profiles and to ascertain if these additional alleles can be ascribed to one or more cultivars.

The Leccino is considered by breeders as a self-sterile cultivar, suggesting that additional alleles in oil profile could be present as a consequence of out-crossing pollination [4]. However, Leccino presented no additional allele in the oil DNA samples for all the primer pairs analyzed, suggesting that probably the absence of additional alleles could be justified by the out-cross with pollinators that present the same electrophoretic profile of the Leccino itself, or by a minimum selfing percentage, which is known to occur also in a cultivar considered to be self-incompatible [25].

The calculation of the genotypic frequency revealed a minimum number of four allele combinations (in case of DCA15, DCA18, EMO90, and EMOL) and a maximum of seven different combinations, one for each cultivar (in UDO43 and GAPU103) (Table 1). A high value of genotypic frequency is usually strictly related to a remarkably high PD of the molecular marker. The PD was calculated on the results obtained from the amplification of leaf DNA, since some monovarietal oil DNA failed to amplify, as reported in Table 2. PD expresses the efficiency of primer pairs in terms of the ability to distinguish between cultivars, and ranged from 0.69 to 0.86, with the highest value for markers with a high genotypic frequency.

For the identification of these seven oils with the lowest number of analyses, the microsatellites having the greater PD were used to create an identification key. UDO43 and GAPU103 showed the highest PD value (0.86) (Table 1), but the latter was discarded because of the failed amplifications. Consequently, a graphic identification key (Fig. 2), based on UDO43 was obtained. To this goal, even if UDO43 amplified 12 alleles, six of them were sufficient to distinguish all the cultivars.

The expected (H_e) and observed (H_o) heterozygosities were calculated only for single-locus primer pairs and varied between 0.58 and 0.73 (0.66 ± 0.05) and 0.57 and 0.86 (0.74 ± 0.11), respectively. In all cases, except one (DCA14) H_o resulted higher than H_e , revealing a high genetic variability among the cultivars investigated. For

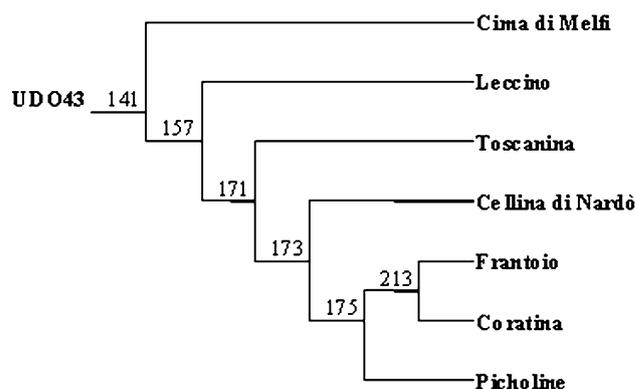
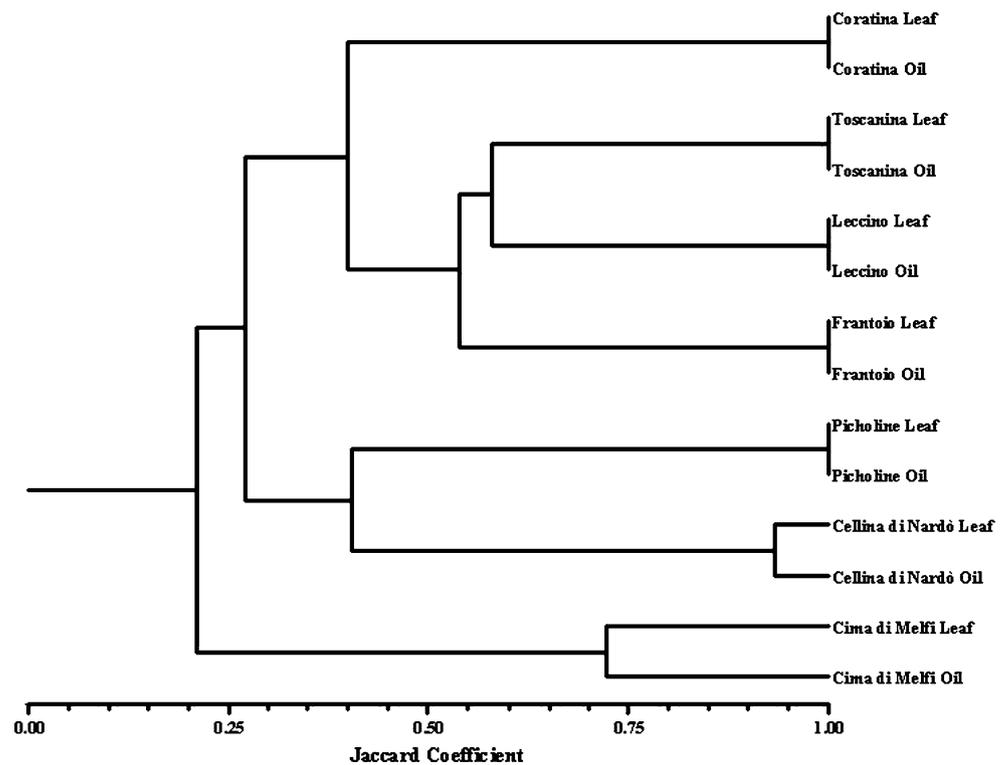


Fig. 2 Identification key of seven olive oil cultivars diffused in Italy based on the SSR marker UDO43. The key was built distinguishing in a dichotomous way cultivars showing the presence of a certain band from those where the band was absent. Each branch corresponds to a different allelic profile, with the *upper* part corresponding to presence and the *lower* to absence of the band reported on each branch. Allele sizes are reported in base pairs

what concerns DCA04, DCA07, GAPU 103 and UDO43, they were not considered for H_e and H_o calculation since they generated multiple bands and, consequently, the attribution of alleles to any locus has not been attempted, in the absence of segregation data. In our study the average expected heterozygosity amounted to 0.66, in line with Diaz et al. [25] who reported how, on the whole, the microsatellites developed in olive, and more in general for the majority of out-cross species which are clonally propagated, seem to be characterized by medium levels of heterozygosity, i.e. never higher than 0.66.

More in general, comparing the obtained allelic profiles of the seven Italian olive cultivars to what reported by other authors, emerged a lack of data referred to the same cultivars analyzed with the same SSRs. Anyway, certain cultivars and SSRs were comparable and our findings resulted consistent to what reported in literature, in other cases a discordance was revealed. An example of complete concordance to what reported on the database <http://www.oleadb.eu/> [26], exclusively for those cultivars for which data are available, is given by DCA03. In other cases, mismatching referred to the number of amplified loci and molecular weights per cultivar were encountered when comparing our results with certain reports. On the contrary, the same results were consistent with other researches. An example is given by the cultivar Frantoio, that revealed a single heterozygous locus 183–189 by applying EMO90 (Table 2). These results are in line with De la Rosa et al. [27], but discordant from [26] that reported, for the cultivar Frantoio, the single heterozygous locus 188/194 by applying EMO90. Furthermore, in this research DCA07 scored some alleles that were not concordant to [26], while they are reported in other studies of genetic diversity in olive [28], revealing the existence of non-homogeneous,

Fig. 3 Dendrogram of genetic similarity among seven leaf DNA and seven monovarietal oil DNA cultivars diffused in Italy based on the UPGMA clustering excluding those SSR primer pairs that failed to amplify in any cultivar



ambiguous or, for some of the cultivars examined in this report, not complete or missing data in the literature. However, intravarietal variation, as reported in some olive cultivars [16, 29, 30], could probably justify, in some cases, these differences in the pattern profiles among reports.

Finally, using data from all the ten microsatellites, a dendrogram was generated to assess the phenetic relationships among cultivars. Moreover, to verify the correspondence between leaf and oil DNA, the dendrogram was obtained by considering each leaf or oil sample as a single genotype. The reported allele mismatches and the failed amplifications explain the differences for some cultivars between leaf and oil DNA (data not shown). In particular, three cultivars (Coratina, Frantoio, and Picholine) showed a Jaccard similarity index of 1 between leaf and oil samples, reflecting what reported in Table 2, where is clearly evident the correspondence of leaf and oil profiles for all the primer pairs. Moreover, as expected, the lower Jaccard index between leaf and oil profiles was showed by the cultivar Cima di Melfi, caused by the mismatches and the failed PCR for DCA14 and DCA18. In order to eliminate the effect of not reliable data, a new cluster analysis was conducted based on six primer pairs, by excluding those ones that failed to amplify in any cultivar, i.e. DCA14, DCA15, DCA18 and GAPU103. The obtained dendrogram, reported in Fig. 3, showed how the cultivars Leccino and Toscanina had a whole correspondence between leaf and oil profile, demonstrating that the differences observed in the first analysis could be

ascribed to missing data deriving from failed PCR. Instead, for Cellina di Nardò and Cima di Melfi the differences between leaf and oil profiles did not change, since they were due to the lost or to the appearance of additional alleles.

In general, the dendrogram reported in Fig. 3 evidenced that no wrong denomination or cases of synonymy occurred, but all the cultivars were distinguishable among them. In particular, Cima di Melfi shared only 20% of the alleles with the other cultivars, representing a single cluster. Two clusters formed at 27%: a first cluster comprised Picholine and Cellina di Nardò, that differed at 40% similarity. A second cluster divided in two subclusters at 40% similarity, in the first of which Coratina was present alone, while the second subcluster was composed by Frantoio (55% similarity), and by Toscanina and Leccino, that showed about 60% similarity.

In conclusion, the study confirmed the possibility to amplify DNA extracted from cell residues present in monovarietal olive oils obtained by processes of mechanic extraction. Anyway, in some cases DNA resulted degraded and thus not amplifiable. The use of capillary electrophoresis by an automatic sequencer facilitated the identification of specific alleles, even for weak signals, so that DNA microsatellites were able to distinguish and identify olive oils from different cultivars.

Acknowledgments This work was conducted within the framework of “Progetto interregionale OLVIVA” funded by Programma

Interregionale—“Sviluppo rurale” sottoprogramma “Innovazione e ricerca” D.M. n. 25279 del 23/12/03”.

References

1. Consolandi C, Palmieri L, Severgnigni M, Maestri E, Marmioli N, Agrimonti C, Baldoni L, Donini P, De Bellis G, Castiglioni B (2008) *Eur Food Res Technol*. doi:10.1007/s00217-008-0863-5
2. Commission Regulation (EC) (1989/2003) *Offic J Eur Union*, L 295/57
3. Pasqualone A, Montemurro C, Summo C, Sabetta W, Caponio F, Blanco A (2007) *J Agric Food Chem* 55:3857–3862
4. Doveri S, O’Sullivan DM, Lee D (2006) *J Agric Food Chem* 54:9221–9226
5. Tautz D (1989) *Nucl Acids Res* 17:6463–6471
6. Muzzalupo I, Perri M (2002) *Eur Food Res Technol* 214:528–531
7. de la Torre F, Bautista R, Cánovas F, Claros G (2004) *Food Agric Environ* 2:84–89
8. Pasqualone A, Montemurro C, Caponio F, Blanco A (2004) *J Agric Food Chem* 52:1068–1071
9. Breton C, Claux D, Metton I, Skorski G, Bervillé A (2004) *J Agric Food Chem* 52:531–537
10. Montemurro C, Pasqualone A, Simeone R, Sabetta W, Blanco A (2008) *Eur J Food Technol* 226:1439–1444
11. Council Directive 92/34/EEC. *Offic J Eur Union*, L 157
12. Commission Directive 93/48/EEC. *Offic J Eur Union*, L 250
13. Doyle JJ, Doyle JL (1990) *Focus* 12:13–15
14. Osborn AM, Moore ERB, Timmis KN (2000) *Environ Microbiol* 2:39–50
15. Sefc KM, Lopes MS, Mendoca D, Rodrigues Dos Santos M, Laimer Da Camara Machado M, Da Camara Machado A (2000) *Mol Ecol* 9:1171–1173
16. Cipriani G, Marrazzo MT, Marconi R, Cimato A, Testolin R (2002) *Theor Appl Genet* 104:223–228
17. Carriero F, Fontanazza G, Cellini F, Giorio G (2002) *Theor Appl Genet* 104:301–307
18. De La Rosa R, James CM, Tobutt KR (2002) *Mol Ecol Notes* 2:265–267
19. Yeh FC, Yang RC, Boyle T (1999) POPGENE. Microsoft Window-based Freeware for Population Genetic Analysis, Version 1.31. Department of Renewable Resources, University of Alberta Edmonton, Canada
20. Nei M (1973) *Proc Natl Acad Sci USA* 70:3321–3323
21. Kloosterman AD, Budowle B, Daselaar P (1993) *Int J Legal Med* 105:257–264
22. Rohlf M (1998) NTSYS-PC. Numerical Taxonomy and Multivariate Analysis System, Version 2.02i. Department of Ecology and Evolution. State University of New York, Setauket, New York
23. Jaccard P (1901) *Bull Soc Vaucloise Sci Nat* 37:547–579
24. Muzzalupo I, Pellegrino M, Perri E (2007) *Eur Food Res Technol* 224:469–475
25. Diaz A, Martin A, Rallo P, Barranco D, de la Rosa R (2006) *J Am Soc Hortic Sci* 131:250–255
26. The Olea Database available on: <http://www.oleadb.eu>
27. De la Rosa R, James CM, Tobutt KR (2004) *Hortic Sci* 39(2):351–354
28. Sarri V, Baldoni L, Porceddu A, Cultrera NGM, Contento A, Frediani M, Belaj A, Trujillo I, Cionini PG (2006) *Genome* 49:1606–1615
29. Belaj A, Rallo L, Trujillo I, Baldoni L (2004) *Hortic Sci* 39:1566–1570
30. Gemas VJV, Almadanim MC, Tenreiro R, Martins A, Feveireiro P (2004) *Genet Resour Crop Evol* 51:501–511