J. Amer. Soc. Hort. Sci. 129(2):266–270. 2004.

Optimal Use of RAPD Markers for Identifying Varieties in Olive (*Olea europaea* L.) Germplasm Collections

A. Belaj,[1] Z. Satovic,[2] L. Rallo, and I. Trujillo
Departamento de Agronomía, ETSIAM, Universidad de Córdoba, Avenida Ménendez Pídal s/n, Apdo 3048, 14080 Córdoba, Spain

Additional index words. confusion probability, cultivar identification, discriminating capacity

Abstract. The aim of this work was to study in depth the resolving power of RAPD markers for rapid and reliable identification of olive cultivars in germplasm collections. The D parameter (the probability that two randomly chosen cultivars have different banding patterns), used for that purpose, showed high values for most of the 21 primers tested and its values ranged from 0.6114 (OPI-13) to 0.9762 (OPK-16) with a mean value of 0.8566. This parameter was used to select the five most discriminating primers: OPK-16, OPA-19, OPX-09, OPF-06 and OPZ-11. The joint confusion probability and the statistical number of indistinguishable pairs of cultivars were estimated for these primers (under independence hypothesis). The combination of three primers (OPK-16, OPA-19 and OPX-09) was found optimal for rapid discrimination of 103 cultivars with a very low value of cumulative confusion probability (1.72×10^{-5}), leaving 0.09 pairs of cultivars indistinguishable. This fact, together with the efficiency of the most discriminating primers combination on an increasing number of cultivars, evidenced the utility of RAPD markers for discrimination of olive cultivars in collections and in nurseries.

The olive (*Olea europaea* L.) tree is a species of great socioeconomic importance in the Mediterranean basin, where 95% of the world production occurs. There are 79 olive collections located in 24 countries which contain about 1200 cultivars with more than 3000 different names (Bartolini et al., 1998). The use of cultivar synonyms and homonyms, the existence of a very large number of varieties and the confusion in naming utility (Barranco and Rallo, 1984) are some of the most important problems in the management of these germplasm collections. This reinforces the need to unambiguously distinguish between cultivars and to clarify synonyms and homonyms for an efficient use and management of olive genetic resources in the collections. Precise and rapid cultivar identification is required in olive breeding programs to accurately identify the parents and further to distinguish new cultivars for proprietary rights protection. Furthermore, a reliable verification of cultivar identity is a very important aspect for the olive growing industry, nurserymen and growers as, similarly to other fruit trees, the cost of the plant material is high and the plantation errors cannot be detected before many years of cultivation (Arus et al., 1993).

The introduction of molecular biology techniques provides an opportunity for genetic characterization that allows direct comparison of different genetic material independent of environmental influences, but the practical utility of any molecular approach for germplasm management is partly determined by the ability to differentiate between large numbers of accessions. That dictates that the protocol used should be quick, uncomplicated and cheap**.** The greatest challenges for the identification of cultivars by means of molecular markers are to reduce the cost of analysis (i.e., the number of amplifications and thus the number of primers) as well as the risk of confusing one of these genotypes with a randomly chosen genotype taken from a larger sample (Tessier et al., 1999).

Different molecular markers such as isoenzymes (Ouazzani et al., 1993; Trujillo et al., 1995), randomly amplified polymorphic DNA (RAPD) markers (Barranco et al., 2000; Belaj et al., 2001; Besnard et al., 2001; Fabbri et al., 1995; Mekuria et al., 1999; Sanz-Cortes et al., 2001; Wiesman et al., 1998), as well as amplified fragment length polymorphisms (AFLP) (Baldoni et al., 2000) and simple sequence repeats (SSR) (Carriero et al., 2002; Cipriani et al., 2002; Rallo et al., 2000; Sefc et al., 2000) have been used for identifying olive cultivars. However, the majority of these studies was conducted on a small number of samples and does not focus on the most important problem, that is the way to optimally apply these techniques for identification purposes in germplasm collections.

The present work is part of a larger study started in 1996 on the World Olive Germplasm Bank of Córdoba (Spain). The first studies mainly aimed at the setting up of PCR-RAPD techinque in olive and the evaluation of its polymorphism for cultivar discrimination. For that purpose, a subsample of 51 cultivars was analyzed by means of 46 primers separated in agarose (Belaj et al., 2001). The objective of this work was to study in depth the optimal use of RAPD markers in Germplasm Banks. The discriminating capacity parameter was used to evaluate the utility of a given RAPD primer for variety identification. The optimal combination of primers to identify a high number of accessions (103) from the World Olive Germplasm Bank of Córdoba with a minimum theoretical risk of confusion was selected. Finally, we estimated the statistical number of indistinguishable pairs when applying a single primer and combination of primers on an increasing number of varieties.

Materials and Methods

Plant material and DNA extraction. One hundred and three cultivars of olive from the World Germplasm Bank of the CIFA "Alameda del Obispo" in Cordoba, Spain, were studied (Table 1). Total genomic DNA was isolated from fresh leaves as described by Belaj et al. (2001).

Received for publication 23 Jan. 2003. Accepted for publication 14 Oct. 2003. Contribution from the Department of Agronomy, University of Cordoba, Cordoba, Spain. We appreciate the help of J. M. Caballero and Carmen del Río from the Germplasm Bank of Córdoba. This research was funded by the INIA (Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria del MAPA), Project CAO 98-001-C3-1. A. Belaj had a PhD grant from the AECI (Agencia Española de Cooperación Internaciional).
[1]Corresponding author; e.mail ag2belay@uco.es.
[2]Faculty of Agriculture, University of Zagreb, Department of Seed Science and Technology, Svetosimunska 25, 10000 Zagreb, Croatia.

Table 1. Olive cultivars analyzed including their register number (R.N.) and the countries of origin.

Cultivar	RN	Origin	Cultivar	R.N.	Origin
Abu Satel Echlot	1040	Syria	Kalamon	105	Greece
Adramitini	102	Greece	Kalinjot	1171	Albania
Alfafara	605	Spain	Kan Çelebi	788	Turkey
Aloreña	829	Spain	Kiraz	679	Turkey
Amigdaloia	228	Greece	Konservolia	219	Greece
Amigdaloia Nana	696	Greece	Koroneiki	218	Greece
Arauco	833	Argentina	Lastovska	704	ExYugoslavia
Arbequina	231	Spain	Leccino	82	Italy
Ascolana Tenera	62	Italy	Leccio del Corno	83	Italy
Ayrouni	134	Lebanon	Lechín de Granada	54	Spain
Ayvalik	97	Turkey	Lechín de Sevilla	5	Spain
Barnea	711	Israel	Lucques	322	France
Beyaz Yaglik	683	Turkey	Manzanilla Cacereña	430	Spain
Blanqueta	11	Spain	Manzanilla de Sevilla	21	Spain
Bouteillan	63	France	Megaritiki	108	Greece
Buga	733	ExYugoslavia	Memeçik	93	Turkey
Cakir	96	Turkey	Menara	836	Morocco
Cañivano Blanco	52	Spain	Merhavia	102	Israel
Carolea	736	Italy	Meski	115	Tunisia
Carrasquenha	125	Portugal	Mixani	1079	Albania
Castellana	576	Spain	Moraiolo	78	Italy
Cellina	179	Italy	Morisca	955	Spain
Chalkidiki	220	Greece	Morrut	224	Spain
Changlot Real	15	Spain	Nabali	158	Palestine
Chemlali	744	Tunisia	Negrinha	123	Portugal
Chemlal de Kabylie	118	Algeria	Oblica	706	ExYugoslavia
Chetoui	113	Tunisia	Ouslati	114	Tunisia
Cobrançosa	124	Portugal	Picholine	70	France
Coratina	79	Italy	Picholine Marocaine	101	Morocco
Cordovil de Serpa	131	Portugal	Picual	9	Spain
Cornicabra	10	Spain	Picudo	3	Spain
Crnica	734	ExYugoslavia	Redondil	127	Portugal
Daebli	1044	Syria	Rosciola	88	Italy
Dam	1003	Syria	Salonenque	73	France
Domat	94	Turkey	Santa Caterina	72	Italy
Elmaçik	686	Turkey	Sevillenca	227	Spain
Empeltre	13	Spain	Sigoise	119	Algeria
Farga	12	Spain	Souri	858	Lebanon
Frantoio	80	Italy	Tanche	74	France
Galega	128	Portugal	Toffahi	721	Egypt
Gemlik	92	Turkey	U. Bardhë i Tiranës	1082	Albania
Gerboui	538	Tunisia	Uslu	95	Turkey
Gordal Sevillana	234	Spain	Valanolia	103	Greece
Grappolo	181	Italy	Vera	660	Spain
Hamed	722	Egypt	Verdale	76	France
Haouzia	835	Morocco	Verdial de Badajoz	988	Spain
Hojiblanca	2	Spain	Verdial de Huevar	6	Spain
Istarska Belica	735	ExYugoslavia	Verdial de Velez Málaga	883	Spain
Itrana	68	Italy	Villalonga	364	Spain
Izmir Sofralik	99	Turkey	Zaity	788	Syria
K. M. Berat	1080	Albania	Zalmati	117	Tunisia
Kaissy	975	Syria			

RAPD analysis. Amplifications were performed as described by Belaj et al. (2001). All the reactions were conducted three times using DNA of different extractions and different lots of the AmpliTaq DNA polymerase Stoffel fragment (Applied Biosystem, Foster City, Calif.). The amplification products were separated on polyacrylamide gels of 18 × 16 cm containing 10% acrylamide, 0.126% piperazine diacrylamide crosslinker in 0.375 M Tris-HCl, pH 8.8, using Tris glycine (0.025 M Tris, and 0.192 M glycine) and were visualized by silver staining as described by Bassam et al. (1991). Twenty one primers from kits A, F, I, J, K, P, Q, X, and Z (Operon Technologies, Alameda, Calif.) were used in the study (Table 2).

Data analysis. RAPD bands were scored as 1 (present) or 0 (absent) in a binary matrix for each primer following a con-

servative criterion for their selection and each gel was scored independently. The number of banding patterns, i.e., the different combinations of bands obtained and their frequencies were calculated for each primer. The amplification banding patterns of two cultivars ('Arbequina' and 'Lechín de Sevilla') were used as standards for comparing different gels.

To evaluate the efficiency of the RAPD markers in varietal identification, the following estimates were used: confusion probability (C_j) of the *j*th given primer as defined by Tessier et al. (1999):

$$C_j = \sum_{i=1}^{I} p_i \frac{(Np_i - 1)}{N - 1}$$

where p_i is the frequency of the *i*th pattern, N = the sample size, and I = the total number of patterns generated by the *j*th assay unit. Joint confusion probability (C_k) for a given set of primers is a product of the C_j of each primer under the independence hypothesis.

The confusion probability of a given primer provides an estimate of the probability that two randomly chosen genotypes from the sample of 103 cultivars have identical banding patterns for that primer. Thus discriminating power (D_j) of the *j*-th primer is equal to $D_j = 1\text{-}C_j$. Limit of D_j as N tends toward infinity was also calculated.

$$D_L = \lim(D_j) = 1 - \sum_{i=1}^{I} p_i^2$$

In a set of N genotypes, it is possible to draw N (N – 1)/2 different pairs; thus theoretically, the total number of non-distinguishable pairs of cultivars is given by $x_j = [N(N-1)/2]\, C_j$. For a given combination of *k* primers, under the hypothesis of independence, this number (X_k) is equal to the product of the x_j of each primer (Tessier et al., 1999). X_k values were calculated on the 103 cultivars included in this study as well as on an increasing number of varieties.

Results and Discussion

RAPD polymorphism. With the 21 primers used, 126 reliable RAPD markers (93% of the total fragments) could be selected (Table 2). The number of polymorphic markers per primer ranged from 3 to 10 and their combination generated from 5 (OPZ-07) to 43 (OPK-16 and OPA-19) banding patterns per primer, with 17.7 banding patterns per primer, on average. The frequency of banding patterns per primer (data not shown) ranged from 0.01 (banding patterns present in only one cultivar) to 0.37 (banding patterns obtained in 38 genotypes).

High values of discrimination power (Table 2) were obtained for the majority of the primers and its values ranged from 0.6114 (OPI-13) to 0.9762 (OPK-16), with a mean value of 0.8566. The high discriminating power of the RAPD primers examined in this study is consistent with that reported for these markers used in cultivar identification of olive (Besnard et al., 2001) and grape (Tessier et al., 1999). However, the average number of banding patterns per primer and the average discriminating power were higher than in these studies. The high diversity found in olive, a better representative in this study of olive cultivar diversity in Mediterranean basin as well as a higher resolution provided by polyacrylamide gels could have influenced these differences.

The efficiency of a given primer depends on the number of fragments it generates as well as on the frequency of their banding patterns. A primer has a maximal discriminating power when it generates patterns at the same frequency (the isofrequency situation, Tessier et al., 1999). D values were fairly close to the limits of discriminating power (D_L) values, as the size sample was rather large (103 cultivars). On the other hand, the D and D_L parameter values were close to the maximal discriminating power (D_{max}) calculated for the isofrequency situation (Table 2). As the number of banding patterns increased the differences between D_{max} and D_L values ($D_{max} - D_L$) decreased ($r = -0.68$; $p < 0.01$).

Table 2. Primer discriminating power calculated from the sample of 103 olive varieties (D), estimated as N tends toward infinity (D_L), and estimated for a given number of banding patterns at the same frequency (D_{max})

Orders of D and D_L	Primer	No. of markers	No. of banding patterns	D	D_L	D_{max}
1	OPA-19	10	43	0.9744	0.9617	0.9767
2	OPK-16	10	43	0.9762	0.9632	0.9767
3	OPX-09	10	32	0.9718	0.9548	0.9688
4	OPF-06	7	26	0.9481	0.9356	0.9615
5	OPZ-11	6	24	0.9457	0.9316	0.9583
6	OPX-03	6	23	0.9231	0.9114	0.9565
7	OPK-17	7	19	0.8747	0.8639	0.9474
8	OPI-06	5	16	0.8785	0.8661	0.9375
9	OPI-12	5	16	0.8135	0.8044	0.9375
10	OPI-14	6	15	0.8591	0.8491	0.9333
11	OPQ-15	6	15	0.8556	0.8440	0.9333
12	OPP-19	6	14	0.8361	0.8263	0.9286
13	OPX-13	5	14	0.8948	0.8809	0.9286
14	OPA-03	6	13	0.9032	0.8889	0.9231
15	OPJ-18	5	13	0.8576	0.8457	0.9231
16	OPX-01	6	11	0.7978	0.7894	0.9091
17	OPZ-10	5	10	0.8773	0.8632	0.9000
18	OPK-07	4	7	0.7735	0.7653	0.8571
19	OPA-01	4	6	0.7386	0.7294	0.8333
20	OPI-13	4	6	0.6114	0.6017	0.8333
21	OPZ-07	3	5	0.6781	0.6688	0.8000
Average		6	17.67	0.8566	0.8450	0.9202

Table 3. Joint confusion probabilities (C_k) for five most discriminating primers and the number of non-distinguishable pairs (X_k) for a single primer or a given combination of primers on a set of 103 varieties (5253 pairs).

No.	Primer	C	Cumulative primers	Ck	X_k
1	OPK-16	0.0238	1	2.38×10^{-2}	124.9
2	OPA-19	0.0256	1 + 2	6.10×10^{-4}	3.2
3	OPX-09	0.0282	1 + 2 + 3	1.72×10^{-5}	9.0×10^{-2}
4	OPF-06	0.0519	1 + 2 + 3 + 4	8.93×10^{-7}	4.7×10^{-3}
5	OPZ-11	0.0543	1 + 2 + 3 + 4 + 5	4.84×10^{-8}	3.0×10^{-4}

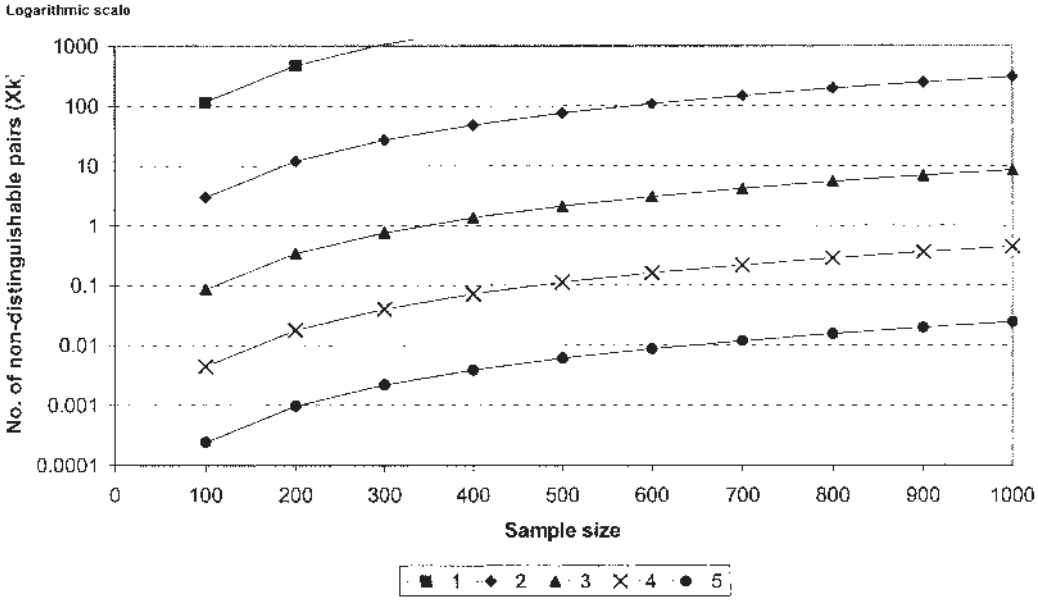
Determining the optimal set of primer combinations for identification purposes. One major application of RAPD markers in olive is the identification of cultivars. Therefore, the potential of each primer to yield different genotypes for as many cultivars as possible with a minimum risk of confusion is of great interest. Selection of the most informative primers reduces the cost of analysis for reliable cultivar distinction. The first five primers (OPK-16, OPA-19, OPX-09, OPF-06 and OPZ-11) were chosen in this study on the basis of their discriminating power (Table 2). The joint confusion probability and theoretical cumulative number of indistinguishable pairs was calculated with these primers under the independence hypothesis (Table 3). As expected, the values of both these parameters decreased as the number of primer combinations increased.

The combination of three primers (OPK-16, OPA-19 and OPX-09) can be considered as a theoretical optimum for rapid and reliable discrimination of 103 cultivars with a cumulative confusion probability value of 1.72×10^{-5}, leaving only 0.0903 pairs of cultivars, from the 5253 theoretical pairs, indistinguishable. The high discriminating capacity of RAPD markers in olive is expected; only three primers make possible the identification of 102 clones with a low confusion probability among cultivars (Besnard et al., 2001) while four primers detected in agarose were needed for the identification of 51 cultivars from the Germplasm Bank of Córdoba, (Belaj et al., 2001).

Despite their very high discriminating capacity, these three primers (OPK-16, OPA-19 and OPX-09) as well as the rest of the primers used in this study were not able to discriminate a few pairs of cultivars between them. These were: 'Cakir-Valanolia'; 'Gordal Sevillana'-'Santa Caterina'; 'Cañivano Blanco'-'Picholine Marocaine'; 'Manzanilla de Sevilla'-'Redondil'; 'Manzanilla Cacereña'-'Negrinha'; 'Grappolo'-'Leccio del Corno'; and 'Cellina'-'Frantoio'. Two hypotheses can explain the fact that even these primers with very high discriminating capacity, could not distinguish a few pairs of cultivars between them. The first is the lack of complete statistical independence of the patterns generated by the primers and the second hypothesis is that these pairs of cultivars are synonyms. Previous investigations using morphological markers (Barranco and Trujillo, personal communication) and isoenzymes (Trujillo et al., 1995) have defined these cultivars as putative synonyms. However, further studies using SSR and AFLP markers are needed in order to confirm the status of these cultivars.

Finally, the expected number of indistinguishable pairs when applying a single primer or a combination of primers on an increasing number of varieties, was calculated (Fig. 1). Both the combination of four (OPK-16, OPA-19, OPX-09 and OPF-06) and five (OPK-16, OPA-19, OPX-09, OPF-06 and OPZ-11) primers could discriminate up to 1000 cultivars, statistically leaving less than one pair of cultivars indistinguishable. According to these data, four primers would theoretically be sufficient to discriminate the total number of accessions (356) found currently at the World Germplasm Bank of Córdoba (del Río, personal communication).

According to our results, under careful laboratory practice and stringent selection of reproducible markers (Belaj et al., 2001), RAPD technology provides an inexpensive and reliable method for routine screening of a large number of cultivars to monitor and manage the genetic resources of olive germplasm collections. The high capacity for cultivar discrimination of the combination of three primers (OPK-16, OPA-19 and OPX-09) is an evidence of



Legend: (1) OPK-16; (2) OPK-16 + OPA-19; (3) OPK-16 + OPA-19 + OPX-09; (4) OPK-16 + OPA-19 + OPX-09 + OPF-06; (5) OPK-16 + OPA-19 + OPX-09 + OPF-06 + OPZ-11.

Fig. 1. Estimation of the number of non-distinguishable pairs (X_k) when applying a single most discriminative primer and primer combinations on an increasing number of varieties. Legend: (1) OPK-16; (2) OPK-16 + OPA-19; (3) OPK-16 + OPA-19 + OPX-09; (4) OPK-16 + OPA-19 + OPX-09 + OPF-06; (5) OPK-16 + OPA-19 + OPX-09 + OPF-06 + OPZ-11.

the utility of RAPD markers for that purpose in olive. For RAPD analysis the problems of reliability and transferability among laboratories should be considered. The molecular data generated from the RAPD/fingerprints of olive germplasm collections should be completed with data obtained by other molecular markers such as SSRs and AFLPs, which may facilitate the interchange of information among different groups and collections.

Literature Cited

Arús, P., R. Messeguer, M.C. de Vicente, P. Puigdomènech, and F. Vargas. 1993. Marcadores moleculares para facilitar la mejora de Prunus. Hortofruticultura 11:54–57.

Baldoni, L., M. Pellegrini, M. Mencuccini, A. Angiolillo, and M. Mulas. 2000. Genetic relationships among cultivated and wild olives revealed by AFLP markers. Acta Hort. 521:275–284.

Barranco, D. and L. Rallo. 1984. Las variedades de olivo cultivadas en Andalucía. M.º de Agricultura, Junta de Andalucía, Madrid, Spain.

Barranco, D., I. Trujillo, and P. Rallo. 2000. Are 'Oblonga' and 'Fantoio' the same cultivar? Hortscience 35:1323–1325.

Bartolini, G., G. Prevost, C. Messeri, and G. Carignani. 1998. Olive germplasm: Cultivars and world-wide collections. Food Agr. Org., Rome.

Bassam, B.J., G. Caetano-Anollés, and P.M. Gresshoff. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal. Biochem. 80:81–84.

Belaj, A., I. Trujillo, R. de la Rosa, L. Rallo, and M.J. Giménez. 2001. Polymorphism and discriminating capacity of randomly amplified polymorphic markers in an olive germplasm bank. J. Amer. Soc. Hort. Sci. 126:64–71.

Besnard, G., C. Breton, P. Baradat, B. Khadari, and A. Bervillé. 2001. Cultivar identification in olive based on RAPD markers. J. Amer. Soc. Hort. Sci. 126:668–675.

Carriero, F., G. Fontanazza, F. Cellini, and G. Glorio. 2002. Identification of simple sequence repeats (SSRs) in olive (*Olea europaea L.*). Theor. Appl. Genet.104:301–307.

Cipriani, G., M.T. Marrazzo, R. Marconi, A. Cimato, and R. Testolin. 2002. Microsatellite markers isolated in olive are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars (*Olea europaea L.*). Theor. Appl. Genet. 104:223–228.

Fabbri, A., J.I. Hormaza, and V.S. Polito. 1995. Random amplified polymorphic DNA analysis of olive (*Olea europaea* L.) cultivars. J. Amer. Soc. Hort. Sci. 120:538–542.

Mekuria, G.T., G.G. Collins, and M. Sedgley. 1999. Genetic variability between different accessions of some common commercial olive cultivars. J. Hort. Sci. and Biotechnol. 74:309–314.

Ouazzani, N., R. Lumaret, P. Villemur, and F. di Guisto. 1993. Leaf allozyme variation in cultivated and wild olive trees (*Olea europaea* L.). J. Hered. 84:34–42.

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

Sanz-Cortés, F., M.L. Badenes, S. Paz, A. Iñiguez, and G. Llácer. 2001. Molecular characterization of olive cultivars using RAPD markers. J. Amer. Soc. Hort. Sci. 126:7–12.

Sefc, K.M., M.S. Lopes, D. Mendonça, M. Rodrigues Dos Santos, M. Laimer Da Câmara Machado, and A. Da Câmara Machado. 2000. Identification of microsatellite loci in olive (*Olea europaea*) and their characterization in Italian and Iberian olive trees. Mol. Ecol. 9: 1171–1173.

Tessier, C., J. David, P. This, J. M. Boursiquot, and A. Charrier. 1999. Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. Theor. Appl. Genet. 98:171-177.

Trujillo, I., L. Rallo, and P. Arus. 1995. Identifying olive cultivars by isozyme analysis. J. Amer. Soc. Hort. Sci. 120:318–324.

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