Theor Appl Genet (2003) 106:521–529
DOI 10.1007/s00122-002-1079-x

B. Khadari · C. Breton · N. Moutier · J. P. Roger
G. Besnard · A. Bervillé · F. Dosba

The use of molecular markers for germplasm management in a French olive collection

Received: 27 April 2002 / Accepted: 7 June 2002 / Published online: 24 October 2002

Abstract With more than 100 accessions, the CBNMP olive collection includes a major part of the French germplasm. We used molecular markers to characterise all accessions and to study genetic relationships between cultivars. Firstly, 497 olive trees were genotyped using 32 RAPD markers. We identified 114 RAPD profiles and detected several cases of mislabelling, synonymy and homonymy. Secondly, for each RAPD profile, one tree was analysed using mtDNA RFLPs to determine the cytoplasmic lineage of each cultivar and using five nuclear SSR loci. French germplasm displayed ME1, MOM and MCK mitotypes with ME1 prevailing (84%). Based on SSR markers, we revealed a slight differentiation between French cultivars growing in the West and the East side of the Rhône Valley. This study allowed us to construct a molecular data-base for the reference collection and to analyse genetic diversity for further prospecting, and for introducing new olive accessions.

Keywords *Olea europaea* · RAPD · mtDNA RFLP · SSR · Molecular characterisation · Genetic relationships

Introduction

For the management of *ex-situ* plant germplasm, two important goals have to be reached. First, all accessions should be characterised in order to eliminate cases of mislabelling and redundancies, and to create a complete data base. Second, to keep a minimum of accessions which should represent a maximum of variability constituting a core collection. The selection of representative accessions is mostly based on agronomic traits of interest, adaptive traits, ecogeographic origin and also on allelic richness (Schoen and Brown 1993). This concern is crucial for species which have a large size of plant germplasm like the grass family. For instance, the rice collection represents more than 80,000 accessions (Virk et al. 1995). Conversely, fruit species propagated by vegetative multiplication display a relatively limited number of cultivated forms and the main goal for germplasm management is to characterise and to collect diversified cultivated forms, particularly at the national and regional level.

Olive is a Mediterranean fruit species that is cultivated mainly for oil but also for canned fruits. Olive orchards cover about 7,000,000 ha in the Mediterranean countries. Because of the dietetic value of olive oil, demand is increasing and there is a regular increase in cultivated surfaces. In general, olive cultivation is traditional and most orchards include several cultivars. Due to the economic demand, these traditional orchards are progressively replaced by modern ones which are characterised by one or two cultivars. As a result of this situation, only a few major cultivars are cultivated in the main production areas, whereas minor cultivars are located in restricted areas and are sometimes threatened. This will lead to a loss in olive genetic diversity.

In addition to the international olive germplasm collection at Cordoba (Spain) preserving major Mediterranean cultivars, several collections exist in other Mediterranean countries. The goal of such collections is to safeguard all cultivars, and particularly the minor ones, to avoid a loss in genetic diversity and to offer an interesting genetic basis for breeding programs.

The olive germplasm collection located in the "*Conservatoire Botanique National Méditerranéen de Porquerolles* (CBNMP)" corresponds to the major part of the national French germplasm. The first part of the

Communicated by H.F. Linskens

B. Khadari (✉) · N. Moutier · F. Dosba
UMR Biologie du Développement des Plantes Pérennes Cultivées, INRA, 2, place Viala, 34 060 Montpellier cedex 1, France
e-mail: Khadari@ensam.inra.fr
Tel.: +33-4-99-61-23-70, Fax: +33-4-67-04-54-15

C. Breton · G. Besnard · A. Bervillé
UMR Diversité et Génome des Plantes Cultivées, INRA, 2, place Viala, 34 060 Montpellier cedex 1, France

B. Khadari · J.P. Roger
Conservatoire Botanique National Méditerranéen de Porquerolles, Parc National de Port-Cros, Castel Ste Claire, F 83418 Hyères cedex, France

collection originated from the INRA orchard of Montpellier which included the main cultivars used for agronomic evaluation. Recently, a national charter of olive genetic resources was established in order to ensure conservation, evaluation and valorisation of French olive germplasm. In this context, field collections are regularly performed in order to introduce new accessions. The final aim is to collect all the cultivars representing the French olive germplasm. Another goal is to use the collection as a reference in France.

Management of the CBNMP collection includes a description of its genetic diversity for a reliable characterisation of all accessions since several cases of mislabelling, homonymy (one denomination for several genotypes) and synonymy (one genotype with several denominations) could exist. In the present study, we used molecular markers to characterise all accessions present in the collection, to build a first molecular data-base and to analyse the genetic relationships between cultivars. We used RAPD markers for genotyping the complete collection and then SSR loci and RFLP of mitochondrial DNA to study the genetic structuring of French germplasm.

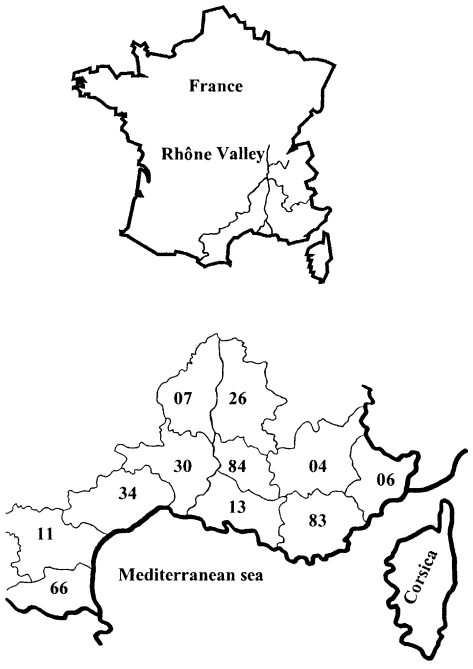

Fig. 1 Map showing French olive-growing areas. The West side of the Rhône Valley corresponds to 66 (Pyrénées Orientales), 11 (Aude), 34 (Hérault), 30 (Gard) and 07 (Ardèche) regions. The East side corresponds to 13 (Bouches du Rhône), 84 (Vaucluse), 26 (Drôme), 04 (Alpes de Haute Provence), 83 (Var) and 06 (Alpes maritimes) regions

Materials and methods

Plant material

Molecular characterisation was performed on 497 olive (*Olea europaea L.*) trees (Fig. 1): 495 trees belonging to 123 accessions and two trees without denomination. We defined an accession as one or several olive trees under the same denomination. The 123 accessions correspond to 95 denominations, 3 clones of *Cailletier*, 3 clones of *Grossane*, 8 clones of *Picholine*, 9 offspring of *Lucques* and 5 offspring of *Verdale* × *Picholine*. The offspring of the *Lucques* cultivar originated from an orchard without controlled pollination. The list of olive accessions analysed, their coordinates (line number and tree number) and the map of the orchard are available on the CBNMP website (http://162.38.231.157/cbnm/collections/).

DNA isolation, nuclear RAPD and mtDNA RFLP procedures

The total DNA-extraction protocol was described by Besnard et al. (2000). All trees were characterised using RAPD markers following the protocol of Quillet et al. (1995). RAPD analysis was performed using four primers (A1, A9, C9 and C15) which had been previously selected for polymorphism and for consistent and reproducible DNA amplification (Besnard et al. 2001a). Reproducibility in RAPD patterns was verified by three different amplifications of DNA from a set of 20 olive accessions.

For each RAPD profile, we determined the mitochondrial type on one tree. The method used for mtDNA RFLP analysis was described by Besnard et al. (2000). Two restriction enzyme/probe combinations (*Hin*dIII/*atp*6 and *Xba*I/*atp*6) were used to identify the four mitotypes ME1, ME2, MOM and MCK previously detected by Besnard et al. (2000).

Microsatellite procedure

A set of SSR loci developed for *Fraxinus excelsior* (Brachet et al. 1999; Lefort et al. 1999) and for *Phillyrea angustifolia* (Saumitou-Laprade et al. 2000), two species of the Oleaceae family, was tested on ten olive cultivars. Five loci were selected because these produced clear and easily readable polymorphic bands (see Table 2). Polymerase chain reaction (PCR) conditions were optimised by adapting annealing temperature (T_m), $MgCl_2$ concentration and the amount of template DNA. The PCR reaction was performed in a mix containing 1.5 to 3 mM of $MgCl_2$, 1 unit per reaction of *Taq* DNA polymerase in buffer, and approximately 50 ng of template DNA in a total reaction volume of 25 µl.

The PCR was carried out using a PTC 100 thermocycler (MJ Research). After 3 min at 94 °C, 30 cycles were performed with 1 min at 94 °C, 1 min at 50–56 °C depending on the primer pair (see Table 2), 1 min at 72 °C and a final extension step of 5 min at 72 °C. Electrophoresis and detection of the PCR products were carried out according to Besnard et al. (2002).

Data analyses

For RAPD data, polymorphic bands were scored as present (1) or absent (0) assuming that each band position corresponds to one locus with two alleles, respectively.

For SSR data, the expected heterozygosity (*He*) was defined as $He = 1 - \sum p_i^2$, where p_i is the allele frequency for the *i*-th allele, and the observed heterozygosity (*Ho*) was calculated using the Genetix V. 4.0 software (Belkhir et al. 1996). Deviations of observed heterozygosity values from Hardy-Weinberg expectations were analysed using the program Genepop (Raymond and Rousset 1995). The probability of null alleles was estimated according to the formula of Brookfield (1996): $r = (He - Ho)/(1 + He)$.

Under the hypothesis of independence between markers, the probability of obtaining a given genotype was calculated as the product of corresponding allele frequencies for each locus of the pattern *i*: $Pi = ∋ fj$. The discriminating power of each RAPD primer and of each SSR locus was computed according to Tessier et al. (1999) as following: $Dj = \sum p_i (Np_i - 1)/N - 1$, where p_i is the frequency of the *i*-th molecular pattern revealed by the primer or by the locus *j*.

Table 1 Codes of the genotypes identified with RAPDs. Designation, origin and mitotype are mentioned for each genotype. Geographic origin is given in parentheses (E = East of Rhône valley; W = West of Rhône valley in South of France; FrC = Corsica, see Fig. 1; Ca = Californie; Gr = Greece; Is = Israel; It = Italy; Mo = Morocco; Sp = Spain; Tu = Tunisia)

Code[a]	Designation and origin	Trees[b]	Mitotype	Code[a]	Designation and origin	Trees[b]	Mitotype
12	*Aglandau*[(c)] (E)	3	ME1	5	*Olivière*[(c)] (W)	6	MCK
10	*Amygdalolia*[(c)] (Gr)	4	ME2	86	*Petite Noire P. R.* (E).	4	ME1
124	*Araban AHP* (E)	4	ME1	2	*Picholine*, *P18*, *P30*, *P35*, *P66*[(c)] (W)	21	MOM
119	*Araban du Var* (E)	4	ME1	64	*Picual*[(c)] (Sp)	1	ME1
17	*Arbequina*[(c)] (Sp)	3	ME1	118	*Pigale*[(c)] (W)	2	MOM
39	*Argental* (W)	4	ME1	52	*Poulo*[(c)] (?)	3	ME2
40	*Ascolana Tenera*[(c)] (It)	4	ME1	121	*Poussou* (E)	3	ME1
24	*Aubenc*[(c)] (W)	4	ME1	111	*Rabeyrolle* (E)	4	ME1
90	*Avellanet* (E)	7	ME1	26	*Rascasset*[(c)] (E)	5	ME1
101	*Baguet* (W)	10	ME1	88	*Razzola* (It)	4	MOM
100	*Barnea 1*[(c)] (Is)	3	ME1	46	*Redouneil*[(c)] (W)	4	ME1
79	*Barnea 2*[(c), (d)] (Is)	3	ME1	25	*Reymet*[(c)] (E)	4	ME1
38	*Barouni*[(c)] (Tu)	5	ME1	131	*Rouge*[(c), (d)] (E)	1	ME1
98	*Béchu de l'Ardèche* (W)	6	MOM	55	*Rougeon*[(c), (d)] (E)	4	ME1
45	*Belle d'Espagne* (It)	5	ME2	58	*Rougette*[(c), (d)] (W)	2	MOM
43	*Berdaneil*, *Poumal*[(c)] (W)	10	ME1	22	*Roussette*, *Verdale de l'Ht*[(c)] (W)	10	ME1
21	*Bid el Hamam*, *Meski*[(c)] (Tu)	9	ME1	35	*Roussette SJG* (W)	5	ME1
93	*Bigarude* (E)	7	ME1	16	*Tanche*[(c)] (E)	2	ME1
11	*Blanc Payzac* (W)	4	ME1	87	*Taulelle* (W)	4	ME1
92	*Blanc Vinezac* (W)	5	ME1	120	*Tripue* (E)	6	ME1
128	*Cailletier*[(c), (d)] (E)	5	ME1	80	*Valensole* (E)	5	ME1
41	*Cayet bleu* (E)	5	ME1	47	*Verdanel*[(c)] (W)	5	ME1
51	*Cayet rouge*[(c)] (E)	4	ME1	50	*Verdelet*[(c), (d)] (W)	4	ME1
20	*Cayon*[(c)] (E)	2	MOM	81	*Vermillau*[(c)] (W)	3	ME1
27	*Celounen*[(c)] (E)	4	ME1	89	*Vilette* (W)	5	ME1
6	*Chemlali*[(c)] (Tu)	4	ME1	1	*Offspring of Lucques*[(1)] (W)	5	ME1
82	*Colombale*[(c)] (E)	5	MOM	3	*Offspring of Lucques*[(1)] (W)	3	ME1
30	*Corniale*[(c)] (W)	4	MCK	4	*Offspring of Lucques*[(1)] (W)	3	ME1
31	*Coucourelle*[(c)] (E)	3	MCK	56	*Galinier*, *L34/1*[(1)] (W)	7	ME1
36	*Courbeil*[(c)] (W)	5	ME1	85	*L11/2*[(1)] (W)	4	ME1
28	*Curnet*[(c)] (E)	5	MCK	49	*L11/48*[(1)] (W)	5	ME1
95	*Darame* (E)	5	ME1	74	*L34/4*, *LINRA*, *L11/25*[(1)] (W)	12	ME1
129	*Dent de Verrat* (E)	1	ME1	103	*VP16*[(2)] (W)	3	ME1
94	*Dorée Vinezac*[(c)] (W)	7	ME1	102	*VP21*[(2)] (W)	6	ME1
76	*Filayre*[(c)] (E)	7	MCK	117	*VP7*[(2)] (W)	3	ME1
37	*Gaïdourolia*[(c)] (Gr)	4	ME1	116	*VP66*[(2)] (W)	3	ME1
99	*Gardisson* (E)	7	ME1	15	2–3, 3–8[(3)]	2	ME1
57	*Germaine*[(c), (d)] (FrC)	2	ME1	29	7–12[(3)]	1	ME1
53	*Grapié*[(c)] (E)	5	ME1	59	5–9[(3)]	1	ME1
32	*Grappola* (It)	5	ME1	62	6–2[(3)]	1	ME1
125	*Grassois* (E)	7	ME1	63	7–5[(3)]	1	ME1
60	*Gros Vert* (E)	2	ME1	65	7–10[(3)]	1	ME1
72	*Grossane*, *G20*, *G28*, *G29*[(c)] (E)	16	MOM	67	9–14[(3)]	1	ME1
33	*Koroneiki Kotreiki*[(c)] (Gr)	10	ME1	68	11–9[(3)]	1	ME1
13	*Lechin de Sevilla* (Sp)	4	ME1	69	12–16[(3)]	1	ME1
127	*Linat* (E)	5	ME1	70	16–9[(3)]	1	ME1
14	*Malaussina*[(c)] (E)	4	ME1	73	17–2, 17–3, 18–6, 13–1, 18–2, 18–3, 18–4, 20–12, 20–13, 20–14, 21–13[(3)]	11	ME1
9	*Manzanilla 1*[(c), (d)] (Sp)	4	ME1	78	18–10, 18–11,18–12, 19–15[(3)]	4	MOM
18	*Manzanilla 2*[(c)] (Sp)	3	ME1	97	25–14[(3)]	1	ME1
7	*Menara*[(c)] (Mo)	5	ME1	104	28–15[(3)]	1	ME1
42	*Menudel* (W)	5	ME1	105	2–13, 4–15[(3)]	1	ME1
19	*Montaurounenque* (E)	4	MOM	107	9–16[(3)]	1	MCK
123	*Moufla* (W)	3	ME1	109	7–1[(3)]	1	ME1
91	*Négret Callian*[(c), (d)] (E)	6	ME1	110	14–13[(3)]	1	ME1
23	*Négrette des Vans*[(c), (d)] (W)	5	ME1	113	13–12[(3)]	1	ME1
44	*Négrette SJB*, *Négrette B*[(c),] (W)	8	ME1	114	15–4[(3)]	1	ME1
8	*Oblonga* (Ca)	5	ME1	126	31–18[(3)]	1	ME1

[a] No of RAPD profiles
[b] Number of olive trees analysed
[(c)] Molecular profiles compared to those obtained by Besnard et al. (2001)
[(d)] Molecular profiles different to those obtained by Besnard et al. (2001)
[(1)] Offspring of *Lucques*
[(2)] Offspring of *Verdale* × *Picholine*
[(3)] Undetermined genotypes

Table 2 Genetic parameters of five SSR loci in olive cultivars of CBNMP germplasm. For each locus, the origin, the range of product size (bp), the number of alleles detected in cultivars (Na), expected (He) and observed heterozygosity (Ho), the probability of exact "Hardy-Weinberg" test (*P*) and the probability of null alleles (*r*) are reported. The value in bold type means that a test is significant

Locus	Origin	Size range (bp)	Number of alleles	Expected heterozygosity	Observed heterozygosity	Probability of exact test	Probability of null alleles
Me 30 Ms	Brachet et al. (1999)	210–238	8	0.822	0.929	0.0891	–0.0587
Femsalt 4	Lefort et al. (1999)	160–202	7	0.699	0.818	0.1440	–0.0700
PA (ATT)2	Saumitou-Laprade et al. (2000)	109–127	6	0.768	0.929	**0.0018**	–0.0912
PA (GA)5	Saumitou-Laprade et al. (2000)	111–129	6	0.795	0.707	0.2747	0.0489
PA (GA)2	Saumitou-Laprade et al. (2000)	106–128	5	0.791	0.404	**0.0000**	**0.2159**

Table 3 Primer discriminating power calculated on 114 olive genotypes

Primer or locus	Number of markers	Number of molecular profiles	Discriminating power (D)
RAPD – A1	11	90	0.993
SSR – Me30Ms	8	23	0.936
SSR – PA(GA)5	6	19	0.933
RAPD – A9	8	27	0.911
SSR – PA(GA)2	5	15	0.901
RAPD – C15	5	19	0.895
SSR – PA(ATT)2	6	15	0.893
SSR – Femsalt4	7	14	0.825
RAPD – C9	8	14	0.709

Genetic relationships between olive genotypes were studied on the basis of SSR data. Genetic distance between cultivars, defined as (1 – proportion of shared alleles), was calculated using the program Microsat (Minch 1997). The corresponding phenogram was drawn based on the UPGMA algorithm using the program Phylip (Felsenstein 1989).

A multiple correspondence analysis (MCA) was performed using the SAS Corresp procedure (SAS Institute 1994). Euclidean distances were calculated on a MCA coordinate matrix for all genotype pairs, and Ward's minimum variance algorithm was used to construct a dendrogram (Ward 1963). The principle of this algorithm is to cluster genotypes or groups at each step by keeping a maximum value of the ratio intergroup sum of squares/total sum of squares (Saporta 1990). Distance/similarity matrix and cluster algorithms were performed using the program developed by John Brzustowski (http://www.biology.ualberta.ca/jbrzusto/cluster.ph).

Results

Identifying molecular patterns and genetic diversity

The four primers used revealed 32 RAPD markers. The number of polymorphic bands selected as RAPD markers varied from five (primer C15) to 11 (primer A1). Based on these 32 RAPD markers, we revealed 114 different RAPD patterns for the 497 olive trees (Table 1). Under the hypothesis of non-linkage between markers, the probability of obtaining a given RAPD pattern was low and ranged from 8.1×10^{-9} for the RAPD pattern no. 113 (olive tree 13–12) to 1.8×10^{-3} for the RAPD pattern no. 10 (*Amygdalolia*). The most-distant RAPD profiles were differentiated by 14 markers (*Aglandau* and *Négrette SJB*) while the most similar ones were differentiated only by 1–3 markers. There were 5, 22 and 56 RAPD profile pairs, which were differentiated by 1, 2 and 3 markers, respectively.

SSR analysis was performed on 114 olive trees corresponding to the 114 RAPD profiles previously determined. The five primer pairs selected for polymorphism and for clear bands revealed a total of 32 alleles ranging from five at the PA(GA)2 locus to eight at the Me30Ms locus, with a mean value of 6.4 alleles per locus (Table 2). The highest frequency exceeding 36% was observed for the allele 166 at the Femsalt-4 locus and for the allele 118 at the PA(ATT)2 locus. The lowest frequency was obtained for the alleles 178 and 190 at the Femsalt-4 locus which were only observed in the *Courbeil* and *Koroneiki* varieties, respectively. Compared to other alleles with a frequency below 10%, these were considered as the only rare ones.

The observed heterozygosity was higher than the expected values under "Hardy Weinberg" equilibrium at loci Me-30-Ms, Femsalt 4 and PA(ATT)2, and lower at the PA(GA)5 locus (Table 2). However, differences between theoretical and observed values were not significant. In contrast, the level of heterozygosity was significantly lower than expected at the PA(GA)2 locus. A possible explanation of such a deficit is the occurrence of null alleles at this locus since the corresponding probability is highly significant (Table 2).

Based on 32 alleles revealed by five SSR loci, we detected 99 SSR profiles. Under the hypothesis of non-linkage between loci, the probability of obtaining a given SSR pattern was very low and ranged from 1.32×10^{-9} for *Courbeil* to 4.34×10^{-4} for *VP66* (Table 3). Among the 4,851 pairwise comparisons, there were 13, 16 and 82 SSR profile pairs which were respectively differentiated by 1, 2 and 3 alleles.

Table 4 Efficiency of a primer/locus combination and probability of getting a genotype under the independence hypothesis

Combination no.	Primer/locus combination	Number of pairs[a]	Probability range of getting a genotype
1	A1	39	1.43×10^{-5}–0.012
2	A1 + Me30Ms	4	2.65×10^{-7}–0.002
3	A1 + Me30Ms + PA(GA)5	2	1.96×10^{-8}–4.2×10^{-4}
4	A1 + Me30Ms + PA(GA)5 + A9	0	9.31×10^{-11}–4.57×10^{-5}
5	A1 + Me30Ms + PA(GA)5 + A9 + PA(GA)2 + C15	0	1.4×10^{-13}–3.91×10^{-6}
6	A1 + Me30Ms + PA(GA)5 + A9 + PA(GA)2 + C15 + PA(ATT)2 + Femsalt4 + C9	0	2.52×10^{-16}–1.58×10^{-8}

[a] Indistinguishable pairs of genotypes

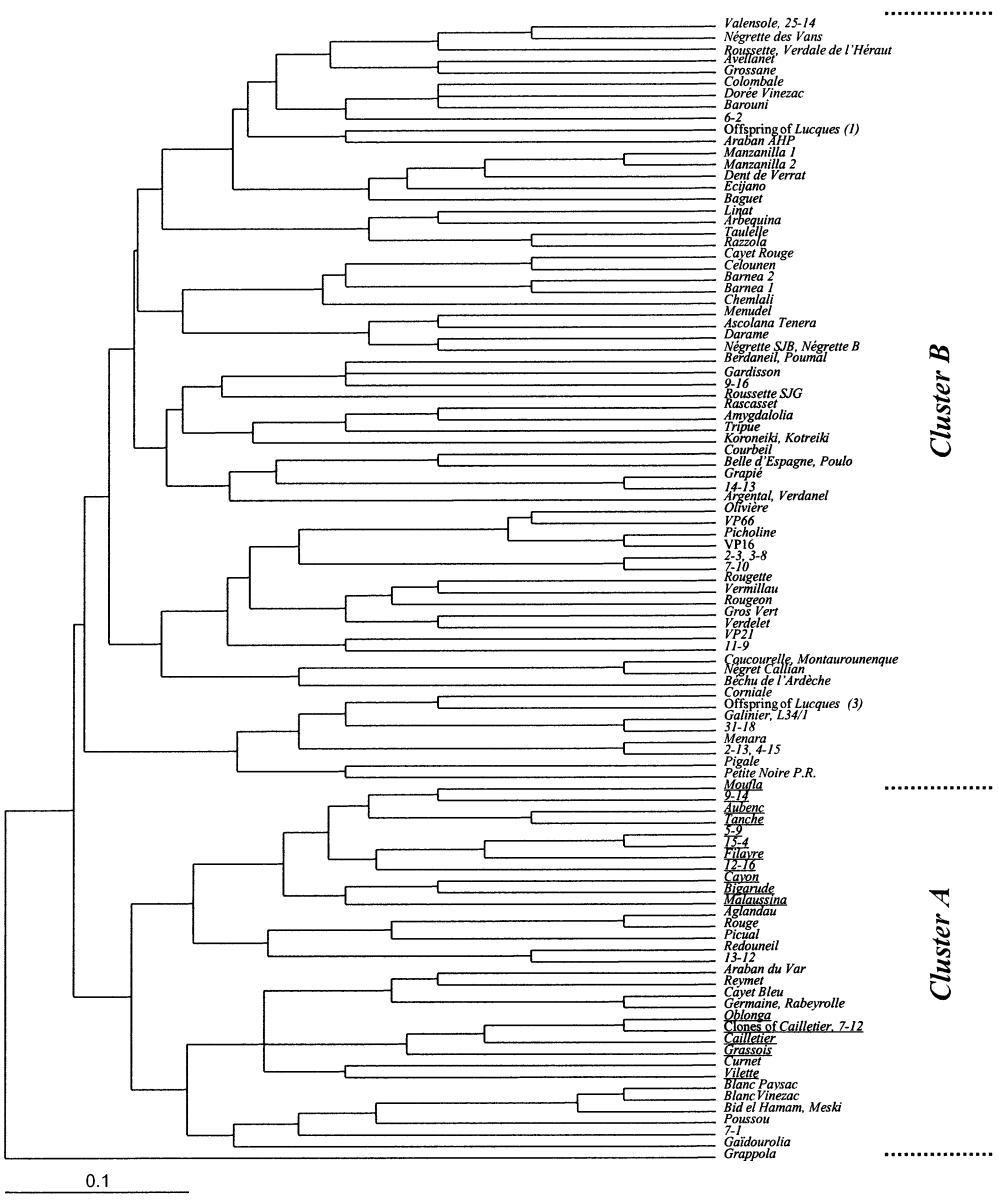

Fig. 2 Relationships between 99 olive genotypes corresponding to the olive germplasm of CBNMP. The phenogram is based on UPGMA cluster analysis of the genetic distance, defined as 1 – proportion of shared alleles. *Underlined genotypes* were grouped in clusters 2 and 3 defined by FCA analysis

Cultivar characterisation

Among the 114 RAPD profiles previously determined, 87 were confirmed by SSR analysis while the 27 remaining ones were classified into 12 SSR profiles. Four pairs of olive cultivars (*Poulo*/*Belle d'Espagne*, *Coucourelle*/*Montaurounenque*, *Argental*/*Verdanel* and *Germaine*/*Rabeyrolle*) which were different by 2 to 5 RAPD markers were not distinguished in the SSR analysis.

Five cases of synonymy and two cases of homonymy were identified (Table 1). Olive trees under the *Barnea* denomination were classified into two molecular profiles

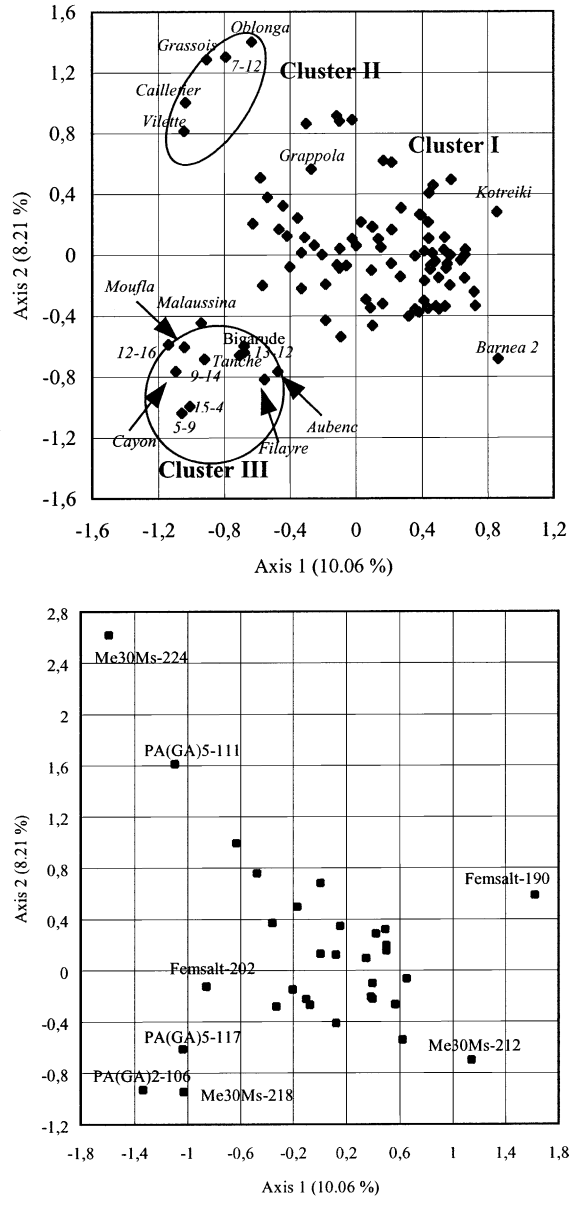

Fig. 3 Relationships between the SSR profiles corresponding to the olive germplasm of CBNMP. The Factorial Correspondence Analysis coordinate is plotted for the first and the second axis. Genotypes are represented in **A**, and alleles in **B**

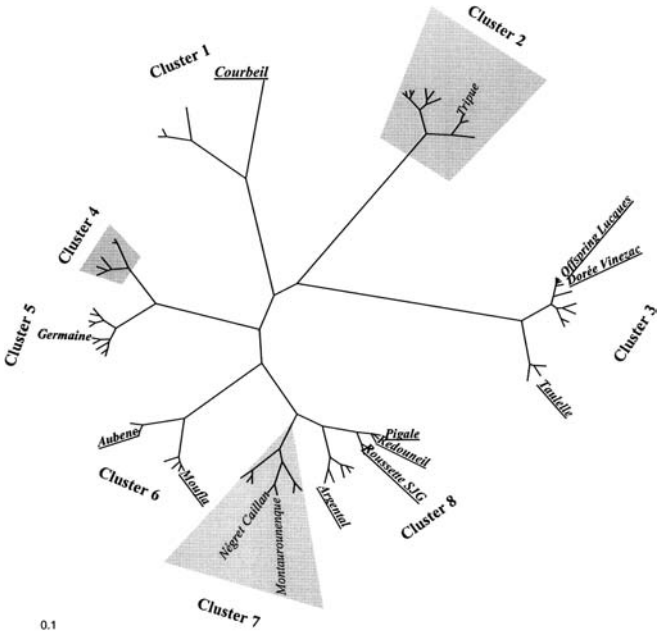

Fig. 4 Unrooted tree based on the Euclidean distance and the minimum variance algorithm (Ward 1963). The *Germaine* cultivar is from Corsica. Clusters in *grey* and cultivars in *underlined bold type* are from the West side of the Rhône Valley. The others are from the East side of the Rhône Valley

which were differentiated by five RAPD markers and two SSR alleles. *Manzanilla* accessions were classified into two molecular profiles which were differentiated by two RAPD markers and one SSR allele (Table 1).

Finally, RAPD and SSR analysis allowed us to classify the CBNMP olive germplasm into 76 molecular profiles corresponding to well-defined varieties, whereas for the remaining molecular profiles, no clear variety assignation was possible. These profiles correspond to accessions considered as clones of *Cailletier* and *Picholine*, and undetermined olive genotypes (Table 1).

To examine the ability of genotype discrimination of the molecular markers used in this study, the discriminating power was computed for each RAPD primer and SSR locus. According to the discriminating power which varied from 0.993 to 0.709 for RAPD primers and from 0.936 to 0.825 for SSR loci (Table 3), six combinations of primers were tested. Starting with the most discriminating, one or two primers were added for each successive combination (Table 4). All the 114 olive genotypes as defined by RAPD analysis were discriminated based on the fourth combination, with a probability of obtaining a given genotype lower than 5×10^{-5} (Table 4). This probability was lowered using further markers as shown by the use of combinations no. 5 and 6 (Table 4).

Genetic relationships between olive cultivars

The average similarity between all studied genotypes is 40% of shared SSR alleles. The phenogram showed two major clusters with the Italian cultivar, *Grappola*, forming an outgroup with only 23.3% of shared alleles (Fig. 2). We noted 13 pairs of genotypes which are most similar (distinct by only one allele) corresponding to some pairs of French cultivars (e.g. *Blanc Payzac* and *Blanc Vinezac*), foreign cultivars (e.g. *Manzanilla* 1 and 2) and genotypes without denomination (e.g. 5–9 and 15–4; Fig. 2).

The plot of Factorial Correspondence Analysis coordinates for the first and the second axes, which explain 10.06% and 8.21% of variance, respectively, showed three clusters of genotypes with a major one with 82 genotypes (Fig. 3A). This structure is defined by two al-

leles at two SSR loci [Me30Ms-224 and PA(GA)5-111] for cluster II and by four alleles at four different loci [Femsalt-202, PA(GA)5-117, PA(GA)2-106 and Me30Ms-218; Fig. 3B] for cluster III. These alleles correspond to the lowest frequency ones. French and foreign cultivars as well as undetermined accessions were grouped in each cluster. Clusters II and III defined by the FCA analysis (Fig. 3A) were found grouped with other accessions in cluster A defined by the UPGMA analysis based on the shared-alleles distance (Fig. 2). Similar results were obtained using the Euclidean distance and the minimum variance algorithm (Ward 1963; data not shown).

All olive genotypes previously defined by RAPD and SSR analyses were analysed by RFLP of mitochondrial DNA (Table 1). Among the 114 RAPD profiles, we noted 95, 3, 10 and 6, displaying ME1, ME2, MOM and MCK mitotypes respectively (Table 1).

Figure 4 shows eight clusters of French olive accessions with cluster 1 taken as an outgroup since it corresponds to a distinct group in the FCA analysis (cluster II in Fig. 3). Figure 4 shows that olive cultivars belonging to the clusters 2, 4 and 7 are growing on the West side of the Rhône Valley except for *Montaurounenque* and *Négret Callian* from the Var, and *Tripue* from the Alpes Maritimes (Fig. 1). Cultivars grouped in clusters 3, 5, 6 and 8 are growing on the East side except those originating from Ardèche (*Dorée Vinezac*, *Taulelle* and *Aubenc*), Gard (*Argental*) and from Hérault (*Pigale*, *Roussette SJG* and *Moufla*). We tested for the genetic differentiation between genotypes located East and West of the Rhône Valley using the Weir and Cockerham F_{st} (1984). No significant differentiation was detected ($F_{st} = 0.023$).

Discussion

The results obtained in this work show that RAPD markers can be effectively used to genotype a collection such as the CBNMP one. In fact, the probability of obtaining a given genotype under the independence hypothesis of markers is below 5×10^{-3} and only five RAPD profile pairs were differentiated by a single marker. Moreover, the RAPD primers used in this work were previously selected for their polymorphism and the clarity of their electrophoretic profiles, and they were therefore proposed as the optimal combination for efficient characterisation of olive cultivars taking into account the Mediterranean cultivar genetic diversity (Besnard et al. 2001a).

Using SSR markers developed in *F. excelsior* L. (Brachet et al. 1999; Lefort et al. 1999) and on *P. angustifolia* L. (Saumitou-Laprade et al. 2000), we demonstrated that SSR technology can be successfully used in olive cultivar characterisation without the laborious and expensive development of markers (construction of a genomic library, clone sequencing, design of primers, selection of informative pairs of primers). The SSR loci used in this work were carefully selected for their polymorphisms among nine pairs of primers. Moreover, Mendelian segregation was confirmed in the *Olivière* × *Arbequina* population for some of these loci and they were easily placed on the olive genetic maps (R. de la Rosa et al., submitted). For the PA(GA)2 locus, the occurrence of null alleles strongly suggested by the significant heterozygosity deficit was not verified throughout the examination of segregation in the *Olivière*× *Arbequina* population. Nevertheless, the assumption of null alleles is plausible because the locus PA(GA)2 was isolated and characterised in *P. angustifolia* (Saumitou-Laprade et al. 2000), and the failure of amplification due to variation in primer sequence could occur between the genera *Phillyrea* and *Olea*. Except for the PA(GA)2 locus, heterozygosity at the other loci ranging from 0.707 to 0.929 was higher than values observed at most of the SSR loci developed in olive (Rallo et al. 2000; Sefc et al. 2000). It was also higher than values obtained at the SSR loci developed in peach and used in apricot (Hormaza 2002). The portability of microsatellites between species has been mostly successful at the genus level as in *Brassica* (Szewc-McFadden et al. 1996), *Actinidia* (Weising et al. 1996) and *Prunus* (Downey and Iezzoni 2000; Hormaza 2002). There are very few reports on the use of SSR loci across genera in the same family, like the family Rosaceae (Yamamoto et al. 2001). Our study demonstrates the portability of SSR loci across genera in the family Oleaceae.

Although the low probability of obtaining a given genotype under the independence between markers hypothesis (below 5×10^{-4}), SSR analysis did not distinguish all the 114 genotypes determined by RAPD analysis since 27 RAPD profiles were classified into 12 SSR profiles. These results could be explained by the limited number of SSR loci examined despite their high polymorphism. In fact, the occurrence of a null allele at the PA(GA)2 locus limited the genetic information, and 13 SSR profile pairs were differentiated by only one allele. We used these few SSR loci because specific markers developed in olive (Rallo et al. 2000; Sefc et al. 2000) had not yet been published during our study. The use of more SSR loci should allow us to distinguish all the genotypes present in the olive CBNMP collection and to validate our molecular characterisation. However, according to the discriminating power, we demonstrated that the combination of RAPD and SSR markers is an efficient tool for genotyping the olive CBNMP collection and could be valid to distinguish other accessions which can be introduced into the collection.

Beyond the characterisation of all genotypes present in the collection, varietal identification for each molecular profile remains unresolved. Denominations reported in Table 1 correspond to accessions under the same denomination displaying an identical molecular pattern. However, due to mislabelling, and the synonymy and homonymy cases occurring in the collection, we cannot guarantee that all these denominations are legitimate. Comparing the molecular profiles of the 25 cultivars analysed both in this study and by Besnard et al. (2001a), we noted the same genotype in only 18 cases

(Table 1). The same genotype was verified for *Menara* in this work and *Picholine marocaine* (Besnard et al. 2001a), confirming that *Menara* is a selected clone (Khadari and Bervillé 2001). As we determined homonymy cases, an identical genotype was obtained for *Barnea* 2/*Barnea* and *Manzanilla* 1/*Manzanilla*. The homonymy observed for *Manzanilla* is probably the result of the polyclonal constitution of the cultivar. Among the most-similar French cultivars, *Blanc Payzac* and *Blanc Vinezac* are growing in the same area (Ardèche, Fig. 1), suggesting that the two cultivars have been selected from the same population as would be the case for *Manzanilla*. Similar results were obtained for some undetermined genotypes which are closely related to well-identified cultivars (i.e. 2–13, 4–15 and *Menara*, Fig. 3). Consequently, such results could be considered as a first basis to identify accessions without passport data present in the CBNMP collection.

For the cultivars *Cailletier*, *Germaine*, and *Verdelet*, genotypes obtained by Besnard et al. (2001a) were different from those determined in this work. For the cultivar *Cailletier*, several trees originating from different nurseries and from the olive CBNMP collection displayed the same molecular profile (Khadari et al. 2001), indicating that the accession analysed by Besnard et al. (2001a) was probably wrongfully assigned to that cultivar. However, for the *Germaine* and *Verdelet* cultivars, there is no objective criterion to determine which genotype typifies the cultivar. This question is also relevant for all other genotypes present in the collection. The crucial question is: how should a genotype be chosen as the reference genotype of a given olive cultivar? When several olive trees, grouped under the same denomination presenting similar morphological characters and originating from different collections, nurseries and orchards, display the same molecular pattern, then this genotype could be considered as the reference genotype for the cultivar. This approach was proposed by Khadari et al. (2001) and will be applied in order to establish the reference genotype for each of the cultivars present in the CBNMP olive germplasm.

Despite the occurrence of some grouped French cultivars, no clear distinction between French and foreign cultivars was obtained since foreign cultivars and undetermined genotypes were grouped with some French cultivars. Based on the RAPD, and a Factorial discriminant, analysis, Besnard et al. (2001b) showed no significant difference between Italy, Sicily, Greece and France despite the regional differentiation of Mediterranean olive cultivars. Our results are in agreement with this previous study.

Previous Phylogeographic studies demonstrated that the MOM and MCK mitotypes are specific to the West Mediterranean, while ME1 and ME2 are characteristic of the East Mediterranean populations (Besnard and Bervillé 2000; Besnard et al. 2002). In the CBNMP olive collection, 86% of the genotypes presented the East Mediterranean cytoplasm, indicating that most olive accessions originated from the East Mediterranean or were hybrid forms between East and West Mediterranean cultivars. Thus, according to mitochondrial information, there is a low contribution of Western populations to the French cultivated germplasm.

Different maternal lineages, and no distinction between foreign cultivars, indicate that French accessions correspond to a diversified olive germplasm. Investigation of genetic structure is of great importance for genetic resources management. Using FCA coordinates of SSR data, we constructed a phenogram based on Euclidean distance and the minimum variance algorithm. In comparison to other phenetic analyses, the advantage of this method is to eliminate redundant information related to the linkage of some markers since it is based on coordinates of multivariate analysis. No clear differentiation between cultivars growing on the West and the East side of the Rhône Valley was observed because some cultivars from areas bordering the Rhône Valley, like Ardèche and Gard (Fig. 1), obscure the East–West structure. Nevertheless, we noted a tendency indicating a restricted cultivation area of the major olive cultivars. These results are in agreement with patterns of traditional and ancient olive cultivation in Southern France.

Conclusion

Our study shows that the use of molecular markers is efficient for olive germplasm management, including the characterisation of accessions and the establishment of genetic relationships between cultivars in the CBNMP olive collection. Beyond this identification, we constructed a molecular data base that can be used to make a reference collection of French olive germplasm by comparing the molecular pattern of each identified accession with samples from different areas. Based on SSR analysis and RFLP of mitochondrial DNA, we showed that French olive germplasm was diversified with a major part of the maternal lineages, indicating an East Mediterranean origin or hybrid forms between the East and the West Mediterranean gene pools. Introducing new accessions by prospecting in different French areas is currently in progress. The choice of which new accessions enter into the CBNMP collection can now take into account our results, both by avoiding duplicates and also by maximising genetic diversity. The molecular pattern of a new accession can be compared to the available data base in order to avoid genotype redundancy. To enlarge the genetic diversity, prospecting and collecting new plant material should focus on the MOM and MCK mitotypes since such cultivars were originally selected in the West Mediterranean and are under-represented in French germplasm. Our study shows that the use of molecular markers is essential during all steps of germplasm management as in the CBNMP olive collection.

Acknowledgements The authors are grateful to M. Tersac and P. Villemur for discussions and to F. Boillot for support and encouragement. They particularly thank P. Baradat, B. Fady, L. Hagen and F. Kjellberg for their comments on our work. Financial

support for this work was provided by CBNM Porquerolles and ONIOL (Office National Interprofessionnel des Oleagineux, Proteagineux et Cultures Textiles).

References

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (1996) Genetix, logiciel sous Windows TM pour la génétique des populations. Version 3.0

Besnard G, Bervillé A (2000) Multiple origins of the Mediterranean olive deduced from mitochondrial DNA polymorphisms. C R Acad Sci Sér III 323:173–181

Besnard G, Khadari B, Villemur P, Bervillé A (2000) Cytoplasmic male sterility in the olive (*Olea europaea* L.). Theor Appl Genet 100:1018–1024

Besnard G, Breton C, Baradat P, Khadari B, Bervillé A (2001a) Cultivar identification in olive (*Olea europaea* L.) based on RAPDs. J Am Soc Hort Sci 126:668–675

Besnard G, Baradat P, Bervillé A (2001b) Genetic relationships in the olive (*Olea europaea* L.) reflect multilocal selection of cultivars. Theor Appl Genet 102:251–258

Besnard G, Khadari B, Baradat P, Bervillé A (2002) *Olea europaea* phylogeography based on chloroplast DNA polymorphism. Theor Appl Genet 105 (in press)

Brachet S, Jubier MF, Richard M, Jung-Muller B, Frascaria-Lacoste N (1999) Rapid identification of microsatellite loci using 5′anchored PCR in the common ash *Fraxinus excelsior*. Mol Ecol 8:160–163

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5:453–455

Downey SL, Iezzoni AF (2000) Polymorphism DNA markers in black cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach, and sour cherry. J Am Soc Hort Sci 125:76–80

Felsenstein J (1989) Phylip – Phylogeny Inference Package. Cladistics 5:164–166

Hormaza JI (2002) Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. Theor Appl Genet 104:321–328

Khadari B, Bervillé A (2001) Genetic diversity of Moroccan cultivated olive using RAPD markers. Acta Hort 546:439–442

Khadari B, Moutier N, Dosba F (2001) Approche moléculaire de la caractérisation des variétés françaises d'olivier: construction d'une base de données de génotypes de référence. Olivae 87:29–32

Lefort F, Brachet S, Frascaria-Lacoste N, Edwards J, Douglas GC (1999) Identification and characterization of microsatellite loci in ash (*Fraxinus excelsior* L.) and their conservation in the olive family (Oleaceae). Mol Ecol 8:1075–1092

Minch E (1997) MICROSAT, version 1.5 b. Stanford University Medical Center, Stanford, California

Quillet MC, Madjidian N, Griveau Y, Serieys S, Tersac M, Lorieux M, Bervillé A (1995) Mapping genetic factors controlling pollen viability in an interspecific cross in *Helianthus* sect. *Helianthus*. Theor Appl Genet 91:1195–1202

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

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Saporta G (1990) Probabilité, Analyse des Données et Statistique. TECHNIP (ed), Paris, France

SAS Institute (1994) SAS/STAT, User's guide, vol 2. SAS Institue, Cary, North Carolina

Saumitou-Laprade P, Vassiliadis C, Epplen JT, Hardt C (2000) Isolation of microsatellites loci for partenity testing in *Phillyrea angustifolia* L. (Oleaceae). Mol Ecol 9:112–114

Schoen DJ, Brown AHD (1993) Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. Proc Natl Acad Sci USA 90:10,623–10,627

Sefc KM, Lopes MS, Mendonça D, Rodrigues Dos Santos M, Laimer Da Câmara Machado M, Da Câmara Machado A (2000) Identification of microsatellite loci in olive (*Olea europaea*) and their characterization in Italian and Iberian olive trees. Mol Ecol 9:1171–1193

Szewc-McFadden AK, Kresovich S, Bliek SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated *Brassica* species. Theor Appl Genet 93:534–538

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

Virk PS, Newbury HJ, Jackson MT, Ford-Lloyd BV (1995) The identification of duplicate accessions within a rice germplasm collection using RAPD analysis. Theor Appl Genet 90: 1049–1055

Yamamoto T, Kimura T, Sawamura Y, Kotobuki K, Ban Y, Hayashi T, Matsuta N (2001) SSRs isolated from apple can identify polymorphism and genetic diversity in pear. Theor Appl Genet 102:865–870

Ward JH Jr (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. Evolution 38:1158–1370

Weising K, Fung RWM, Keeling DJ, Atkinson RG, Gardner RC (1996) Characterisation of microsatellites from *Actinidia chinensis*. Mol Breed 2:117–131