

# Identification of the Worldwide Olive Germplasm Bank of Córdoba (Spain) using SSR and morphological markers

Isabel Trujillo · Maria A. Ojeda · Nieves M. Urdiroz · Daniel Potter ·  
Diego Barranco · Luis Rallo · Concepcion M. Diez

Received: 31 October 2012 / Revised: 23 September 2013 / Accepted: 25 September 2013 / Published online: 13 October 2013  
© Springer-Verlag Berlin Heidelberg 2013

**Abstract** Olive is one of the most ancient crop plants and the World Olive Germplasm Bank of Cordoba (WOGBC), Spain, is one of the world's largest collections of olive germplasm. We used 33 SSR (Simple Sequence Repeats) markers and 11 morphological characteristics of the endocarp to characterise, identify and authenticate 824 trees, representing 499 accessions from 21 countries of origin, from the WOGBC collection. The SSR markers exhibited high variability and information content. Of 332 cultivars identified in this study based on unique combinations of SSR genotypes and endocarp morphologies, 200 were authenticated by genotypic and morphological markers matches with authentic control samples. We found 130 SSR genotypes that we considered as molecular variants because they showed minimal molecular differences but the same morphological profile than 48 catalogued cultivars. We reported 15 previously described and 37 new cases of synonyms as well as 26 previously described and seven new cases of homonyms. We detected several errors in accession labelling, which may have occurred at any step during establishment of plants in the collection. Nested sets of 5, 10 and 17 SSRs were proposed to progressively and efficiently identify all of the genotypes studied here. The study provides a

useful protocol for the characterisation, identification and authentication of any olive germplasm bank that has facilitated the establishment of a repository of true-to-type cultivars at the WOGBC.

**Keywords** *Olea europaea* L · Genetic diversity · Microsatellite · Ex-situ conservation · Clonal cultivars

## Introduction

The olive tree (*Olea europaea* L.) is currently the most widely cultivated temperate fruit crop in the world and is characterised by an extensive legacy of clonally propagated traditional cultivars (Bartolini and Petrucelli 2002; FAO 2008; Rallo 2005). This great genetic diversity is the result of empirical and local selection of exceptional trees since the olive was domesticated about 6,000 years ago in the Middle East (Zohary and Spiegel-Roy 1975). Commercial shipping and human migrations spread the olive westward across the Mediterranean Basin, leading to complex genetic relationships among cultivars (Besnard et al. 2013).

In the last 20 years, important socioeconomic changes in many Mediterranean countries have driven significant technological improvements in olive cultivation. These changes are increasing the risk of genetic erosion of olive germplasm because local traditional cultivars are being replaced by a few cultivars that are suitable for the new mechanically harvested plantations. Therefore, the identification and conservation of traditional olive cultivars are currently high-priority tasks that are needed to ensure the sustainable use of those cultivars in the future (Rallo et al. 2013).

Germplasm banks are facilities that are designed to achieve this goal by providing characterisation and long-term “ex situ” conservation of genetic resources. Clonally propagated fruit crops such as olive are typically conserved in “live

---

Communicated by G. G. Vendramin

**Electronic supplementary material** The online version of this article (doi:10.1007/s11295-013-0671-3) contains supplementary material, which is available to authorized users.

---

I. Trujillo (✉) · M. A. Ojeda · N. M. Urdiroz · D. Barranco ·  
L. Rallo · C. M. Diez

Departamento de Agronomía, Campus Universitario de Rabanales,  
Edificio Celestino Mutis. Ctra. Madrid-Cádiz, Km. 396,  
14014 Córdoba, Spain  
e-mail: ag2tmai@uco.es

D. Potter  
Department of Plant Sciences Mail Stop 2, University of California,  
One Shields Avenue, Davis, CA 95616, USA

collections”, which are suitable selected field plantations where the crop can fulfil its normal biological cycle (van Hintum et al. 2000). Almost 100 collections of olive genetic resources have been established, both in the Mediterranean Basin and in new olive-growing regions of the world, such as North and South America, Australia, China and South Africa (Bartolini et al. 2005). The World Olive Germplasm Bank of Cordoba (WOGBC) was established in 1970 and it is currently one of the largest olive germplasm banks; WOGBC contained 499 accessions from 21 countries at the time that this study was undertaken (Caballero et al. 2006). In 2003, a second world olive germplasm bank was established at the experimental orchard of Tessaout, in Marrakech, Morocco. This bank contains olive cultivars from other collections, such as the WOGBC, as well as local genetic resources (Haouane et al. 2011).

Initially, olive cultivars were named for their outstanding morphological traits or utility of production. Denominations are also frequently based on the locality of origin of the propagating material (Rallo 2005). Consequently, synonymies (different names for the same cultivar) and homonyms (the same name for different cultivars) are extremely frequent among and within olive-growing countries (Barranco et al. 2000a). Additionally, the occurrence of clonal mutations, which may or may not have some phenotypic expression, make the characterisation of olive cultivars a challenging process that requires experience in both morphological and molecular identification (Cantini et al. 2008; Corrado et al. 2009; Díez et al. 2012).

To ensure an efficient conservation of olive genetic resources, the accessions of the WOGBC have been characterised using both, morphological and molecular markers. Initially, morphological markers were used to characterise the accessions by applying the morphological scheme proposed by Barranco and Rallo (1984). Later, a simplified scheme proposed by Barranco et al. (2000a) was adopted as reference by the International Union for the Protection of New Varieties of Plants (1991). This morphological scheme allowed for the identification of 272 different cultivars from Spain (Barranco et al. 2005). Additionally, various molecular markers including isozymes (Trujillo et al. 1995). RAPD (Belaj et al. 2002, 2003a, b) and SSR markers, have been applied to complete the morphological descriptions. SSR markers became the marker of choice for the identification of the entire collection after study their performance in the collection (Belaj et al. 2003c) and following successful experiences in other fruit crops, such as grape, sweet cherry or pear (Bowers et al. 1996; Kimura et al. 2002; Wünsch and Hormaza 2002).

The WOGBC has paid special attention to the authentication of olive cultivars. The concept of cultivar authentication has primary been used in the context of modern food technology to guarantee that the commercial, edible product matches

the cultivar specified on the label (Downey and Boussion 1996; Melchiade et al. 2007; Mouly et al. 1997). Similarly, we considered an accession of the WOGBC to be authentic if it matched samples coming from the area of origin of the putative cultivar to which it belongs. Therefore, authentication guarantees that a cultivar distributed worldwide corresponds to the original cultivar growing in its area of origin. The conservation of a reference collection of endocarps has proved to be a useful tool for this purpose. Unfortunately, authentication is a pending task in most olive collections. Nevertheless, authentication should be an essential pre-requisite for exchanging plant material among collections, researchers and the nursery industry to avoid the extended confusion among denominations and true-to-type cultivar names reported in most olive cultivar collections around the world (Bartolini et al. 2005).

This is the first study describing the complete process of characterisation, identification and authentication of a world olive germplasm collection. We used 33 select SSRs complemented by the use of morphological markers to achieve three main goals: (a) the characterisation, identification and authentication of the olive cultivars in the germplasm collection; (b) the establishment of a consistent, easy to manage and affordable protocol for the management of olive germplasm banks; and (c) the detection of common errors arising from any misnaming and mislabelling that occurs during the processes of sending, receiving or propagating plants.

In summary, this study attempts to ease and encourage the labour of other olive germplasm banks by describing a management pipeline to efficiently preserve olive genetic resources. To support this overarching goal, we also provide the SSR profiles and endocarp morphological descriptions of the cultivars analysed herein.

## Materials and methods

Several common terms related to the management of a germplasm bank could result in confusion for readers unfamiliar with this field. For this reason, we have included a glossary with critical terms used in this study (Online resource 1).

## Plant material

We studied 824 trees belonging to 499 accessions of cultivated olive (*O. europaea* L.) from 21 countries of origin (Table 1). These accessions are conserved in live collections in the WOGBC located at the Instituto de Investigacion y Formacion Agraria y Pesquera de Andalucía centre “Alameda del Obispo”, Cordoba (Southern Spain). Each accession was given a unique identifier composed of the letters

**Table 1** The number of olive accessions and trees analysed per country and the number of genotypes and cultivars that were identified and authenticated after the identification process

Origin	No. of accessions	No. of trees	No. of genotypes	No. of identified cultivars	No. of authentic cultivars
Albania	12	26	11	10	3
Argentina	2	2	1	2	1
Chile	1	1	1		
Cyprus	3	6	2	1	1
Algeria	2	3	1	1	
Egypt	5	8	3	3	1
France	10	18	10	8	6
Greece	18	34	18	15	7
Croatia	7	8	7	7	3
Iran	5	10	5	5	5
Israel	9	16	3	3	2
Italy	36	65	30	20	17
Lebanon	2	2	2	1	1
Morocco	4	4	5	1	1
Mexico	7	8	2	2	
Portugal	10	16	8	6	4
Spain	279	468	239	186	134
Syria	56	87	41	37	1
Tunisia	7	7	6	7	7
Turkey	19	24	17	14	7
USA	4	7	2	2	
Unknown <sup>a</sup>	1	4	1	1	
Total	499	824	411 <sup>b</sup>	332	200

<sup>a</sup> The origin was unknown in four trees belonging to one accession labeled as “Unknown”

<sup>b</sup> Four genotypes coming from different countries were duplicated

COR plus six digits corresponding to the chronological arrival of the sample to the WOGBC (i.e. COR000789).

For 260 accessions, only one tree was available for analysis, while each of the remaining 239 accessions from two to eight trees were analysed to confirm their identity (Online resource 2). All the accessions were planted on their own roots except four accessions that were grafted onto the cv. Oblonga.

### DNA extraction and amplification

Total genomic DNA was extracted from fresh leaves using the CTAB method described by de la Rosa et al. (2002). DNA quality and quantification were assessed by electrophoresis on 0.8 % (w/v) agarose gels.

A set of 33 SSR markers were selected based on their polymorphism level, PCR amplification reproducibility and

easy interpretation among 77 SSR markers developed for the olive (Carriero et al. 2002; Cipriani et al. 2002; de la Rosa et al. 2002; Gil et al. 2006; Sefc et al. 2000). The selection of the most suitable set of SSR markers was carried out by testing their performance (clear amplification, easy interpretation and polymorphism) on a set of 48 different cultivars from the WOGBC. These cultivars were selected according to their variability and representativeness of the original collection (Díez et al. 2012).

Differences of 1 bp between alleles were checked by re-amplification to establish whether a coding error had occurred. Whenever present, replicates of the same accessions were compared. In cases of mismatch, both replicates were analysed again using newly collected plant material. These rare differences were considered correct SSR profiles if the differences were confirmed in the second amplification.

The SSR amplification was performed in a total volume of 20 µl, containing 2 ng of genomic DNA, 1× supplied PCR buffer (Biotools, Spain), 200 µM of each dNTP (Roche), 0.25 units of Taq DNA polymerase (Biotools, Spain) and 0.2 µM of forward (fluorescently labelled) and reverse primers. The PCR reactions were carried out on a thermal cycler (Perkin-Elmer-9600) using the following program: denaturation at 94 °C for 5 min, 35 cycles of 94 °C for 20 s, 50 °C for 30 s and 72 °C for 30 s and a final extension at 72 °C for 7 min. Detection of amplification products was carried out with an automated sequencer ABI 3130 Genetic Analyser (Applied Biosystems/HITACHI) using the internal standard GeneScan 400 HD-Rox. Two cultivars, Arbequina and Frantoio, were used as controls in all runs.

### Data analysis

#### Genetic characterisation by SSR markers

The allele profiles were sized (basepairs) and characterised using Genescan 3.7 (Applied Biosystems). The following parameters were calculated for each SSR locus using the Power Marker V3.23 (Liu and Muse 2005) software package: average number of alleles; number of alleles making up each genotype; number of unique alleles; observed heterozygosity (Ho); expected heterozygosity (He) and polymorphism information content (PIC) (Botstein et al. 1980). Null allele frequency per locus was tested using the program Cervus v 2.0 (Marshall et al. 1998).

The genotypes were discriminated by the pair-wise comparison of their SSR profiles using “Excel Microsatellite Toolkit” (Park 2001). To evaluate the genetic relationships among the different genotypes, a matrix containing only the different SSR profiles was built, with amplified alleles scored as present (1) or absent (0). This matrix was used to perform a cluster analysis based on the unweighted pair group method

with arithmetic mean (UPGMA) algorithm using Dice's similarity index (Dice 1945) implemented in the statistical software NTSYS-PC v2.02 (Rohlf 1998).

The characterisation of the accessions by SSR markers (Online resource 2) was complemented by the use of morphological markers to (a) authenticate the putative cultivars; (b) check whether small allelic differences between pairs of genotypes led to differential phenotypic expression in the selected endocarp traits; (c) confirm cases of synonymy; and (d) detect possible errors during the propagation and establishment of the collection.

### Morphological characterisation

The morphological characterisation was independently carried out by three trained observers, using a representative sample of 40–50 endocarps per tree and during at least 2 years. We evaluated a minimum of 11 characteristics of the endocarp included in the pomological scheme developed and described by Barranco et al. (2000a, 2005): weight, shape in position A, symmetry in positions A and B, position of maximum transverse diameter in position B, shape of apex in position A, shape of base in position A, roughness of surface, number of grooves on basal end, distribution of the grooves on basal end and presence of mucro (Online resource 3). The morphological profile of each sample was the combination of its level of expression for each one of the 11 endocarp traits that were evaluated. We confronted the morphological profiles of the samples conducting pair-wise comparisons between them. The traits of endocarps are the most discriminating and stables ones, while other characteristics, such as those of the fruit are more influenced by environmental conditions. Moreover, endocarps may also be conserved for a long time and they are easily exchanged among collections. For these reasons, the description of the endocarp has been frequently used to catalogue olive cultivars (Barranco et al. 2000a, 2005; Fendri et al. 2010; D'Imperio et al. 2011)

### Authentication and denomination of the cultivars

The identified cultivars were authenticated by their comparisons with control samples of endocarps coming from the corresponding countries and areas of origin. Recently, a comparison with control DNA samples has been added to the authentication process to complement the morphological control. These control samples are part of the reference collection, which progressively is being established in the Department of Agronomy at the University of Cordoba, Spain. The accessions involved in cases of synonymy and homonymy as well as those whose SSR profiles did not match any cultivar present in the WOGBC, were re-named following previously described criteria (Barranco and Rallo 1984; Barranco et al. 2000a, 2005; Caballero et al. 2006) (Online resource 2).

## Results

### Overall genetic diversity and nested sets of SSRs for identification purposes

A total of 466 alleles were amplified from the entire collection, of which 67 were unique alleles (present in only a single genotype) (Table 2). The number of alleles per SSR ranged from 5 (GAPU82) to 36 (ssrOeUA-DCA10), with an average of 14.12 alleles per locus. Allele frequencies varied between 0.001 and 0.93; it is noteworthy that 254 alleles (53.3 %) showed a frequency below or equal to 0.01. The allelic differences between pairs of different genotypes ranged from 1 to 59 alleles, with an average of 40.74 alleles (Online resource 4).

The  $H_o$  varied between 0.973 (GAPU-101) and 0.168 (GAPU-11e17), with an average of 0.65 (Table 2). The  $H_e$  ranged from 0.875 (UDO99-043) to 0.324 (GAPU82), with a mean of 0.69. Twenty-six SSR markers had a PIC value higher than 0.5 (Table 2).

Despite the valuable information given by the set of 33 SSRs in terms of genetic variability, to use it for routine identifications could be quite time-consuming and expensive. For this reason, we also defined three nested sets of SSRs for identification purposes that were selected according to their discrimination capacity in the WOGBC (Fig. 1). Five SSRs (UDO99-043, ssrOeUA-DCA9, ssrOeUA-DCA16, ssrOeUA-DCA3 and GAPU101) could be used to distinguish between 79 % of the accessions. Ten SSRs (UDO99-043, ssrOeUA-DCA9, ssrOeUA-DCA16, ssrOeUA-DCA3, GAPU101, ssrOeUA-DCA11, ssrOeUA-DCA4, UDO99-005, sseOeIGP07 and GAPU89) could be applied to discriminate between 93 % of the accessions. A larger set of 17 SSRs (UDO99-043, ssrOeUA-DCA9, ssrOeUA-DCA16, ssrOeUA-DCA3, GAPU101, ssrOeUA-DCA11, ssrOeUA-DCA4, GAPU103, UDO99-005, ssrOeIGP7, GAPU89, ssrOeUA-DCA18, ssrOeUA-DCA8, ssrOeUA-DCA10, GAPU82, UDO99-042 and ssrOeUA-DCA15) was needed to discriminate between 100 % of the accessions from the WOGBC.

### Identification of the WOGBC by SSR markers

The set of 33 SSR markers amplified 411 different genotypes among the 824 trees belonging to 499 accessions analysed herein (Table 1). The process of identifying each accession is compiled in the Online resource 2; each different genotype is coded with an ordinal number (Online resources 2 and 5).

To discriminate between the 411 different genotypes, we first compared the SSR profiles of the trees within accessions and those accessions represented by only one tree. The results for the 499 accessions (Online resource 2) were distributed between three groups as follows: (a) 297 accessions gave rise to unique SSR profiles (not duplicated in any other part of the

**Table 2** Diversity parameters of the 33 SSR markers used in this study characterising the entire collection of olive germplasm: size range (base pairs), number of alleles (Na), number of unique alleles (Nu), observed (Ho) and expected (He) heterozygosity, null allele frequency (An), polymorphic information content (PIC) and whether or not the SSRs was involved in the description of molecular variants (MV)

Locus <sup>a</sup>	Size range	Na	Nu	Ho	He	An	PIC	MV
UDO99-043	162–225	24	2	0.883	0.875	−0.0077	0.861	Yes
UDO99-042	133–162	11	1	0.251	0.78	0.5192	0.745	Yes
UDO99-027	110–197	9	1	0.571	0.463	−0.1145	0.404	No
UDO99-024	164–203	14	1	0.57	0.636	0.0497	0.596	No
UDO99-017	152–173	6	–	0.77	0.783	0.0025	0.748	No
UDO99-015	100–132	8	–	0.451	0.688	0.1939	0.639	Yes
UDO99-012	152–163	6	1	0.639	0.598	−0.0401	0.556	Yes
UDO99-011	103–142	14	3	0.949	0.829	−0.0729	0.807	Yes
UDO99-005	130–145	6	–	0.792	0.598	−0.1576	0.527	Yes
UDO99-004	132–147	7	–	0.491	0.53	0.0355	0.419	No
ssrOeUA-DCA9	160–214	22	2	0.949	0.867	−0.0481	0.853	Yes
ssrOeUA-DCA8	123–168	18	3	0.617	0.805	0.1381	0.778	Yes
ssrOeUA-DCA5	191–211	10	–	0.384	0.392	0.0139	0.382	Yes
ssrOeUA-DCA4	116–198	28	2	0.644	0.813	0.1185	0.793	Yes
ssrOeUA-DCA3	227–255	15	1	0.937	0.846	−0.0546	0.825	Yes
ssrOeUA-DCA18	158–193	16	1	0.922	0.82	−0.0619	0.796	Yes
ssrOeUA-DCA16	122–228	32	12	0.961	0.848	−0.0663	0.829	Yes
ssrOeUA-DCA15	243–267	7	1	0.226	0.51	0.3792	0.458	Yes
ssrOeUA-DCA11	126–185	26	5	0.922	0.823	−0.0622	0.799	Yes
ssrOeUA-DCA10	138–260	36	4	0.204	0.812	0.6044	0.795	Yes
ssrOeUA-DCA1	204–274	14	4	0.767	0.613	−0.1289	0.555	Yes
ssrOeIGP7	120–152	12	2	0.76	0.753	−0.0139	0.719	Yes
ssrOeIGP15	177–202	7	–	0.694	0.595	−0.08	0.543	No
GAPU89	156–209	21	5	0.915	0.756	−0.1137	0.729	Yes
GAPU82	173–188	5	1	0.097	0.324	0.5422	0.291	Yes
GAPU71A	206–246	10	4	0.478	0.426	−0.0651	0.371	No
GAPU71-B	118–147	9	1	0.936	0.802	−0.0798	0.77	Yes
GAPU59	194–227	10	2	0.644	0.627	−0.0326	0.584	Yes
GAPU45	178–194	7	–	0.547	0.487	−0.0792	0.438	Yes
GAPU103	133–208	26	4	0.744	0.818	0.0433	0.795	Yes
GAPU101	183–219	13	1	0.973	0.834	−0.0807	0.811	Yes
GAPU-11e17	178–221	7	1	0.168	0.681	0.6087	0.617	No
EMO-90	181–208	10	2	0.702	0.643	−0.0537	0.607	No
<i>Mean</i>		14.12	2.58	0.65	0.69		0.65	
<i>Total</i>		466	67					

<sup>a</sup> The characteristic parameters of the initial 77 SSRs can be provided by the authors upon personal request

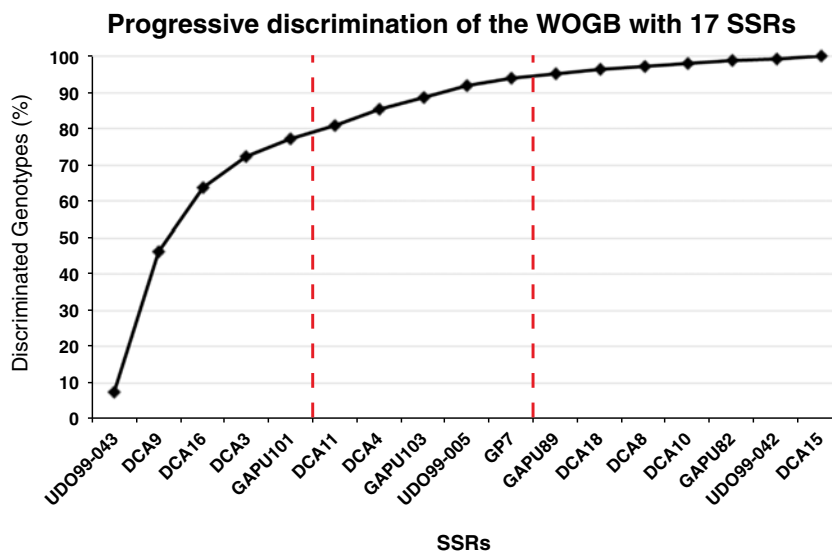
entire collection); (b) 166 accessions had SSR profiles in common with other accessions resulting in the identification of 80 different SSR profiles. For instance, we detected the same genotype among accessions coming from different countries, including Picholine Marocaine (COR00101 and COR001479) from Morocco, Mission de San Vicente (COR001133) from Mexico and Mission Nieland (COR000716) from the USA. We considered these last two accessions to be duplicated and a case of synonymy because their denominations were accepted and well known in their respective areas of origin; (c) 36 accessions accounted for more than one genotype, being some of them duplicated within this group. Therefore, this third group gave rise to 34 different SSR

profiles after pairwise comparisons (Online resource 2). We found, for example, that the accession Manzanilla Picua (COR000377) from Spain showed two different SSR profiles and that Zaity (COR000788) from Syria had three different profiles. These cases could be due to the accidental mixture of plants and mislabelling during the propagation phase.

Complementing SSRs with morphological markers to identify olive cultivars

The analysis of the entire WOGBC by 33 SSR markers was complemented with the evaluation of 11 endocarp traits,

**Fig. 1** Percentages of genotypes discriminated using an increasing number of SSR markers (one to seventeen SSRs). The red dashed lines indicate the five and ten SSRs sets to appreciate their discrimination capacity



allowing the identification of 332 different olive cultivars (Online resources 2 and 3).

The 11 endocarp traits were polymorphic, discriminating 246 different morphological profiles that were coded with an ordinal number (Online resources 1, 2 and 3). As the morphological traits were approximately half as powerful as the use of molecular markers (246 morphological vs. 411 SSRs profiles), we used the evaluation of the endocarp as a complement for the identification of olive cultivars.

We found 61 different genotypes sharing the same endocarp profile than other cultivars in the collection. In those cases, the information given by the SSRs markers was taken as main criteria to consider the genotypes to be different cultivars. We finally applied the same criteria to 25 cultivars without data for endocarp profiles until their endocarps are available in the collection.

We paid special attention to the pairs of different genotypes, which presented small allelic differences between them sharing high similarity indexes (0.9). We checked the consistency and reproducibility of these small allelic variations by re-amplification using new plant material. Up to 25 SSRs were involved in the amplification of these allelic differences, independently of their length and repetitive motif (Table 2). If the SSR variability was confirmed, we proceeded to compare the endocarps of the accessions with their closest cultivars. If no morphological differences were observed, we considered the accessions to be molecular variants, representing cases of intracultivar variability of their closest cultivar. Alternatively, if the endocarp morphology was different, we consider them to be different cultivars. In these latter cases, previous information about the morphological characterisation of other organs as well as about their agronomic performance, if available, was also taken into account to take the decision. Therefore, we found 130 genotypes that could be considered as molecular variants of 48 different cultivars because any

morphological difference in the endocarp was observed between them (Table 3; Figs. 2 and 3a). By contrast, only four pairs of cultivars (Chemlali-744-Chetoui, Azulejo-Manzanilla Cacereña, Cordovil Castello Branco-Verdial de Badajoz and Zarza-Lechín de Sevilla) showed identical or nearly identical SSR profiles but presented morphological differences (Fig. 3b). We would like to make clear that the cv. Chemlali mentioned above, is not the cv. Chemlali de Sfax, one of the most important cultivars in Tunisia. The accession CO000744 arrived at the WOGBC with the generic denomination of Chemlali, a common homonymy in Tunisia; however after the identification process, it was determined to be almost identical to cv. Chetoui, also from Tunisia. To avoid confusion between this accession and the well-known cv. Chemlali de Sfax, we added a code number to the original accession name, being from now on Chemlali-744.

#### Authentication process, synonyms and homonyms

We performed the authentication of 200 cultivars, 172 by comparisons of their endocarps with those of the same cultivar coming from its originating country and 28 by both endocarp and SSR profiles. For example, we authenticated the accessions COR000231 and COR001477, named Arbequina, by comparing their SSR profiles and endocarps with the authentic control sample of the cv. Arbequina coming from Catalonia, Spain, its natural area of origin. By contrast, 28 cultivars did not match their respective authentic control samples (Table 1; Online resource 2). The remaining 104 cultivars (31 %) could not be authenticated because their corresponding authentic reference samples of endocarps from the countries of origin were not available in our collection.

The authentic control samples were also very helpful to determine 37 new and 15 previously described, cases of synonyms among cultivars (different names for the same

**Table 3** Information on cultivars with molecular variants including their area of cultivation, the number of molecular variants (no. MV) and Dice similarity index (SI) range

Cultivar	Cultivation area	No. MV	SI range
Alameño de Montilla	Spain	3	0.95–0.97
Alfajara	Spain	2	0.97
Belluti	Turkey	2	0.98
Blanqueta	Spain	2	0.98
Bolvino	Spain	4	0.91–0.98
Callosina	Spain	2	0.98
Carolea	Italy	2	0.98
Carrasquenho de Elvas	Portugal	2	0.96
Cerezuela	Spain	2	0.97
Chalchali	Syria	4	0.93–0.98
Changlot Real	Spain	2	0.97
Cirujal	Spain	5	0.90–0.98
<i>Corbella-817</i>	Spain	2	0.98
Cordovil de Serpa	Portugal	2	0.93
Dolce Agogia	Italy	2	0.97
Frantoio	Italy	5	0.93–0.97
Gemlik	Turkey	2	0.98
Gerboui	Tunisia	2	0.98
Gordal de Granada	Spain	3	0.97–0.98
Gordal de Hellín	Spain	2	0.96
GordalSevillana	Spain	4	0.94–0.96
Grappolo	Italy	2	0.97
Hojiblanca	Spain	3	0.97–0.99
Leccino	Italy	2	0.97
Lechín de Sevilla	Spain	2	0.98
Kato Drys	Cyprus	2	0.96
ManzanillaCacereña	Spain	3	0.95–0.96
Manzanilla de Agua	Spain	2	0.96
Manzanilla de Sevilla	Spain	8	0.92–0.99
Mawi	Syria	2	0.98
Mollar de Cieza	Spain	4	0.96–0.97
Moraiolo	Italy	2	0.98
Morona	Spain	2	0.98
Morrut	Spain	2	0.97
Negrillo de Arjona	Spain	2	0.98
Ocal	Spain	4	0.96–0.99
Pavo	Spain	2	0.99
Pequeña de Casas Ibañez	Spain	2	0.98
Picholine	France	2	0.98
PicholineMarocaine	Morocco	5	0.97–0.98
Picual	Spain	3	0.97–0.98
Picudo	Spain	2	0.97
Uovo di Piccione	Italy	2	0.99
Valanolia	Greece	3	0.96–0.98
Varudo	Spain	3	0.95–0.98
Verdale	France	2	0.98
Verdial de Badajoz	Spain	4	0.97–0.98

**Table 3** (continued)

Cultivar	Cultivation area	No. MV	SI range
Verdial de Huévar	Spain	2	0.98
<i>TOTAL 48</i>		<i>130</i>	

cultivar used in different growing areas). For instance, a new synonymous group was identified that included four names of the cv. Picholine Marocaine, originally from Morocco; this group was formed by the accessions Alameño de Marchena (COR000254) and Cañivano Blanco (COR000052) from Spain, Mission Nieland (COR000716) from the USA, Haouzia (COR000835) and Menara (COR000836) from Morocco and Sigoise (COR000119) from Algeria (Table 4).

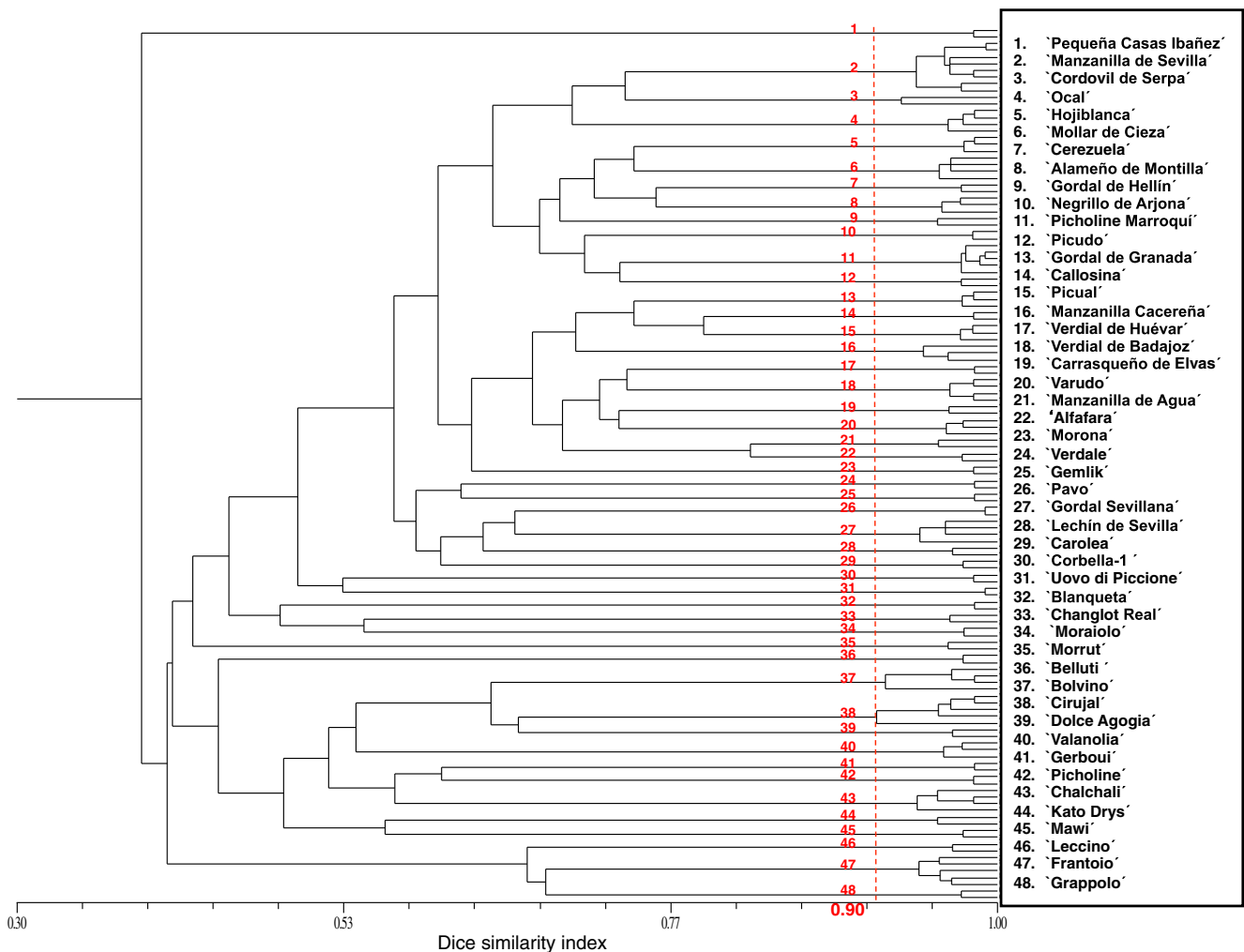
When a synonym was described, the cultivar was given the denomination that it holds in its wider and original area of cultivation. For instance, in the previously mentioned case, the name of the cultivar was Picholine Marocaine. Similarly, we previously found that Frantoio and Oblonga were synonymous denominations used for the same cultivar in Italy and the USA, respectively. Nevertheless, Frantoio was chosen, as the reference name because Italy is the putative area of origin for this cultivars and it is cultivated in a significantly greater area in Italy compared with in the USA.

Several possible synonymous cases are pending confirmation because of the lack of authentic control samples from their countries of origin, including the pairs: Abbadi-AbadiShalal, AbadiAbou Gabra-1033-Bent Al kali, Adramitini-Ayvalik, Ayrouni-Verdial de Huévar, Kokerrmadh Berati-Frantoio, Kusha-Mixani, Pecosó-Pico Limón, Torcio de Huelma-Nevado Rizado and Sevillana-Sevillana de Abla, which could be synonymous of the cultivars Bent al Kali, Ayvalik, Verdial de Huevar, Frantoio, Mixani, Pico Limón, Sevillana de Abla, respectively (Online resource 2).

Additionally, seven new cases of homonyms (the same name used for different cultivars) were discovered. For example, the denomination Toffahi, which included two different cultivars, cv. Toffahi from Egypt and cv. Toffahi-1,000 from Syria. Similarly, the denomination Trylia included two cultivars, cv. Gemlik from Turkey and cv. Trylia-992 from Syria (Table 5). The accessions that did not match any identified cultivar from the WOGBC kept their original name followed by the significant digits of their accession code at the bank. For example, the accession COR000361, named Chorro de Castro del Río, was renamed as Chorro de Castro del Río-361 after ensuring that it did not match the cv. Chorro de Castro del Río (Online resource 2).

#### Propagation errors and mislabelling

Possible errors (within and among accessions), which might occur at any step during the establishment of the plants in the



**Fig. 2** UPGMA dendrogram based on the Dice similarity index of 48 different cultivars with molecular variants. Numbers in red indicate groups of molecular variants of each cultivar whose name is specified in the rightmost box. The morphological characterisations are provided in Online resource 2

collection, were identified. For example, the accession Picual (COR000303) did not match the control cv. Picual at either



**Fig. 3** Contrasting patterns of molecular and morphological differences among accessions. **a** Endocarps belonging to three synonymous accessions (Mission Nieland, Sigoise and Menara) that exhibit subtle genetic differences (similarity index, 1–0.983) but no morphological differences from their closest cultivar “Picholine Marocaine.” **b** Endocarps belonging to cultivars “Zarza” and “Lechín de Sevilla,” which exhibited subtle genetic differences (SI=0.991) and clear morphological differences

morphological or molecular markers; however, it matched cv. Pico Limon. Similarly, the accession Hamed (COR000722) from Egypt was reliably identified as cv. Manzanilla de Sevilla from Spain and three accessions labelled as Desconocida (Unknown; COR000954, COR001481 and COR001464) were identified as cvs. Verdial de Huevar, Caballo and Uovo di Piccione (Online resource 2).

## Discussion

Overall genetic diversity and nested sets of SSRs for identification purposes

The main purpose of this paper was the identification of the WOGBC and to propose a consistent pipeline to fulfil this goal based on our experience. However, the task of a germplasm bank is not only the identification and conservation of their accessions but also the characterisation of their



**Table 4** Cases of synonyms found in the identification process at the WOGBC

Cultivar	Synonyms	Original report and comments
Alfafara	Royal de Villena	
Ayvalik	Edremit YagliK	Barranco et al. (2000a)
Azapa	Arauco	Barranco et al. (2000a)
Belluti	Yun Gelebi	
Bosana	Peranzana	Erre et al. (2010)
Cerezuela	Morejona	
Changlot Real	Torcio de Cabra	
Cordovil de Serpa	Madural	Trujillo et al. (1995)
Frantoio	Frantoio A. Corsini, Maelia, Oblonga and Razzola	Oblonga was previously described by Barranco et al. (2000b)
Gemlik	Samsun Tuzlamalik	
Gerboui	Marsaline	
Gordal de Granada	Gordalejo, Manzanilla de Jaén, Manzanilla del Centro and Nabali	
Gordal Sevillana	Santa Caterina	Trujillo et al. (1995) and Belaj et al. (2002)
Grappolo	Leccio di Corno	Belaj et al. (2002)
Kato Drys	Klirou	
Manzanilla Cacereña	Azeitera and Negrinha	Barranco et al. (2000a) and Belaj et al. (2002)
Manzanilla de Sevilla	Chesna and Redondil	Manzanilla de Sevilla and Redondil were described by Belaj et al. (2002)
Menya	Manzanilla Picua	
Mixani	Ulliri i Bardhe Berat	
Mollar de Cieza	Meloncillo, Ojúa and Verdálón	
Moraiole	Moraiole T. Corsini and Carboncella	Carboncella was previously described by (Barranco et al. 2000a)
Morona	Llorón de Ronda	
Ocal	Gordal de Archidona and Verdál de Alhama	
Rechino	Palomillo	
Picholine Marocaine	Alameño de Marchena, Cañivano Blanco, Haouzia, Menara, Mission Nieland and Sigoise	Cañivano Blanco was previously described by Trujillo et al. (1995), Barranco et al. (2000a) and Belaj et al. (2002). Sigoise were previously described by Besnard et al. (2001). Haouzia, Menara were described by (Charafi et al. (2008) and Zine El Aabidine et al. (2010)
Picual	Picual de Hoja Oscura, Picual de Hoja Clara and Olivo Macho de Santisteban Pto	Picual de Hoja Oscura and Picual de Hoja Clara were described by Barranco et al. (2005)
Picudo	Picudo de Fruto Rojo	Barranco et al. (2005)
Varudo	Picudo de Montoro	
Uovo di Piccione	Novo	
Valanolia	Cakir	Barranco et al. (2000a) and Belaj et al. (2002)
Verdial de Badajoz	Corneja and Llorón de Iznalloz	
Villalonga	Branquita de Elvas	Barranco et al. (2000a)
Total (30)	Total (52)	

The original report is cited in the case of synonyms previously described

variability (van Hintum et al. 2000). For this reason, we increased the number of SSRs commonly used for identification purposes (Baldoni et al. 2009; Haouane et al. 2011) not

only to improve this task, but also to provide optimised markers to further characterise the WOGBC genetic variability. To do so, we selected a set of 33 SSRs from among 77

**Table 5** Cases of homonyms found in the identification process at the WOGBC

Cultivar <sup>a</sup>	Homonym	Original report
AbbadiShalal, Bent al Kadi <sup>a</sup> , AbbadiAbou Gabra-842 and AbbadiShalal	Abbadi	Belaj et al. (2003a)
About Choki-1115, About Choki-1126 and AboutKanani	About	
Alameño Blanco, Alameño de Cabra, Picholine Marocaine <sup>a</sup> , Alameño de Montilla and Lechín de Sevilla	Alameño	Barranco et al. (2005)
Cañivano Negro and PicholineMarocaine <sup>a</sup>	Cañivano	Barranco et al. (2005)
Carrasqueño de Elvas <sup>a</sup> , Carrasqueño de Alcaudete, Carrasqueño de Jumilla, Carrasqueño de la Sierra, Carrasqueño de Porcuna, Manzanilla Cacereña, Manzanilla de Sevilla and Picudo	Carrasqueño	Barranco et al. (2005)
Cordovil de Castelo Branco and Cordovil de Serpa	Cordovil	
Callosina, Cornicabra, Cornicabra de Jerez Caballeros and Cornicabra de Mérida	Cornicabra	Barranco et al. (2005)
Escarabajuelo de Atarfe, Escarabajuelo de Posadas and Escarabajuelo de Úbeda	Escarabajuelo	Barranco et al. (2005)
Gordal de Granada, Gordal de Hellín, Gordal de Vélez Rubio, Gordal Sevillana, Mollar de Cieza, Ocal, Picual de Almería and Picudo	Gordal	Barranco et al. (2005)
Grosal de Cieza, Grosal Vimbodi <sup>a</sup> and Picual	Grosal	Barranco et al. (2005)
Habichuelero de Baena and Habichuelero de Grazalema	Habichuelero	Barranco et al. (2005)
Jlot and Jlot-841	Jlot	
Lechín de Granada and Lechín de Sevilla	Lechín	Barranco et al. (2005)
Llorón de Atarfe, Verdial de Badajoz <sup>a</sup> , Morona <sup>a</sup> and Ocal	Llorón	Barranco et al. (2005)
Mahati-846 and Mahati-1010	Mahati	
Majhol-152, Majhol-1059, Majhol-1063 and Majhol-1122	Majhol	
Castellana, Bodoquera, Lechín de Granada, Manzanilla Cacereña, Manzanilla de Abla, Manzanilla de Agua, Manzanilla de Almería, Manzanilla de Hellín, Manzanilla de Montefrío, Manzanilla de Sevilla, Manzanilla del Piquito, Menya <sup>a</sup> , Manzanilla Prieta, Mollar de Cieza, Verdial de Badajoz and Villalonga	Manzanilla	Barranco et al. (2005)
Mission Moojeski and Picholine Marocaine <sup>a</sup>	Mission	Soleri et al. (2010)
Gordal Sevillana, Mollar Basto, Mollar de Cieza and Verdial de Badajoz	Mollar	Barranco et al. (2005)
Morisca and Morisca de Mancor	Morisca	Barranco et al. (2005)
Manzanilla Cacereña, Negrillo de Arjona, Negrillo de Estepa, Negrillo de Iznalloz, Negrillo de la Carlota and Negrillo Redondo	Negrillo	Barranco et al. (2005)
Nevadillo de Santisteban Pto and Picual	Nevadillo	Barranco et al. (2005)
Nevado Azul, Nevado Basto, Nevado Rizado <sup>a</sup> and Picual	Nevado	Barranco et al. (2005)
Pico Limón and Pico Limón de Grazalema	Pico Limón	Barranco et al. (2005)
Picual and Picual de Almería	Picual	Barranco et al. (2005)
Picudo and Varudo <sup>a</sup>	Picudo	Barranco et al. (2005)
Manzanilla del Piquito, Picudo and Picudo Blanco de Estepa	Picudo Blanco	Barranco et al. (2005)
Bodoquera, Manzanilla Cacereña, Picual, Redondilla de Grazalema and Redondilla de Logroño	Redondilla	Barranco et al. (2005)
Changlot Real, Royal de Calatayud, Royal de Cazorla, Royal de Sabiñan <sup>a</sup> and Alfara <sup>a</sup>	Royal	Barranco et al. (2005)
Toffahi and <i>Toffahi-1000</i>	Toffahi	
Gemlik and <i>Trylia-992</i>	Trylia	
Mixani <sup>a</sup> , Ulliri BardheiTiranes and Ulliri Kuq	Ulliri	Belaj et al. (2003b)
Morisca, Verdial de Badajoz, Verdial de Huévar, Verdial de Cádiz and Verdial de Vélez-Málaga	Verdial	Barranco et al. (2005)
	Total (33)	

The original report is cited in the case of homonyms previously described

<sup>a</sup> Cultivars which were not previously defined homonyms

polymorphic markers based on their high variability, reproducible patterns, easy interpretation, and polymorphism.

The overall levels of polymorphism, heterozygosity and PIC obtained for the entire collection were consistent with those observed in previous studies, thus corroborating the high variability and heterozygosity values of olive germplasm (Belaj et al. 2003c; Bracci et al. 2009; Cantini et al. 2008; Erre et al. 2010; Koehmstedt et al. 2010; Sarri et al. 2006). The large number of alleles amplified in the entire collection, many of which appeared at low frequencies, also demonstrated this high variability. The international origin of the cultivars, as well as the occurrence and accumulation of punctual mutations throughout the long history of clonal reproduction of these cultivars might have strongly contributed to this variability (Charafi et al. 2008; Díez et al. 2011; Soleri et al. 2010).

The heterozygote deficiency detected at some SSRs loci, such as GAPI82, GAPI-11e17, *ssrOeUA-DCA10*, *ssrOeUA-DCA15* and UDO99-042, could be related to the presence of null alleles at these markers (Table 2). These markers were still selected for identification purposes because of their polymorphism level, clarity of amplification and ease of interpretation. Nevertheless, these markers should be used with caution in parentage and population structure analyses because of their null allele frequency (Dakin and Avise 2004).

The application of regular identification protocols in large collections is costly and labour intensive. Therefore, one of the goals of this study was propose a gradual use three nested sets of SSRs according to their discriminant capacity for an efficient and progressive identification of the olive accessions in a germplasm bank (Fig. 1). The application of five and ten SSRs was enough to discriminate between 79 and 93 % of the genotypes, respectively and 17 SSRs were able to discriminate between all of them. Thus, when a set of samples coming from the survey of a region is received in the WOGBC, the first task would be to eliminate duplicated samples. This step could be quickly, easily and most importantly, economically done applying from five to ten SSRs, depending on the level of similarity of the samples. Afterwards, the different genotypes should be better characterised applying larger sets of SSRs with the aim of accurately evaluate their genetic diversity.

Most of the SSRs have already been systematically used at the WOGBC to conduct preliminary identification tasks and have been successfully employed in previous studies focused on the characterisation of olive germplasm (Belaj et al. 2007; Bracci et al. 2009; Díez et al. 2011; Erre et al. 2010; Noormohammadi et al. 2007). It is worth of mentioning that our 5-SSR set was included in the set of 11 SSRs proposed by Baldoni et al. (2009) for olive identification purposes. However, only five and seven markers of our 10-SSR and 17-SSR sets, respectively, were represented in the set of Baldoni et al. (2009). Reciprocally, the 11 SSRs proposed by Baldoni et al. (2009) were included in our large set of 33 SSR markers, with the exception of *ssrOeUA-DCA14*. Despite the

congruency of the minimum 5-SSR between both studies, the different selection of markers in the subsequent sets could be due to two main reasons. First, we screened SSRs that were not tested by Baldoni et al. (2009), for example those reported by Gil et al. (2006). Second, the SSRs proposed by Baldoni et al. (2009) were selected using a set of 21 cultivars and tested in a sample of 77 accessions. By contrast, our SSR sets were selected using a group of 48 cultivars and tested on the 499 accessions contained in the WOGBC.

#### Identification of olive cultivars

The SSRs profiles were complemented with the evaluation of the endocarp traits to identify 332 olive cultivars in the WOGBC; similar methods have been employed in other fruit species, such as chestnut, sweet cherry, strawberry or grape (García et al. 2002; Ganopoulos et al. 2011; Martín et al. 2009; Zulini et al. 2005).

The information given by the SSRs markers was taken as main criteria to considered the different genotypes to be different cultivars having into account that: (a) the discrimination capacity of the SSRs markers was the double than the capacity exhibited by the endocarp traits (411 vs. 246 profiles); (b) the phenotypic changes derived from genetic differences might be expressed in other not evaluated organs or features, such as leaves, tree architecture or fatty acid compounds. This last reason remarks the relevance of the phenotypic characterisation of the cultivars in a germplasm bank.

Coupling between molecular and phenotypic differences is a classic controversial topic in the identification of cultivars (Staub and Meglic 1993). Qualitative morphological traits have been successfully used for cataloguing the olive cultivars of Spain (Barranco et al. 2005). Indeed, endocarp traits were successfully used to overcome the initial confusion originated by cases of synonymy and homonymy for two main reasons: first, they were the most discriminating and stables morphological traits (Barranco et al. 2000a, 2005); and second, endocarps could be indefinitely preserved and exchanged among collections. Despite endocarp traits present these positive features; SSR markers have greatly overcome their discrimination capacity discarding environmental variation effects. Nevertheless morphological descriptors are still necessary to complement the UPOV descriptions and they have been very helpful in the identification process of the WOGBC. We described four pairs of cultivars (Chemlali-744-Chetoui; Azulejo-Manzanilla Cacereña; Cordovil CasteloBranco-Verdial de Badajoz and Zarza-Lechin de Sevilla) (Fig. 3b) with the same or highly similar SSR profiles that showed morphological and even agronomical differences (Barranco et al. 2005). We hypothesised that these pairs of cultivars might have originated from the same cultivar through punctual somatic mutations, which might trigger major morphological changes without affecting the amplified SSR regions.

However, this is only a hypothesis, and further research would be necessary to disentangle the responsible genetic mechanisms underlying this variation. Nonetheless, we detected cases such as Lechin de Sevilla and Zarza (Fig. 3b), where the morphological differences were so evident that the effect of environmental variation might be just a secondary force. These cases highlighted the usefulness for complementing molecular markers with morphological descriptors to characterise and identify olive cultivars.

#### Different cultivars and molecular variants

We described 130 genotypes that could be considered as molecular variants of 48 different cultivars because there were no morphological differences in the endocarp and only small genotypic differences (Table 3; Fig. 2). Despite the possibility of genotyping errors, which must always be taken into account, the small genotypic differences could be due to somatic mutations and consider intracultivar variation. Then, having into account this phenomenon and in terms of cultivars, the application of the set of five SSRs previously described, discriminate between 93 % of the cultivars; 10 SSRs were sufficient to distinguish between 100 % of the cultivars, albeit without revealing intracultivar variation; and 17 SSRs were able to discriminate between all cultivars the collection including intracultivar variation. This singularity has been frequently observed in olive using different molecular markers, including RAPD, AFLP and SSR markers (Banilas et al. 2003; Charafi et al. 2008; Cipriani et al. 2002; Díez et al. 2011; Garcia-Diaz et al. 2003; Khadari et al. 2008). Several cases of somatic mutations giving rise to variation within cultivars has often been described in grape, a fruit crop that shows a remarkable resemblance with olive in regard to its diffusion history, propagation system and diversity of cultivars (Riaz et al. 2002; This et al. 2006). The profuse detection of somatic mutations in these crops might be due to two facts: (a) traditional cultivars have been continuously clonally propagated and could have accumulated somatic mutations without accompanying phenotypic consequences in crop morphology and agronomic performance (Díez et al. 2011; Riaz et al. 2002; and (b) mutations are more likely to occur in highly variable and neutrally evolving genomic regions such as SSRs. Additionally, highly polymorphic di-nucleotide SSRs are the most widely used for identification purposes in olive and other clonal fruit crops (Baldoni et al. 2009; Haouane et al. 2011; Irish et al. 2010; Laucou et al. 2011; Motilal et al. 2011; Zhang et al. 2009). For these reasons, new SSRs with core repeats from three to six nucleotides long have recently been developed in grape (Cipriani et al. 2008) as well as in olive (González-Plaza et al. 2011). The polymorphism, stability and transferability of these markers are currently being assessed in accessions from the WOGBC.

#### Authentication process, synonyms and homonyms

The authentication process guarantees that a cultivar, which is being distributed worldwide, corresponds to the original cultivar growing in its putative area of origin. Therefore, authentication should be an essential requisite before a germplasm bank is able to distribute plant material to other collections, researchers and nursery plant certification agencies.

In this study, 200 accessions were authenticated, 172 based only on their identity with control samples of the endocarp from their original growing area and 28 by the comparison of both endocarp and DNA control samples. DNA profiles are a valuable complement to the authentication of cultivars with subtle morphological differences or when the technicians are not well trained in the morphological characterisation. The authentication of the entire collection was not possible due to the lack of control samples in the WOGBC (Table 1; Online resource 2). Several factors hinder the possibility of developing a complete collection of authentic control samples of all endocarps such as: (a) the absence of characterisation studies of olive cultivars in many olive growing countries; (b) the fact that several old local cultivars are disappearing from their autochthonous areas of origin and c) the incomplete or erroneous passport data associated with many accessions received in germplasm banks that make their comparison with control samples impossible (Trujillo et al. 2006).

Authentic control samples were helpful resources to compare and describe cases of synonyms and homonyms. We found 43 new cases of synonyms among the cultivars (Table 4). Olive growing has been traditionally linked to human migrations, which may have blurred the fingerprints of independent domestication events and led to complex relationships among cultivars (Baldoni et al. 2006; Bracci et al. 2009; Díez et al. 2011; Koehmstedt et al. 2010; Sarri et al. 2006; Soleri et al. 2010). Lately, the globalisation process has reinforced this trend by increasing the movement of olive cultivars within and among countries. The synonymy case revealed by this study involving the accessions named Alameño de Marchena (COR000254; Spain), Cañivano Blanco (COR000052; Spain), Haouzia (COR000835; Morocco), Menara (COR000836; Morocco), Mission Nieland (COR000716; USA) and Sigoise (COR000119; Algeria), which all corresponded to the main Moroccan cultivar cv. Picholine Marocaine, is a clear example of this phenomenon. This case is especially remarkable since the cv. Mission was thought to have originated from California (USA) and to be part of the olive production identity of this area (Soleri et al. 2010). According to our results, cv. Mission Nieland and cv. Picholine Marocaine are synonymous; this result is in accord with those obtained in the USA by Koehmstedt et al. (2010) between Zitoun, a well-known synonymous name of Picholine Marocaine (Barranco et al. 2005; Khadari et al. 2008), and Mission Nieland.

The correct identity of each tree is essential to avoid the propagation of mislabelled cultivars that could have harmful consequences for the nursery industry, research and breeding programs. Despite the actions taken to avoid errors, we found that trees of 36 (7.2 %) accessions corresponding to 34 different genotypes (8.2 %), did not share the same SSR profile. Similarly, errors due to mislabeling have been reported for French olive cultivars (Khadari et al. 2003), apple (Evans et al. 2011), *Cicer* (Shan et al. 2005), persimmon (Badenes et al. 2003) and cacao (Motilal and Butler 2003).

#### WOGBs and international networks

International initiatives are currently being established to identify and authenticate the accessions included in the WOGBs of Cordoba and Marrakech using a common protocol and for gathering a complete worldwide collection of authentic control samples.

The molecular characterisation efforts at the WOGB of Cordoba and Marrakech used 11 SSR markers in common (Haouane et al. 2011). Nine of these markers were slightly more polymorphic and had greater allelic richness in the Marrakech collection than in Cordoba; however, these values were not significantly different (Kruskal–Wallis test;  $P > 0.05$ ; Online resource 6). This similarity in genetic diversity is remarkable given the distinctive composition and geographical areas represented in each collection. The collections only share the names of 153 accessions (Haouane et al. 2011). The WOGBC includes a large proportion of samples from the northern shore of the Mediterranean Basin along with a large collection of Spanish cultivars, but North of Africa is poorly represented. By contrast, the Marrakech collection, with 505 olive genotypes from 24 countries, has a more balanced composition of the Mediterranean olive growing areas, with countries such as Morocco, Algeria and Egypt well represented (Haouane et al. 2011). The distinctive and rich composition of both WOGBs highlights the need for cooperative international projects to characterise the collections and to guarantee the efficient conservation of their genetic resources.

Despite the international efforts, we will never be able to authenticate the cultivars that have already disappeared from the areas where they were once sampled. This real case of genetic erosion should be taken into account when promoting the characterisation and conservation of olive cultivars in those countries where this task has not even begun. The combined application of a minimum of morphological traits and SSRs can be an efficient and affordable tool for the identification of olive cultivars in those countries where germplasm characterisation is still in the early stages. Recently, high-throughput markers have been applied to the characterisation of olive cultivars (Belaj et al. 2012) and other fruit crops, such as grape (Myles et al. 2011). These methods

provide powerful tools for both deep genetic characterisation and association mapping to be applied in the near future.

#### Conclusions

The identification of the accessions of any olive germplasm bank should be compulsory before distribution of any plant material from that bank. Only diffusion of true to type cultivars will avoid the worldwide confusion between denominations and cultivars existing in almost any world germplasm collection (Bartolini et al. 2005). This work illustrates how the use of 17 SSR and 11 endocarp traits has allowed for the identification of different sources of error (propagation, mislabeling, inconsistent denominations). In addition, the method led to the establishment of synonymies, homonymies, intracultivar molecular variability and partial authentication of cultivars housed in one of the world's largest olive germplasm banks. Our protocol could be useful for managing olive germplasm banks and identifying the true to type cultivars to be preserved and exchanged.

**Acknowledgements** This work was financially supported by the projects RF01-006 and RF-2009-00011-00-00 (INIA, the Ministerio de Ciencia e Innovación of Spain) and Oleagen (Fundación Genoma España, Junta de Andalucía and Corporación Tecnológica de Andalucía). The research carried out by I. Trujillo was partially performed during a sabbatical leave at the University of California, Davis, which was funded by the Ministerio de Educación of Spain. Programa Salvador de Madariaga, No. Ref. PR-2009-0527. The authors are grateful to Dr. Juan M. Caballero and to the late Dr. Carmen Del Río, successive curators of the WOGBC, for facilitating us to carry out this work.

#### References

- Badenes M, Garcés A, Romero C, Romero M, Clave J, Rovira M, Llacer G (2003) Genetic diversity of introduced and local Spanish persimmon cultivars revealed by RAPD markers. *Genet Resour Crop Evol* 50:579–585
- Baldoni L, Tosti N, Ricciolini C, Belaj A, Arcioni S, Pannelli G, Germana MA, Mulas M, Porceddu A (2006) Genetic structure of wild and cultivated olives in the central Mediterranean basin. *Ann Bot* 98: 935–942
- Baldoni L, Cultrera NG, Mariotti R, Ricciolini C, Arcioni S, Vendramin GG, Buonamici A, Porceddu A, Sarri V, Ojeda MA, Trujillo I, Rallo L, Belaj A, Perri E, Salimonti A, Muzzalupo I, Casagrande A, Lain O, Messina R, Testolin R (2009) A consensus list of microsatellite markers for olive genotyping. *Mol Breed* 24:213–231
- Banilas G, Minas J, Gregoriou C, Demoliou C, Kourti A, Hatzopoulos P (2003) Genetic diversity among accessions of an ancient olive variety of Cyprus. *Genome* 46:370–376
- Barranco D, Rallo L (1984) Las variedades de olivo cultivadas en Andalucía. Ministerio de Agricultura, Junta de Andalucía, Madrid, Spain
- Barranco D, Cimato A, Fiorino P, Rallo L, Touzani A, Castañeda C, Serafini F, Trujillo I (2000a) World olive catalogue of olive varieties. International Olive Oil Council, Madrid, Spain

- Barranco D, Trujillo I, Rallo L (2000b) Are ‘Oblonga’ and ‘Frantoio’ the same cultivar? *HortSci* 35:1323–1325
- Barranco D, Trujillo I, Rallo L (2005) *Elaiografia Hispanica*. In: Rallo L, Barranco D, Caballero JM, Del Rio C, Martin A, Tous J, Trujillo I (eds) *Varietades de olivo en España*. Mundi-Prensa, Madrid
- Bartolini G, Petrucelli R (2002) Classification, origin, diffusion and history of the olive. FAO, Rome
- Bartolini G, Prevost G, Messeri C, Carignani C (2005) Olive germplasm: cultivars and world-wide collections. In: FAO SaPGRSo (ed) FAO
- Belaj A, Satovic Z, Rallo L, Trujillo I (2002) Genetic diversity and relationships in olive (*Olea europaea* L.) germplasm collections as determined by randomly amplified polymorphic DNA. *Theor Appl Genet* 105:638–644
- Belaj A, Caballero JM, Barranco D, Rallo L, Trujillo I (2003a) Genetic characterization and identification of new accessions from Syria in an olive Germplasm Bank by means of RAPD markers. *Euphytica* 134:261–268
- Belaj A, Satovic Z, Ismaili H, Panajoti D, Rallo L, Trujillo I (2003b) RAPD genetic diversity of albanian olive germplasm and its relationships with other mediterranean countries. *Euphytica* 130:387–395
- Belaj A, Satovic Z, Cipriani G, Baldoni L, Testolin R, Rallo L, Trujillo I (2003c) Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive. *Theor Appl Genet* 107:736–744
- Belaj A, Munoz-Diez C, Baldoni L, Porceddu A, Barranco D, Satovic Z (2007) Genetic diversity and population structure of wild olives from the north-western Mediterranean assessed by SSR markers. *Ann Bot* 100:449–458
- Belaj A, Dominguez-García MC, Atienza SG, Martín Urdíroz N, De la Rosa R, Satovic Z, Martín A, Kilian A, Trujillo I, Valpuesta V, Del Río C (2012) Developing a core collection of olive (*Olea europaea* L.) based on molecular markers (DARs, SSRs, SNPs) and agronomic traits. *Tree Genet Genomes* 8:365–378
- Besnard G, Breton C, Baradat P, Khadari B, Bervillé A (2001) Cultivar identification in the olive (*Olea europaea* L.) based on RAPDs. *J Am Soc Hortic Sci* 126:668–675
- Besnard G, Khadari B, Navascués M, Fernández-Mazuecos M, El Bakkali A, Arrigo N, Baali-Cherif D, Brunini-Bronzini de Caraffa V, Santoni S, Vargas P, Savolainen V (2013) The complex history of the olive tree: from Late Quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant. *Proc R Soc B Biol Sci* 280(1756)
- Botstein D, White R, Skolnick M, Davis R (1980) Construction of a genetic-linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314–331
- Bowers JE, Dangl GS, Vignani R, Meredith CP (1996) Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). *Genome* 39(4):628–633
- Bracci T, Sebastiani L, Busconi M, Fogher C, Belaj A, Trujillo I (2009) SSR markers reveal the uniqueness of olive cultivars from the Italian region of Liguria. *Sci Hortic* 122:209–215
- Caballero JM, del Rio C, Barranco D, Trujillo I (2006) The Olive World Germplasm Bank of Cordoba, Spain. *Olea* 25:14–19
- Cantini C, Cimato A, Autino A, Redi A, Cresti M (2008) Assessment of the Tuscan olive germplasm by microsatellite markers reveals genetic identities and different discrimination capacity among and within cultivars. *J Am Soc Hortic Sci* 133:598–604
- Carriero F, Fontanazza G, Cellini F, Giorio G (2002) Identification of simple sequence repeats (SSRs) in olive (*Olea europaea* L.). *Theor Appl Genet* 104:301–307
- Charafi J, El Meziane A, Moukhli A, Boulouha B, El Modafar C, Khadari B (2008) Menara gardens: a Moroccan olive germplasm collection identified by a SSR locus-based genetic study. *Genet Resour Crop Evol* 55:893–900
- Cipriani G, Marrazzo MT, Marconi R, Cimato A, Testolin R (2002) Microsatellite markers isolated in olive (*Olea europaea* L.) are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars. *Theor Appl Genet* 104:223–228
- Cipriani G, Marrazzo MT, Di Gaspero G, Pfeiffer A, Morgante M, Testolin R (2008) A set of microsatellite markers with long core repeat optimized for grape (*Vitis* spp.) genotyping—art. no. 127. *BMC Plant Biol* 8:127–127
- Corrado G, La Mura M, Ambrosino O, Pugliano G, Varricchio P, Rao R (2009) Relationships of Campanian olive cultivars: comparative analysis of molecular and phenotypic data. *Genome* 52:692–700
- Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* 93(5):504–509
- de la Rosa R, James C, Tobutt KR (2002) Isolation and characterization of polymorphic microsatellite in olive *Olea europaea* L. and their transferability to other genera in the Oleaceae. *Mol Ecol Notes* 2: 265–267
- Dice LR (1945) Measures of the amount of ecologic association between species. *Ecology* 26:297–302
- Diez CM, Trujillo I, Barrio E, Belaj A, Barranco D, Rallo L (2011) Centennial olive trees as a reservoir of genetic diversity. *Ann Bot* 108:797–807
- Diez CM, Imperato A, Rallo L, Barranco D, Trujillo I (2012) Worldwide core collection of olive cultivars based on simple sequence repeat and morphological markers. *Crop Sci* 52:211–221
- D’Imperio M, Viscosi V, Scarano MT, D’Andrea M, Zulo BA, Pilla F (2001) Integration between molecular and morphological markers for the exploitation of olivegermplasm (*Olea europaea*). *Sci Hortic* 130:229–240
- Downey G, Boussion J (1996) Authentication of coffee bean variety by near-infrared reflectance spectroscopy of dried extract. *J Sci Food Agric* 71(1):41–49
- Erre P, Chessa I, Munoz-Diez C, Belaj A, Rallo L, Trujillo I (2010) Genetic diversity and relationships between wild and cultivated olives (*Olea europaea* L.) in Sardinia as assessed by SSR markers. *Genet Resour Crop Evol* 57:41–54
- Evans KM, Patocchi A, Rezzonico F, Mathis F, Durel CE, Fernández-Fernández F, Boudichevskaia A, Dunemann F, Stankiewicz-Kosyl M, Gianfranceschi L, Komjanc M, Latuer M, Madduri M, Noordijk Y, Van de Weg WE (2011) Genotyping of pedigreed apple breeding material with a genome-covering set of SSRs: trueness-to-type of cultivars and their parentages. *Mol Breed* 28:535–547
- FAO (2008) The Statistical Database (FAOSTAT). Food and Agriculture Organization of the United Nations (FAO), Rome
- Fendri M, Trujillo I, Trigui A, Rodriguez-García MI, Ramirez JDA (2010) Simple sequence repeat identification and endocarp characterization of olive tree accessions in a Tunisian germplasm collection. *Hortscience* 45:1429–1436
- Ganopoulos IV, Kazantzis K, Chatzicharis I, Karayiannis I, Tsiftaris AS (2011) Genetic diversity, structure and fruit trait associations in Greek sweet cherry cultivars using microsatellite based (SSR/ISSR) and morpho-physiological markers. *Euphytica* 181:237–251
- García MG, Ontivero M, Diaz Ricci JC, Castagnaro A (2002) Morphological traits and high resolution RAPD markers for the identification of the main strawberry varieties cultivated in Argentina. *Plant Breed* 121:76–80
- García-Díaz A, Oya R, Sanchez A, Luque F (2003) Effect of prolonged vegetative reproduction of olive tree cultivars (*Olea europaea* L.) in mitochondrial homoplasmy and heteroplasmy. *Genome* 46:377–381
- Gil FS, Busconi M, Machado AD, Fogher C (2006) Development and characterization of microsatellite loci from *Olea europaea*. *Mol Ecol Notes* 6:1275–1277
- González-Plaza JJ, Muñoz-Mérida A, Ortiz-Martín I, Domínguez-García MC, Martín-Urdíroz N, Diez-Muñoz C, Sánchez-Sevilla JL, Rallo L, Trujillo I, Belaj A, de la Rosa R, Trelles O, Valpuesta V, Beuzón CR (2011) Identification of molecular markers in olive tree for the analysis of agronomical traits. In: 4th International Conference for Olive Tree and Olive Products. OliveBioteq, Chania, Grecia

- Haouane H, El Bakkali A, Moukhli A, Tollon C, Santoni S, Oukabli A, El Modafar C, Khadari B (2011) Genetic structure and core collection of the World Olive Germplasm Bank of Marrakech: towards the optimised management and use of Mediterranean olive genetic resources. *Genetica*, Amsterdam
- International Union for the Protection of New Varieties of Plants. 1991. International convention for the protection of new varieties of plants of 2 December 1961, as revised at Geneva on 10 November 1972, on 23 October 1978 and on 19 March 1991. <http://www.upov.int/en/publications>. Accessed 6 February 2011
- Irish BM, Goenaga R, Zhang DP, Schnell R, Brown JS, Motamayor JC (2010) Microsatellite fingerprinting of the USDA-ARS tropical agriculture research station cacao (*Theobroma cacao* L.) germplasm collection. *Crop Sci* 50:656–667
- Khadari B, Breton C, Moutier N, Roger JP, Besnard G, Berville A, Dosba F (2003) The use of molecular markers for germplasm management in a French olive collection. *Theor Appl Genet* 106:521–529
- Khadari B, Charafi J, Moukhli A, Ater M (2008) Substantial genetic diversity in cultivated Moroccan olive despite a single major cultivar: a paradoxical situation evidenced by the use of SSR loci. *Tree Genet Genomes* 4:213–221
- Kimura T, Zhong Shi Y, Shoda M, Kotobuki K, Matsuta N, Hayashi T, Ban Y, Yamamoto T (2002) Identification of Asian pear varieties by SSR analysis. *Breed Sci* 52:115–121
- Koehmstedt AM, Aradhya MK, Soleri D, Smith JL, Polito VS (2010) Molecular characterization of genetic diversity, structure, and differentiation in the olive (*Olea europaea* L.) germplasm collection of the United States Department of Agriculture. *Genet Resour Crop Evol* 58:519–531
- Laucou V, Lacombe T, Dechesne F, Siret R, Bruno JP, Dessup M, Dessup T, Ortigosa P, Parra P, Roux C, Santoni S, Vares D, Peros JP, Boursiquot JM, This P (2011) High throughput analysis of grape genetic diversity as a tool for germplasm collection management. *Theor Appl Genet* 122:1233–1245
- Liu KJ, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21:2128–2129
- Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* 7:639–655
- Martín MA, Alvarez JB, Mattioni C, Cherubini M, Villani F, Martín LM (2009) Identification and characterisation of traditional chestnut varieties of southern Spain using morphological and simple sequence repeat (SSRs) markers. *Ann Appl Biol* 154:389–398
- Melchiade D, Foroni I, Corrado G, Santangelo I, Rao R (2007) Authentication of the ‘Annurca’ apple in agro-food chain by amplification of microsatellite loci. *Food Biotechnol* 21(1):33–43
- Motilal L, Butler D (2003) Verification of identities in global cacao germplasm collections. *Genet Resour Crop Evol* 50:799–807
- Motilal LA, Zhang DP, Umaharan P, Mischke S, Pinney S, Meinhardt LW (2011) Microsatellite fingerprinting in the International Cocoa Genebank, Trinidad: accession and plot homogeneity information for germplasm management. *Plant Genet Resour-C* 9:430–438
- Mouly PP, Gaydou EM, Faure R, Estienne JM (1997) Blood orange juice authentication using cinnamic acid derivatives. Variety differentiations associated with flavanone glycoside content. *J Agric Food Chem* 45(2):373–377
- Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK, Prins B, Reynolds A, Chia JM, Ware D, Bustamante CD, Buckler ES (2011) Genetic structure and domestication history of the grape. *Proc Natl Acad Sci U S A* 108:3530–3535
- Noomohammadi Z, Hosseini-Mazinani M, Trujillo I, Ratio L, Belaj A, Sadeghzadeh M (2007) Identification and classification of main Iranian olive cultivars using Microsatellite markers. *Hortscience* 42:1545–1550
- Park SDE (2001) Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection. University of Dublin
- Rallo L (2005) Variedades de olivo en España: una aproximación cronológica. In: Rallo L, Caballero JM, Del Río C, Martín A, Tous J, Trujillo I (eds) Variedades de olivo en España. Junta de Andalucía, MAPA and Ediciones Mundi-Prensa, Madrid
- Rallo L, Barranco D, Castro-García S, Connor DJ, Gómez-del-Campo M, Rallo P (2013) High-density olive plantations. *Hortic Rev* 41:303–384
- Riaz S, Garrison KE, Dangl GS (2002) Genetic divergence and chimerism within ancient asexually propagated winegrape cultivars. *J Amer Hort Sci* 127(4):508–514
- Rohlf FJ (1998) NTSYS-pc. Numerical taxonomy and multivariate analysis system. Version 2.00. Exeter Software. Setauket, New York
- Sarri V, Baldoni L, Porceddu A, Cultrera NGM, Contento A, Frediani M, Belaj A, Trujillo I, Cionini PG (2006) Microsatellite markers are powerful tools for discriminating among olive cultivars and assigning them to geographically defined populations. *Genome* 49:1606–1615
- Sefc KM, Lopes S, Mendonca D, Dos Santos MR, Machado MLD, Machado AD (2000) Identification of microsatellite loci in olive (*Olea europaea*) and their characterization in Italian and Iberian olive trees. *Mol Ecol* 9:1171–1173
- Shan F, Clarke HC, Plummer JA, Yan G, Siddique KHM (2005) Geographical patterns of genetic variation in the world collections of wild annual Cicer characterized by amplified fragment length polymorphisms. *Theor Appl Genet* 110:381–391
- Soleri D, Koehmstedt A, Aradhya MK, Polito V, Pinney K (2010) Comparing the historic olive trees (*Olea europaea* L.) of Santa Cruz Island with contemporaneous trees in the Santa Barbara, CA area: a case study of diversity and structure in an introduced agricultural species conserved in situ. *Genet Resour Crop Evol* 57:973–984
- Staub JE, Meglic V (1993) Molecular genetic markers and their legal relevance for cultivar discrimination: a case study in cucumber. *HortTechnology* 3(3)
- This P, Lacombe T, Thomas MR (2006) Historical origins and genetic diversity of wine grapes, vol. 22 No. 9. Elsevier, Amsterdam
- Trujillo I, Rallo L, Arus P (1995) Identifying olive cultivars by isozyme analysis. *J Am Soc Hortic Sci* 120:318–324
- Trujillo I, Ojeda MA, Baldoni L, Belaj A (2006) Olive cultivar identification by means of microsatellites (SSR). *Olea* 25:24–27
- van Hintum TJJ, Brown AHD, Spillane C, Hodgkin T (2000) Core collections of plant genetic resources. IPGRI Technical bulletin, p 48
- Wünsch A, Hormaza JI (2002) Molecular characterisation of sweet cherry (*Prunus avium* L.) genotypes using peach [*Prunus persica* (L.) Batsch] SSR sequences. *Heredity* 89:56–63
- Zhang DP, Mischke S, Johnson ES, Phillips-Mora W, Meinhardt L (2009) Molecular characterization of an international cacao collection using microsatellite markers. *Tree Genet Genomes* 5:1–10
- Zine El Aabidine M, Charafi J, Grout C, Doligez A, Santoni S, Moukhli A, Jay-Allemand C, El Modafar C (2010) Construction of a genetic linkage map for the olive based on AFLP and SSR markers. *Crop Sci* 50(6):2291–2302
- Zohary D, Spiegel-Roy P (1975) Beginnings of fruit growing in the old world. *Science* 187:319–327
- Zulini L, Fabro E, Peterlunger E (2005) Characterisation of the grapevine cultivar Picolit by means of morphological descriptors and molecular markers. *Vitis* 44(1):35–38