TÜBİTAK

Turkish Journal of Agriculture and Forestry

http://journals.tubitak.gov.tr/agriculture/

Research Article

Turk J Agric For
(2016) 40: 583-596
© TÜBİTAK
doi:10.3906/tar-1602-27

Molecular characterization of genetic variability and structure of olive (*Olea europaea* L.) germplasm collection analyzed by agromorphological traits and microsatellite markers

Karim SORKHEH[1,*], **Esmaeil KHALEGHI**[2]
[1]Department of Agronomy and Plant Breeding, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran
[2]Department of Horticulture, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran

Received: 06.02.2016 • **Accepted/Published Online:** 28.03.2016 • **Final Version:** 14.06.2016

Abstract: We studied 200 trees belonging to 20 accessions of cultivated olive (*O. europaea* L.) from 4 regions of origin, evaluated by means of agromorphological traits and simple sequence repeat (SSR) markers. The agromorphological traits showed high variation between genotypes and significant correlation coefficients were obtained among the values recorded in two consecutive years, 2013 and 2014. The maximum coefficient of variation for the quantitative agronomic traits was observed in fruit weight wet (13.45%), while the lowest was found in stone width (3.18%). Fruit shape index, leaf length, leaf width, and lenticel size also showed variability. With both DNA-based and agromorphological descriptors, higher levels of variability were found. Genetic variation observed among the olive germplasm at the DNA level was higher than that of the agromorphological traits, indicating the efficiency of SSR markers for detecting genetic diversity among olive genotypes and their relationships. The lack of consistency between the relationship studies performed with molecular and morphological markers could indicate that each marker system measures different aspects of olive genetic variability. Molecular data obtained by SSR markers together with morphological and agronomical characterization of olive trees confirmed the high diversity and their potential use for olive breeding.

Key words: Genetic variability, microsatellite markers, *Olea europaea* L., STRUCTURE

1. Introduction

Olive (*Olea europaea* subsp. *europaea* var. *europaea*) is one of the oldest agricultural tree crops in the Mediterranean basin with remarkable cultural and economic importance. The richness of the cultivated olive germplasm is an unusual case among horticultural crops, as a consequence of tree longevity and lack of turnover with new breeding genotypes (Barranco et al., 2005; Bartolini et al., 2005; Baldoni and Belaj, 2009). In spite of the richness of cultivated germplasm, olive cultivars exhibit lower genetic diversity than their wild relatives (Lumaret et al., 2004; Breton et al., 2006; Belaj et al., 2010), indicating that the latter could enrich the genetic basis of cultivated material.

To date, most work has concentrated on evaluating the distribution of variability between cultivated and wild olives (Baldoni et al., 2006; Breton et al., 2006; Belaj et al., 2007; Erre et al., 2010) and on establishing the genetic relationships among the different *O. europaea* subspecies that are distributed beyond the Mediterranean area (Besnard et al., 2007; García-Verdugo et al., 2010). Due to recent advances in DNA technologies most of these studies have been performed by means of molecular markers, simple sequence repeats (SSRs) being the most widely used. However, in spite of the drawbacks of traditional morphological description, such as environmental influences, and the need for extensive observations of mature plants, the joint use of both morphoagronomic traits and SSR markers could give the opportunity to exploit the complementary natures of these two methods (Karp et al., 1997) in evaluating the genetic diversity of wild olive trees.

It is largely accepted that olive cultivar discrimination based on morphological descriptions is not completely reliable (Belaj et al., 2002, 2007, 2010, 2011); therefore, DNA molecular markers, and particularly microsatellites (SSRs), are today widely used (Bracci et al., 2011; Noormohammadi et al., 2014) to complement morphological analyses and to unambiguously identify the accessions held in collections. Genetic variation has been reported among naturally occurring olive clones in the literature with molecular markers. Clones were identified with RAPD and ISSR (Gemas et al., 2004; Gomes et al., 2008; Martins-Lopes

* Correspondence: karimsorkheh@gmail.com

et al., 2009), with AFLP (Strikic et al., 2010), and with microsatellites (Lopes et al., 2004; Muzzalupo et al., 2010; Albertini et al., 2011; Zaher et al., 2011; Ipek et al., 2012; Marra et al., 2013; Caruso et al., 2014; Noormohammadi et al., 2014; Abdessemed et al., 2015). Although currently there is intense research to develop reliable techniques for detecting mutations in genes, clone identification is still predominantly based on the study of phenotypic traits, integrated with molecular analyses.

The present work reports the employment of morphoagronomic traits and SSR markers to investigate genetic diversity and relationships among 20 olive cultivars from different regions with high olive oil production in the world. Comparisons among these approaches were also made to assess their efficiency in evaluating genetic diversity levels and relationships among the olive trees under study. To the best of the authors' knowledge this is the first morphoagronomic and microsatellite-based study on collecting olive cultivars.

2. Materials and methods

2.1. Plant material

We studied 200 trees belonging to 20 accessions of cultivated olive (*O. europaea* L.) from 4 regions of origin: Abosatl (Syria), T5 (unknown), Dezfoli (Iran), Mishen (USA), Masabei (Syria), Konservolia (Europe), Kaeisei (Europe), Kaylit (Europe), T21 (Europe), Khoseari (Syria), Zard (Iran), Roghani (Iran), T7 (Europe), Manzanila (Europe), Kavi (Syria), T2 (Europe), Balidi (Syria), Mari (Syria), Foji (Europe), and Koroneiki (Europe), with 12 individuals representing each region. All samples were collected from young orchards composed of olive trees 5–10 years old (Table 1).

2.2. Morphoagronomic characterization

Field expeditions were carried out in autumn 2013 and 2014 in Ahvaz (south of Iran). Morphological characterization was based on olive descriptors developed by the International Olive Council (Mulas, 1999; Barranco

Table 1. Cultivar name, geographic origin, and use of fruits.

No.	Cultivar	Origin	Kind of use
1	Zard	North Iran	Double use
2	Roghani	North Iran	Oil
3	Dezfoli	North Iran	Double use
4	T5	Unknown	Table olive
5	Balidi	Syria	Table olive
6	Kavi	Syria	Table olive
7	T2	Europe	Table olive
8	Abosatl	Syria	Double use
9	Mari	Syria	Double use
10	Khoseari	Syria	Table olive
11	Masabei	Syria	Double use
12	Mishen	USA	Table olive
13	Kaylit	Europe	Table olive
14	T21	Europe	Table olive
15	Conservalia	Europe	Table olive
16	Foji	Europe	Double use
17	T7	Europe	Double use
18	Manzanila	Europe	Oil
19	Kaeisei	Europe	Table olive
20	Kroniki	Europe	Table olive

et al., 2000; García-Donas Díaz, 2001; Trujillo et al., 2014). The morphological characters evaluated included 28 qualitative and quantitative traits (Table 2). For that purpose, a random sample of 30 fully expanded leaves and 30 fruits were collected from each of the selected trees in both sampling periods.

2.2.1. Qualitative traits

For determination of the diversity and discriminative power of qualitative morphological traits according to Belaj et al. (2011), we used Shannon's information index as a measure of morphological trait diversity across olive trees, calculated for each trait as:

$$H_i = \sum_{k=1}^{k} \log_2 P_k$$

where *pk* is the frequency of the *k*th state and *K* is a total number of states of the *j*th trait (Lewontin, 1972). The confusion probability (C) of a given morphological trait provides an estimate of the probability that two randomly chosen individuals from the sample of 20 genotypes share identical states for a given morphological trait. This parameter is calculated as described by Tessier et al. (1999) for molecular marker data:

$$C_j = \sum_{k=1}^{k} P_k \frac{NP_k - 1}{N - 1}$$

where Pk is the frequency of the *k*th state, N is the sample size, and *K* is the total number of states of the *j*th morphological trait. The discriminating power (D_j) of the *j*th morphological trait is equal to $D_j = 1 - C_j$, denoting the proportion of distinguishable pairs of individuals. Joint discriminating power (i.e. the proportion of distinguishable pairs in the sample of 20 genotypes) of an increasing set of qualitative traits was calculated by adding the traits ordered according to their discriminating power.

In order to explore the associations among 18 qualitative morphological traits used in the analysis of 20 olive cultivars, the correspondence analysis of categorical data (also known as HOMALS, for homogeneity analysis by means of alternating least squares; Sutherland et al.,

Table 2. Trait state, number of observed states (Kj), Shannon's information index (Hj), and discriminating power (Dj) of 18 qualitative morphological traits used in the analysis of 20 olive trees.

No.	Trait	Abbr.	Traits states				Kj	Hj	Dj
			1	2	3	4			
1	Leaf blade length	LBL	Short	Medium	Long		2	0.564	0.254
2	Leaf blade width	LBW	Narrow	Medium	Wide		3	1.235	0.325
3	Leaf shape (length/width)	LSH	Elliptic	Elliptic-lanceolate	Lanceolate		2	0.948	0.354
4	Fruit symmetry (position A)	FS	Symmetrical	Slightly asymmetrical	Asymmetrical		3	1.325	0.687
5	Fruit apex shape (position A)	FASH	Pointed	Rounded			2	0.568	0.403
6	Fruit base shape (position A)	FBSH	Rounded	Truncate			1	0.589	0.258
7	Fruit position of maximum diameter (B position)	FMD	Toward base	Central	Toward apex		3	1.024	0.847
8	Fruit shape	FSH	Spherical	Oval	Longer		3	1.048	0.654
9	Fruit weight	FW	Low (<0.5 g)	Medium (0.5–1 g)	High (>1 g)		2	0.798	0.568
10	Stone symmetry (position A)	SSA	Symmetrical	Slightly asymmetrical			1	-	-
11	Stone symmetry (position B)	SSB	Symmetrical	Slightly asymmetrical			2	0.882	0.658
12	Stone apex shape (A)	SASH	Pointed	Rounded			3	1.056	0.784
13	Stone base shape (position A)	SBSH	Rounded	Truncate	Pointed		3	0.687	0.332
14	Stone position of maximum diameter (B position)	SMD	Toward base	Central	Toward apex		2	0.333	0.235
15	Stone distribution of grooves	SDG	Regular	Irregular			2	1.632	0.248
16	Stone shape (position A)	SSH	Spherical	Oval	Elliptic	Longer	4	1.326	0.658
17	Stone weight	SW	Low (<0.15 g)	Medium (0.15–0.3 g)	High (>0.3 g)		3	1.348	0.458
18	Stone surface	SS	Smooth	Rough	Knotty		1	1.318	0.625
	Average							0.981	0.491
	Min							0.333	0.248
	Max							1.632	0.84

2000) was applied using PROC CORRESP in SAS (SAS Institute, 2004).

2.2.2. Quantitative traits

PROC GLM in SAS was used for analysis of variance for each trait. Variance components were estimated for each agronomic trait with PROC VARCOMP using the REML method. Pearson's correlation coefficients were computed for each trait between years as well as between all traits by PROC CORR in SAS in order to study the associations among traits.

2.3. DNA extraction and microsatellite analysis

Total DNA was extracted from mature leaves collected from the upper parts of trees following the protocol described by De la Rosa et al. (2002) with the minor modifications of Sorkheh et al. (2007). Eleven microsatellite loci (Sefc et al., 2000; Cipriani et al., 2002; De la Rosa et al., 2002) that were successfully used in previous studies (Belaj et al., 2007, 2010) were genotyped in the olive samples collected (Table 3). PCR amplifications were carried out according to Omrani-Sabbaghi et al. (2007) in PCR reactions in a total volume of 25 mL, containing 20 ng of genomic DNA, 1X PCR buffer (10 mM Tris–HCl (pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl), 200 mM dNTPs, 0.5 U of Taq DNA polymerase (Roche), and 0.2 mM of each primer. The amplification profile was composed of an initial denaturation at 94 °C for 5 min, followed by 10 cycles of touchdown PCR (95 °C/45 s, 68.5–58.5 °C/45 s, 72 °C/1 min) and 25 cycles of 95 °C/45 s, 58.5 °C/45 s, and 72 °C/1 min, with a final extension step of 7 min at

Table 3. Overall mean, standard deviation, range, and coefficient of variation (CV) of 10 quantitative agronomic traits in 20 olive trees measured in 2 years (2013 and 2014).

Traits	Abbr.	Min	Max	Mean	Variance
Fruit shape index	FSHI	1.18	1.93	1.45	0.041
Fruit length	FL	15.09	26.91	20.9	10.63
Fruit width	FW	8.99	18.41	14.60	5.97
Weight of wet	WW	0.74	5.73	3.07	1.45
Fruit shape	FSH	1.00	3.68	2.29	0.53
Fruit symmetry (position A)	FSA	1.00	2.68	1.47	0.22
Fruit position of maximum diameter (B position)	FMD	1.4	3.00	2.21	0.15
Fruit apex shape (position A)	FASH	1.00	2.00	1.44	0.19
Fruit base shape (position A)	FBSH	1.00	3.00	1.72	0.47
Location appearance of fruit color change	LAFCC	1.00	3.00	2.13	0.36
Fruit lenticels	FL	1.00	2.00	1.65	0.12
Lenticel size	LS	1.00	2.00	1.3	0.17
Stone shape index	SSHI	1.49	2.87	2.03	0.16
Stone length	SL	12.72	22.93	16.76	8.17
Stone width	SW	5.95	9.44	8.33	0.89
Stone shape (position A)	SSHA	1.67	4.00	2.93	0.52
Stone symmetry (position A)	SSA	2.08	3.00	2.46	0.06
Fruit symmetry (position B)	FSB	1.72	2.00	1.95	0.01
Stone position of maximum diameter (B position)	SMD	1.44	3.00	2.37	0.27
Stone apex shape(A)	SASH	1.00	1.67	1.04	0.03
Stone base shape (position A)	SBSH	1.88	3.00	2.38	0.09
Stone surface	SS	1.32	3.00	2.03	0.22
Stone distribution of grooves	SDG	1.00	2.00	1.59	0.21
Leaf length	LL	2.96	5.68	1.41	0.59
Leaf width	LW	0.91	2.53	1.19	0.03
Leaf shape (length/width)	LSH	2.51	5.22	3.83	0.55
Bending leaf	BL	1.4	3.96	2.48	0.53
Leaf shape index	LSHI	1.00	2.08	1.44	0.13

72 °C. PCR products were checked by agarose gel (2%) electrophoresis and then were separated on 6% denaturing polyacrylamide gels containing 1X TBE buffer and 7.5 M urea. Banding patterns were visualized using a silver staining method (Bassam et al., 1991). The gels were then dried by exposure overnight in the laboratory according to Sorkheh et al. (2007). Digital images of gels were made using an A4 scanner. Allele sizes were determined by DNA Ladder VIII (Roche).

For SSR diversity assessment at loci level, Power Marker V3.23 software (Liu, 2002) was used to calculate genetic diversity parameters: the average number of alleles per locus (Na), the observed heterozygosity (HO), and the expected heterozygosity or gene diversity (HE) per microsatellite locus. Polymorphism information content (PIC) for each locus (Botstein et al., 1980) and discriminating power (D) of each locus (Tessier et al., 1999) were also calculated.

The allelic richness, Nar, as a measure of the number of alleles per locus independent of sample size was calculated by FSTAT v. 2.9.3.2 (www.unil.ch/izea/softwares/fstat.html), while the number of private alleles (Npr) per population was assessed by MICROSAT (Minch et al., 1997). GENEPOP 4.0 (Raymond and Rousset, 1995) was used to estimate the inbreeding coefficients, FIS, and to test population genotypic frequencies across all loci for conformance to Hardy–Weinberg (HW) expectations (multilocus test). Pairwise FST and their respective P-values for significant differences from zero were calculated in FSTAT.

2.4. Combined use of morphoagronomic and molecular data

Phenotypic dissimilarities based on qualitative traits between pairs of individual trees were calculated using the proportion-of-shared-alleles distance as implemented in MICROSAT. In our case, as each individual tree can have only one state for a given trait, the results obtained by using the proportion-of-shared-alleles distance formula are identical to those obtained by simple matching coefficient: DSM = 1 – SSM = 1 – (m/n), where m is the number of qualitative morphological traits shared between a pair of accessions and n is the total number of traits according to Belaj et al. (2011).

Euclidean distances were calculated between all pairs of individual trees based on standardized values of ten quantitative agronomic traits and the distance matrix was used for cluster analysis (UPGMA) in NTSYS-pc version 2.10s (Rohlf, 2005).

The proportion-of-shared-alleles distance (Bowcock et al., 1994) between pairs of individuals was calculated using MICROSAT. Cluster analysis based on distance matrix was performed using the unweighted pair group method with arithmetic mean (UPGMA) as implemented in NTSYS-pc version 2.10s (Rohlf, 2005). The reliability of the UPGMA topology was assessed via bootstrapping (Felsenstein, 1985) over 1000 replicates generated by MICROSAT. Pairwise distance matrices among individual trees as obtained based on (A) qualitative morphological, (B) quantitative agronomic, and (C) molecular data were compared by calculating correlation coefficients and by performing Mantel's test (Mantel 1967). The randomization procedure as implemented in NTSYS-pc included 1000 permutations.

A model-based clustering method was applied to multilocus microsatellite data to infer genetic structure and to define the number of clusters (gene pools) in the dataset using STRUCTURE 2.1 software (Pritchard et al., 2000; Falush et al., 2003). Given a value for the number of clusters, this method assigns individual genotypes from the entire sample to clusters in a way in which linkage disequilibrium (LD) is maximally explained. Ten runs of STRUCTURE were performed by setting the number of clusters (K) from 1 to 14 (one more than the number of sampled populations). Each run consisted of a burn-in period of 200,000 steps followed by 106 Monte Carlo Markov chain replicates, assuming an admixture model and correlated allele frequencies. No prior information was used to define the clusters. The choice of the most likely number of clusters (K) was carried out by calculating an ad hoc statistic, DK, based on the rate of change in the log probability of data between successive K values, as described by Evanno et al. (2005).

2.5. Components of the diversity among and within populations

Both the phenotypic dissimilarity matrix (based on qualitative traits) and genetic distance matrix (based on molecular data) were subjected to analysis of molecular variance (AMOVA; Excoffier et al., 1992) using ARLEQUIN version 2.000 (Schneider et al., 2000). AMOVA was used to partition the total diversity among populations and within populations. The variance components were tested statistically by nonparametric randomization tests using 10,000 permutations.

3. Results

3.1. Morphoagronomic characterization

The 18 qualitative traits included in the study exhibited considerable morphological variability with a number of observed states per trait ranging from 1 (monomorphic traits) to 4 (Table 3). Three traits, fruit position of maximum diameter in B position (FMD), stone symmetry in B position (SSB), and stone surface (SS), did not show any variability in the 20 olive cultivars under study. These three monomorphic traits were excluded from further analysis. Regarding the variability of the remaining 20 traits, leaves were mostly short in length (LBL), medium

in width (LBW), and with elliptical shape (LSH). Fruits of most olives trees were long shaped (FSH), and with low weight (FW). The shapes of both the apex (FASH) and base (FBSH) of the fruits were found to be round in the majority of wild olive trees. Our observations revealed that stones of most of the olive trees sampled were slightly asymmetrical (SSA), with rounded shape of apex (SASH) and base (SBSH) and with low weight (SW) and high number of grooves (SNG), which were irregularly distributed (SDG). As far as their shape (SSH), all but three trees displayed almost equally oval and elliptical states.

Relatively high values of Hj and Dj were observed for the qualitative morphological traits measured. Fruit symmetry (FS), stone shape (SSH), stone apex shape A (SASH), stone distribution of grooves (SDG), stone shape (position A) (SSH), stone surface (SS), and leaf blade width (LBW) were the most discriminative traits and showed the highest values of diversity, giving a proportion of distinguishable pairs higher than other traits. Thus, the olive trees included in the study could be discriminated and identified by means of 18 traits.

The maximum coefficient of variation for the quantitative agronomic traits was observed in wet fruit weight (FWW, 13.45%), while the lowest was found for stone width (SWI, 3.18%). Fruit shape index (FSHI), leaf length, leaf width, and lenticel size (LS) also showed variability. The study of the variance components (Table 3) indicated that most of the variance found within the 11 agronomical traits was significantly related to the olive trees under study. The total variance explained by this source of variation ranged from 10.45% (SW) to 68.13% (FW). Differences among olive trees were also important for the traits of leaf blade width (LBW) and fruit weight (FW).

The correlation analysis between the 10 quantitative traits showed a strong association among all the fruit and stone dimensions (Table 4). Interestingly, a high and significant correlation of fruit and stone dimensions and wet fruit weight was also observed. Leaf dimensions were lowly correlated among them and with the rest of traits measured. The correlation analysis of the 18 qualitative traits is presented in Table 5.

On the other hand, Supplementary Tables 1 and 2 present the standardized canonical coefficient values of the assayed quantitative traits. The first canonical variable accounted for 28.47% of the total variation existing among the groups, while the second and the third canonical variables reduced the measure of the total variation further by 21.27% and 11.53%, respectively. On the other hand, the values of the canonical variant coefficients showed that stone shape index (SSHI) is a major discriminating coefficient among the clusters with the wet weight (WW), fruit lenticels (FL), and stone position of maximum diameter (SMD) making smaller contributions. The second canonical variant was found to be largely dominated by the fruit traits including fruit length (FL, 0.82), fruit width (FW, 0.87), and fruit wet weight (WW, 0.91). The third canonical variant revealed that fruit position of maximum diameter (FMD, 0.70) and stone position of maximum diameter (SMD, 0.58) play much larger roles in separating genotypes based on the studied traits.

Table 4. Correlation coefficients for 10 quantitative agronomic traits obtained from 20 olive trees based on individual tree averages.

	FSHI	FL	FW	WW	SSHI	SL	SW	LL	LW	LSHI
FSHI	1.0									
FL	0.30	1.0								
FW	–0.53*	0.63**	1.0							
WW	–0.25	0.82**	0.93**	1.0						
SSHI	0.87**	0.62*	–0.15	0.16	1.0					
SL	0.58*	0.92**	0.33	0.58*	0.80**	1.0				
SW	–0.60*	0.34	0.78**	0.61*	–0.47	0.13	1.0			
LL	0.08	–0.38	–0.41	–0.45	–0.13	–0.27	–0.22	1.0		
LW	0.04	0.02	–0.04	–0.11	0.05	0.07	0.05	0.29	1.0	
LSHI	0.06	–0.38	–0.37	–0.34	–0.14	–0.32	–0.30	0.67**	–0.49	1.0

**Significant at $P < 0.01$.
*Significant at $P < 0.05$.

Table 5. Correlation coefficients for 18 qualitative agronomic traits obtained from 20 olive trees based on individual tree averages.

	FSH	SFA	FMD	FASH	FBSH	LAFCC	FLE	LS	SSH	SSA	SFB	SMD	SASH	SBSH	SS	SG	BL	LSH
FSH	1.0																	
SFA	0.67**	1.0																
FMD	0.55*	0.56*	1.0															
FASH	–0.49	–0.63**	–0.02	1.0														
FBSH	–0.19	–0.22	–0.43	–0.04	1.0													
LAFCC	–0.20	–0.57*	–0.19	0.57*	–0.23	1.0												
FLE	–0.20	0.04	–0.06	0.13	0.35	–0.08	1.0											
LS	–0.24	–0.45	0.04	0.44	0.08	0.48	0.17	1.0										
SSH	0.81**	0.44	0.35	–0.36	–0.23	0.001	–0.07	–0.04	1.0									
SSA	0.43	0.53*	0.43	–0.30	0.08	–0.54*	0.37	0.25	0.23	1.0								
SFB	–0.20	0.18	0.27	0.10	–0.07	–0.16	0.63**	0.35	–0.27	0.51*	1.0							
SMD	0.16	0.02	0.74**	0.35	–0.47	0.11	–0.15	0.21	0.22	0.02	0.03	1.0						
SASH	–0.42	–0.39	–0.16	0.39	0.39	0.39	0.39	0.46	–0.42	–0.14	0.26	–0.19	1.0					
SBSH	–0.43	–0.31	–0.49	0.36	0.32	0.39	0.24	0.003	–0.28	–0.24	–0.03	–0.39	0.42	1.0				
SS	–0.52*	–0.27	–0.43	0.16	0.24	0.06	0.25	0.31	–0.51*	–0.40	0.14	–0.30	0.43	0.30	1.0			
SDG	0.46	0.31	0.01	–0.25	0.29	–0.28	0.32	–0.13	0.46	0.50*	0.06	–0.22	–0.33	–0.06	–0.54*	1.0		
BL	–0.04	–0.26	–0.25	–0.04	–0.07	–0.06	–0.50*	–0.31	0.07	–0.18	–0.53*	0.01	–0.25	–0.08	–0.31	0.01	1.0	
LSH	0.22	+0.06	0.46	0.13	–0.47	0.16	–0.61*	–0.03	0.10	0.10	–0.10	0.43	–0.19	–0.16	–0.24	–0.39	0.76	1.0

**Significant at P < 0.01.

*Significant at P < 0.05.

3.2. Molecular markers

All 11 microsatellite loci were polymorphic and a total of 93 alleles were found in the 20 olive cultivars analyzed. The average number of alleles per each locus was 8.45, ranging from 3 at locus ssrOeUA–DCA18 to 14 at locus EMO–90 (Table 6). The discrimination power (Dj) varied from 0.789 (ssrOeUA–DCA18) to 0.996 (EMO–90), with an average of 0.9232. All microsatellite loci displayed high values of PIC (from 0.62 to 0.95), permitting the identification of all the trees. Relatively high levels of SSR diversity were also observed for the olive trees under study (Table 6).

3.3. Combination of morphoagronomic and SSR data

Mantel's matrix correspondence test was used to compare the distance matrices based on qualitative and quantitative morphological traits as well as SSR markers. The correlation coefficient between matrices based on qualitative and quantitative morphological traits was significant but relatively low ($r = 0.53$; $P < 0.001$). No significant correlation was found between matrices based on SSR markers and qualitative traits ($r = 0.02$; $P = 0.18$) nor between matrices based on SSR markers and quantitative traits ($r = 0.01$; $P = 0.33$).

The three markers gave different dendrograms in which the 20 olive cultivars were separated into two main groups with different numbers of subgroups. In general, the grouping of the olive cultivars did not reflect a close relationship with their sampling sites. However, a certain tendency of grouping of the olive trees according to their geographical distribution was observed in the dendrograms.

In the tree obtained with qualitative morphological traits based on 16 of the 20 cultivars studied, 4 genotypes of 20 cultivars were studied since it lacked some traits (such as flowering, fruit), respectively. The cluster analysis based on morphological characteristics data were obtained from 16 cultivars (Figure 1A) and two main groups were observed: Group 1, including 13 olive trees, and Group 2, including the rest (3). Thirteen (81%) out of 16 olive trees in Group 1 were from Syria, Iran, and Europe, while the rest of the olive trees (3) from Iran and Europe clustered together in Group 2. This dendrogram and the one obtained with quantitative traits (Figure 1B) showed a certain degree of similarity in dendrogram topologies, though differences in the positions of olive trees within and between the main groups were observed.

The dendrogram obtained with SSR markers (Figure 2) was to some extent less similar at the subgroup level to the ones obtained with qualitative and quantitative traits.

Table 6. Source, repeat motifs, size ranges, number of alleles (NA), number of genotypes (NG) (or banding patterns), size range (SR), polymorphic information content (PIC), number of cultivars with null alleles as observed (NCNA), and number of unique genotypes (NUG) at 11 SSR loci in the 20 olive genotypes.

Source	Locus	Repeat motif	SR	PIC	NA	NG	NUG	NCNA	Dj
Sefc et al. (2000)	ssrOeUA–DCA09	(GA)23	164–204	0.95	8	26	11	0	0.806
Sefc et al. (2000)	ssrOeUA–DCA16	(GT)13(GA)29	122–210	0.81	11	23	13	0	0.982
Sefc et al. (2000)	ssrOeUA–DCA03	(GA)19	227–253	0.96	8	22	15	4	0.992
Sefc et al. (2000)	ssrOeUA–DCA05	(GA)15	195–211	0.94	8	5	8	2	0.953
Sefc et al. (2000)	ssrOeUA–DCA18	(CA)4CT(CA)3(GA)19	158–182	0.91	3	22	16	0	0.789
Sefc et al. (2000)	ssrOeUA–DCA14	(CA)18A6(TAA)7	170–187	0.62	5	10	16	3	0.942
Cipriani et al. (2002)	UDO99–043	(GT)12	166–222	0.91	10	25	14	0	0.989
Carriero et al. (2002)	GAPU101	(CT)9	175–215	0.89	9	14	18	18	0.924
Carriero et al. (2002)	GAPU103A	(TC)26	130–184	0.80	9	29	27	25	0.896
Carriero et al. (2002)	GAPU71B	GA(AG)6(AAG)8	121–144	0.92	8	34	8	8	0.934
De la Rosa et al. (2002)	EMO–90	(CA)7	184–196	0.95	14	18	2	10	0.946
Average				0.87	8.45	20.7	13.45	6.36	0.923

AMOVA analysis (Table 7) showed that most of the genetic diversity was attributable to differences among (38.67%) and within (58.08%) populations for molecular markers. However, significant P-values for among-population components were observed, suggesting the existence of weak population differentiation.

Tests for linkage disequilibrium performed for microsatellite loci as implemented in POPGENE did not show any significant association between them. Therefore, these molecular markers were considered to meet the assumptions for applying the Bayesian method implemented in the program STRUCTURE. The Q-matrix plot of STRUCTURE analysis (Figure 3) showed the presence of 3 subgroups, which was also supported by the Evanno test (K = 3)

On the basis of the molecular data, the results from Bayesian clustering analysis using STRUCTURE software confirmed the groupings that we observed in the UPGMA dendrogram and PCoA (Supplementary Figure 1). The most likely value of K (as chosen by Evanno's DK method) in Bayesian clustering analysis was 3 and this indicates the division of variation into three clusters, indicating the most appropriate three main clusters within the samples studied and confirming the clustering of the UPGMA dendrogram and PCoA. The first cluster (red color) consisted of Iranian and Syrian olive trees. Syrian olive trees were placed in the second cluster (green color), while European olive trees were placed in the third cluster (violet color) with one olive tree include as an outgroup (yellow color).

4. Discussion

4.1. Morphoagronomic traits and SSR diversity

The present study has revealed a considerable level of diversity in olive trees from different geographical distributions (Europe, Syria, Iran, USA) at both morphoagronomic and molecular levels. The results of morphological characterization coincide with previous studies carried out for olive trees (García-Donas Díaz, 2001; Mulas et al., 2004; Hannachi et al., 2008) and related subspecies (Hannachi et al., 2009; García-Verdugo et al., 2010; Koehmstedt et al., 2011). However, contrary to the results obtained by Hannachi et al. (2008) and in total agreement with those obtained by García-Donas Díaz (2001) and Belaj et al. (2011), we found that the olive genotypes under study showed only smooth stone surfaces.

The significant and high correlations recorded in the sampling for each quantitative morphological trait indicate the consistency of data and low environmental influence. To the best of our knowledge, this is the first study performed with morphological descriptors in this olive tree germplasm.

As in the case of qualitative traits, similar results to those previously reported in studies with wild material were obtained for leaf and fruit traits (Mulas, 1999; García-Donas Díaz, 2001; Mulas et al., 2004; Hannachi et al., 2008; Belaj et al., 2011).

Correlations between quantitative traits showed a strong association among the fruit and stone dimensions, as previously reported in studies with wild (Hannachi

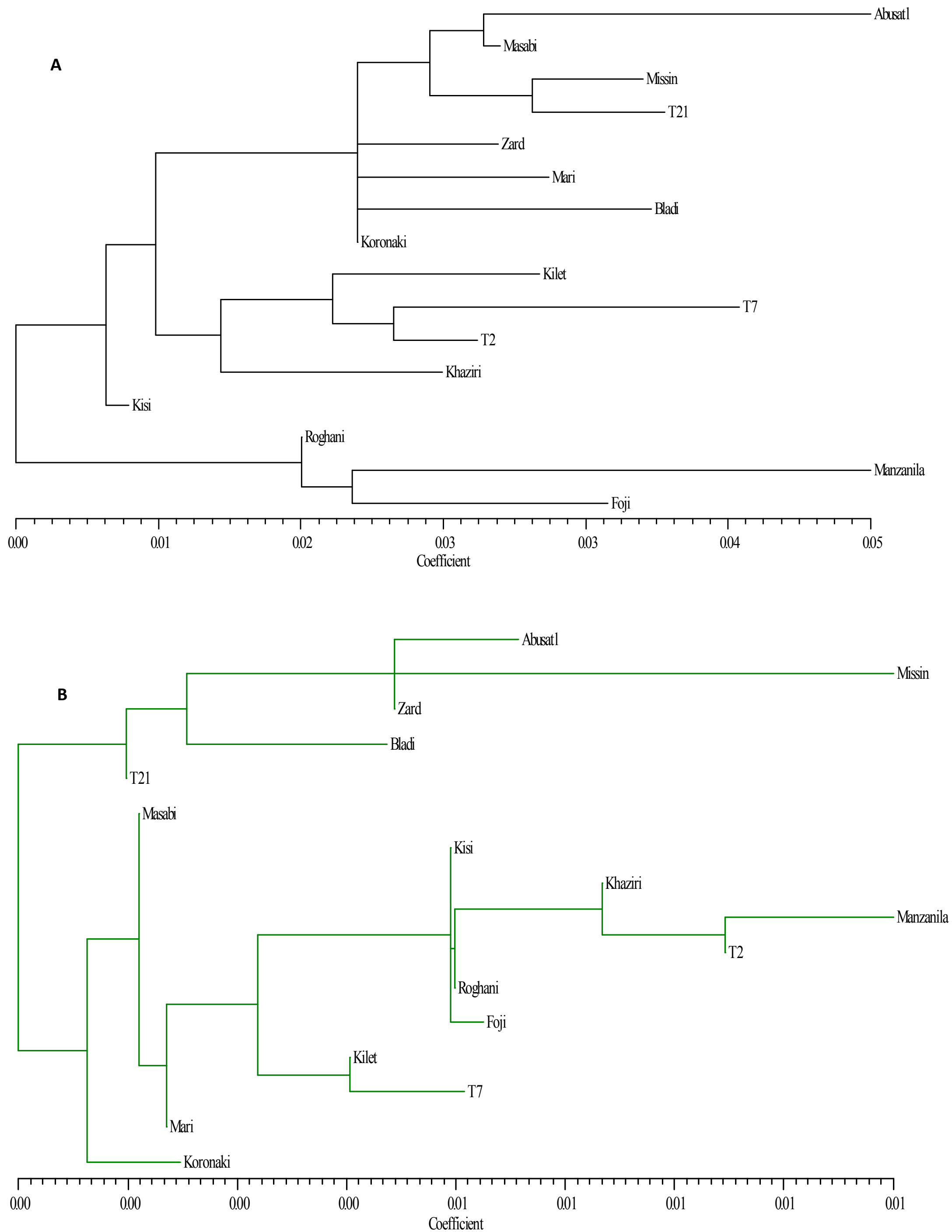


Figure 1. UPGMA dendrograms based on (A) qualitative morphological data and (B) quantitative traits of 16 of 20 olive cultivars.

et al., 2008) and cultivated olive trees (Cantini et al., 1999). Contrary to the results obtained with cultivated material (Del Río et al., 2005) or progenies from crosses (León et al., 2004a), a high and significant correlation of fruit dimensions and stone was observed. This could be of interest when using olives in breeding programs. In addition, the PCA performed for agronomic traits was useful for identifying the most important traits associated with variation among the olive trees.

Analysis of variance evidenced that the major source of variance variability for the 10 quantitative traits was related to the individual olive trees under study, which was

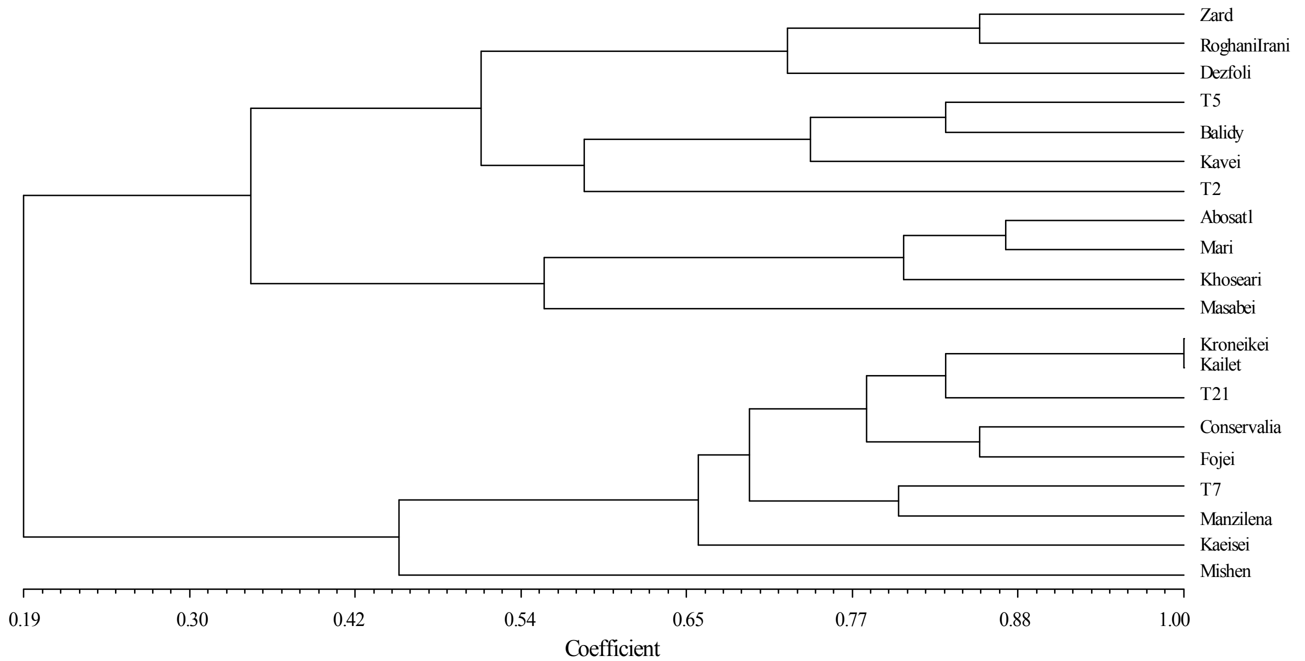


Figure 2. UPGMA dendrogram showing the genetic diversity of 20 the olive trees based on microsatellite markers.

Table 7. Analysis of molecular variance (AMOVA) for 20 olive trees based on SSR markers.

Source of variation	df	Variance components	% Total variance	P-value[a]
Variance among groups[b]	1	0.6623	4.87	P < 0.001
Variance among population	1	8.4783	38.67	P < 0.001
Variance within population	17	12.5640	58.08	P < 0.001
Total	19			

[a]Significance of variance component expressed as the probability of obtaining a more extreme random value computed from nonparametric procedures (1000 data permutations).
[b]Two groups consist of the accessions of olive trees from Europe and Asia; for more details, see Section 2.

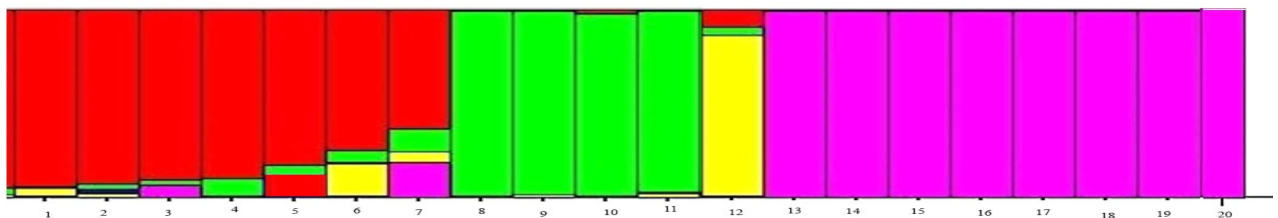


Figure 3. Q-matrix of STRUCTURE analysis based on SSR data. Numbers are 1 to 7 (Iranian and Syrian olive trees), 8 to 11 (Syrian olive trees), and 12 to 20 (American and European olive trees). Numbers are according to samples in Table 1.

much higher than the year effect. This is in total agreement with AMOVA analysis based on SSR data and previous studies in olive, suggesting that most of the variation is maintained within populations (León et al., 2004b; Baldoni et al., 2006; Belaj et al., 2007, 2010). Apart from revealing the diversity found in the olive germplasm, at both qualitative and quantitative levels, the use of these descriptors evidenced differences between wild and

cultivated olive trees. For instance, average values of 0.46 g, 2.09, and 20.30% were obtained for fruit weight, flesh/stone ratio, and oil content of dry matter respectively for wild olives, while these values were 4.0 g, 7.0, and 43.8% and 2.9 g, 6.4, and 43.7% for olive cultivars and progenies, respectively (León et al., 2004b; Del Río et al., 2005). In spite of that, it is worth mentioning that olive trees with fruit weights (3.07 g) comparable to the values found for some olive cultivars were also found, although a possible feral origin of these trees cannot be excluded (Sedgley, 2000; Hannachi et al., 2008). The increase of fruit size might have been an important criterion of multilocal selection during the domestication process (Besnard et al., 2001; Belaj et al., 2002; Hagidimitirou et al., 2005; Breton et al., 2006). In addition, leaves of olive trees in our study would be classified as being of short length and medium width in comparison to the cultivated germplasm (Barranco et al., 2005).

As expected (Breton et al., 2006; Belaj, et al. 2007, 2010, 2011; Erre et al., 2010), high levels of molecular diversity were found at both individual and populations levels by means of SSR markers. The total number of alleles, PIC values for each SSR locus, and diversity parameters found in the olive trees were similar to those of previous studies carried out in wild and cultivated olive trees (Breton et al., 2006; Hannachi et al., 2008; Erre et al., 2010).

4.2. Comparison between marker systems

The olive trees included in this study were clearly discriminated by the three marker systems considered (SSRs and qualitative and quantitative morphological traits). Similar to previous studies (Sedgley, 2000; García-Donas Díaz, 2001; Mulas et al., 2004; Hannachi et al., 2008), the qualitative morphological traits used in our research proved to be useful for the distinguishing of olive trees. The combinations of the five most discriminative traits [fruit symmetry (FS), stone shape (SSH), fruit flesh/stone ratio (FST), fruit shape index (FSHI), and leaf shape index (LSHI)] made possible the discrimination of all but 5% of possible pairs of olive trees.

In spite of the successful use of qualitative morphological descriptors for olive tree discrimination, consistently lower values of discrimination power were observed in comparison to the SSR primers used in the study. The efficiency of a given trait/primer depends on the number of states/fragments it generates, as well as on their frequency.

Similar estimates of genetic relationships among wild olive trees were obtained for qualitative and quantitative morphology traits, while no significant correlations between morphoagronomic traits and molecular data were found. Although there was sufficient variability to discriminate all the wild olive trees by means of both morphological and agronomical data, the dendrogram based on SSR markers did identify a clearer structure of the tree grouping, suggesting that DNA markers are more informative in depicting genetic relationships.

Each marker system measures different aspects of this genetic variability and this may explain the lack of consistency in genetic diversity and relationship studies. The detected DNA variation, which is neutral, was often not correlated to the phenotypic and agronomical variation of olive cultivars (Hagidimitirou et al., 2005; Corrado et al., 2009; Rao et al., 2009). However, in spite of the powerfulness of SSR markers to detect genetic variability and genetic relationships, they should not be seen as a substitute for traditional morphoagronomic descriptors (Karp et al., 1997).

All these marker systems should be considered as complementary tools to provide a more complete understanding of the diversity available in wild olive populations and the ways in which it can best be used for olive breeding and conservation strategies.

4.3. Breeding perspectives

The traditional method of agromorphological plant characterization is a common step in plant breeding for selection of parents and it also represents the first choice used for describing and classifying the germplasm. In this sense, our results give some insights into the potential value of olive trees as a source of morphoagronomically interesting traits, which have only occasionally been evaluated (Baldoni and Belaj, 2009). Our results indicate a high level of morphoagronomic and SSR variability that is harbored in the olive collection. This variability would be interesting for broadening the genetic base of olive breeding programs.

This collection will be characterized at molecular and phenotypic levels and evaluated for agronomic traits to test the possibility of introgression of new and superior alleles into cultivated varieties.

Acknowledgments

This study was supported by Shahid Chamran University of Ahvaz, Iran. We thank the anonymous reviewer for his/her most helpful comments. We are grateful to Kh Chenaneh-Hanoni for kind help in undertaking this study. Thanks are also due to Angjelina Belaj, Centro IFAPA 'Alameda del Obispo', Cordoba, Spain, for helpful comments on an earlier draft of the manuscript. This study was supported by the project "Fingerprinting of olive trees collection" of Shahid Chamran University of Ahvaz.

References

Abdessemed S, Muzzalupo L, Benbouz H (2015). Assessment of genetic diversity among Algerian olive (*Olea europaea* L.) cultivars using SSR marker. Sci Hortic 192: 10-20.

Albertini E, Torricelli R, Bitocchi E, Raggi L, Marconi G, Pollastri L, Di Minco G, Battistini A, Papa R, Veronesi F (2011). Structure of genetic diversity in *Olea europaea* L. cultivars from central Italy. Mol Breed 27: 533-547.

Baldoni L, Belaj A (2009). Olive. In: Vollmann J, Rajcan I, editors. Oil Crops. Berlin, Germany: Springer, pp. 397-421.

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.

Barranco D, Cimato A, Fiorino P, Rallo, L, Touzani A, Castaneda C, Serafíni F, Trujillo I (2000). World Catalogue of Olive Varieties. Madrid, Spain: International Olive Oil Council.

Barranco D, Trujillo I, Rallo L (2005). Elaiografía Hispánica. In: Rallo L, Barranco D, Caballero JM, del Rio C, Martín A, Tous J, Trujillo I, editors. Variedades de olivo en Espana. Junta de Andalucía. Madrid, Spain: MAPA-Ediciones Mundi-Prensa, pp. 48-76 (in Spanish).

Bartolini G, Prevost G, Messeri C, Carignani C (2005). Olive Germplasm: Cultivars and World-wide Collections. Rome, Italy: FAO/Plant Production and Protection.

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991). Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 196: 80-83.

Belaj A, Leon L, Satovic Rosa R (2011). Variability of wild olives (*Olea europaea* subsp. *europaea* var. *sylvestris*) analyzed by agro-morphological traits and SSR markers. Sci Hortic 129: 561-569.

Belaj A, Munoz-Diez C, Baldoni L, Porceddu A, Barranco D, Satovic Z (2007). Genetic diversity and population structure of wild olives from North-Western Mediterranean assessed by SSR markers. Ann Bot 100: 449-458.

Belaj A, Munoz-Diez C, Baldoni L, Satovic Z, Barranco D (2010). Genetic diversity and relationships of wild and cultivated olives at regional level in Spain. Sci Hortic 124: 323-330.

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.

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

Botstein D, White RL, Sholnick M, David RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32: 314-331.

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

Bracci T, Busconi M, Fogher C, Sebastiani L (2011). Molecular studies in olive (*Olea europaea* L.): overview on DNA markers applications and recent advances in genome analysis. Plant Cell Rep 30:449-462.

Breton C, Tersac M, Bervillé A (2006). Genetic diversity and gene flow between the wild olive (oleaster Olea europaea L.) and the olive: several Plio-Pleistocene refuge zones in the Mediterranean basin suggested by simple sequence repeats analysis. J Biogeogr 33: 1916-1928.

Cantini C, Cimato A, Sani G (1999). Morphological evaluation of olive germplasm present in Tuscany region. Euphytica 109: 173-181.

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.

Caruso T, Marra FP, Costa F, Campisi G, Macaluso L, Marchese A (2014). Genetic diversity and clonal variation within the main Sicilian olive cultivars based on morphological traits and microsatellite markers. Sci Hortic 180: 130-138.

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

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.

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 Oleacea. Primer note. Mol Ecol Notes 2: 265-267.

Del Río C, Caballero JM, García-Fernández MD, Tous J, Romero A (2005). Rendimiento graso de la aceituna. In: Rallo L, Barranco D, Caballero JM, del Rio C, Martín A, Tous J, Trujillo I, editors. Variedades de olivo en Espana. Junta de Andalucía. Madrid, Spain: MAPA-Ediciones Mundi-Prensa, pp. 347-356 (in Spanish).

Erre P, Chessa I, Munoz-Diez C, Belaj A, Rallo L, Trujillo I (2010). Genetic diversity and relationships between wild and cultivated olives in Sardinia as assessed by SSR markers. Genet Res Crop Evol 57: 41-54.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611-2620.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial restriction sites. Genetics 131: 479-491.

Falush D, Stephens M, Pritchard JK (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7: 574-578.

Felsenstein J (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

García-Donas Díaz M (2001). Caracterización morfológica, agronómica y elaiotécnica de los acebuches de la provincia de Cádiz. Trabajo profesional de fin de carrera. Cordoba, Spain: University of Cordoba (in Spanish).

García-Verdugo C, Forrest AD, Balaguer L, Fay MC, Vargas P (2010). Parallel evolution of insular *Olea europaea* subspecies based on geographical structuring of plastid DNA variation and phenotypic similarity in leaf traits. Bot J Linn Soc 162: 54-63.

Gemas VJV, Almadanim MC, Tenreiro R, Martins A, Fevereiro P (2004). Genetic diversity in the olive tree (Olea europaea L. subsp. europaea) cultivated in Portugal revealed by RAPD and ISSR markers. Genet Res Evol 51: 501-511.

Gomes S, Martins-Lopes P, Lima-Brito J, Meirinhos J, Lopes J, Martins A, Guedes-Pinto H (2008). Evidence of clonal variation in olive 'Verdeal-Transmontana' cultivar using RAPD, ISSR and SSR markers. J Hortic Sci Biotechnol 83: 395-400.

Guerin J, Mekuria G, Burr M, Collins G, Sedgley M (2003). Selection of olive cultivars. Acta Hortic 622: 231-234.

Hagidimitirou M, Katsiotis A, Menexes G, Pontikis C, Loukas M (2005). Genetic diversity of major Greek olive cultivars using molecular (AFLPs and RAPDs) markers and morphological traits. J Am Soc Hortic Sci 130: 211-217.

Hannachi H, Breton C, Msallem M, Ben El Hadj S, El Gazzah M, Bervillé A (2008). Differences between native and introduced cultivars as revealed by morphology of drupes, oil composition and SSR polymorphism; a case study in Tunisia. Sci Hortic 116: 280-290.

Hannachi H, Sommerlatte H, Breton C, Msallem M, El Gazzah M, Ben El Hadj S, Bervillé A (2009). Oleaster (var. *sylvestris*) and subsp. *cuspidata* are suitable genetic resources for improvement of the olive (*Olea europaea* subsp. *europaea* var. *europaea*). Genet Res Crop Evol 56: 393-403.

Ipek A, Barut E, Gulen H, Ipek M (2012). Assessment of inter- and intra-cultivar variations in olive using SSR markers. Sci Agric 69: 327-335.

Karp A, Kresovich S, Bhat KV, Ayad WG, Hodgkin T (1997). Molecular Tools in Plant Genetic Resources Conservation: A Guide to the Technologies. IPGRI Technical Bulletin No. 2. Rome, Italy: International Plant Genetic Resources Institute.

Koehmstedt AM, Aradhya MK, Soleri D, Smith JL, Polito VS (2011). 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.

León L, Martín LM, Rallo L (2004a). Phenotypic correlations among agronomic traits in olive progenies. J Am Soc Hortic Sci 129: 271-276.

León L, Rallo L, Del Rio C, Martín LM (2004b). Variability and early selection on the seedling stage for agronomic traits in progenies from olive crosses. Plant Breed 123: 73-78.

Lewontin RC (1972). The apportionment of human diversity. Evol Biol 6: 381-398.

Liu J (2002). POWERMARKER. A Powerful Software for Marker Data Analysis. Raleigh, NC, USA: North Carolina State University Bioinformatics Research Center.

Lopes MS, Mendonc AD, Sefc MK, Gil FS, Da Camara Machado A (2004). Genetic evidence of intra-cultivar variability within Iberian olive cultivars. Hort Sci 39: 1562-1565.

Lumaret R, Ouazzani N, Michaud H, Vivier G, Deguilloux MF, Di Giusto F (2004). Allozyme variation of oleaster populations (wild olive tree) (*Olea europaea* L.) in the Mediterranean basin. Heredity 92: 343-351.

Mantel N (1967). The detection of disease clustering and a generalized regression approach. Cancer Res 27: 209-220.

Marra FP, Caruso T, Costa F, Di Vaio C, Mafrica R, Marchese A (2013). Genetic relationships, structure and parentage simulation among the olive tree (*Olea europaea* L. subsp. *europaea*) cultivated in Southern Italy revealed by SSR markers. Tree Genet Genomes 9: 961-973.

Martins-Lopes P, Gomes S, Lima-Brito J, Lopes J, Guedes-Pinto H (2009). Assessment of clonal genetic variability in *Olea europaea* L. 'Cobrançosa' by molecular markers. Sci Hortic 123: 82-89.

Minch E, Ruiz-Linares A, Goldstein D, Feldman M, Cavalli-Sforza LL (1997). MICROSAT. A Computer Program for Calculating Various Statistics on Microsatellite Allele Data. Ver. 1.5d. Stanford, CA, USA: Stanford University.

Mulas M (1999). Characterization of olive wild ecotypes. Acta Hortic 474: 121-124.

Mulas M, Fadda A, Cauli E (2004). Prime osservazioni su cloni di oleastro (*Olea europaea* var. sylvestris Hoff E Link) selezionati per l'utilizzo forestale. Italus Hortus 11: 214-217 (in Italian).

Muzzalupo I, Chiappetta A, Benincasa C, Perri E (2010). Intra-cultivar variability of three major olive cultivars grown in different areas of central-southern Italy and studied using microsatellite markers. Sci Hortic 126: 324-329.

Noormohammadi Z, Trujillo I, Belaj A, Ataei A, Hosseini-Mazinan M (2014). Genetic structure of Iranian olive cultivars and their relationship with Mediterranean's cultivars revealed by SSR markers. Sci Horti 178: 175-183.

Omrani-Sabbaghi A, Shahriari M, Falahati-Anbaran M, Mohammadi SA, Nankali A, Mardi M, Ghareyazie B (2007). Microsatellite markers based assessment of genetic diversity in Iranian olive (*Olea europaea* L.) collections. Sci Hortic 112: 439-447.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Rao R, La Mura M, Corrado G, Ambrosino O, Foroni I, Perri E, Pugliano G (2009). Molecular diversity and genetic relationships of southern Italian olive cultivars as depicted by AFLP and morphological traits. J Hortic Sci Biotechnol 84: 261-266.

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

Rohlf FJ (2005). NTSYS-PC. Numerical Taxonomy and Multivariate Analysis System, Version 2.10s. Setauket, NY, USA: Department of Ecology and Evolution, State University of New York.

SAS Institute (2004) SAS/STAT® 9.1 User's Guide. Cary, NC, USA: SAS Institute Inc.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN. A Software for Population Genetic Data Analysis. Version 2000. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.

Sedgley M (2000). Wild Olive Selection for Quality Oil Production, Vol. 116. Barton, Australia: Rural Industries Research and Development Corporation.

Sedgley M, Wirthensohn M (2000). The Australian olive improvement program. Olivae 83: 27-30.

Sorkheh K, Shiran B, Gradziel TM, Epperson BK, Martinez-Gomez P, Asadi E (2007). Amplified fragment length polymorphism as a tool for molecular characterization of almond germplasm: genetic diversity among cultivated genotypes and related wild species of almond, and its relationships with agronomic traits. Euphytica 156: 237-344.

Strikic F, Mavsar DB, Perica S, Cmelik Z (2010). Genetic variation within the olive (*Olea europaea* L.) cultivar Oblica detected using amplified fragment length polymorphism (AFLP) markers. Afr J Biot 9: 2880-2883.

Sutherland EK, Hale BJ, Hix DM (2000). Defining species guilds in the Central Hardwood Forest, USA. Plant Ecol 147: 1-19.

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.

Trujillo I, Ojeda MA, Urdiroz NM, Potter D, Barranco D, Rallo L, Diez CM (2014). Identification of the Worldwide Olive Germplasm Bank of Córdoba (Spain) using SSR and morphological markers. Tree Genet Genomes 10: 141-155.

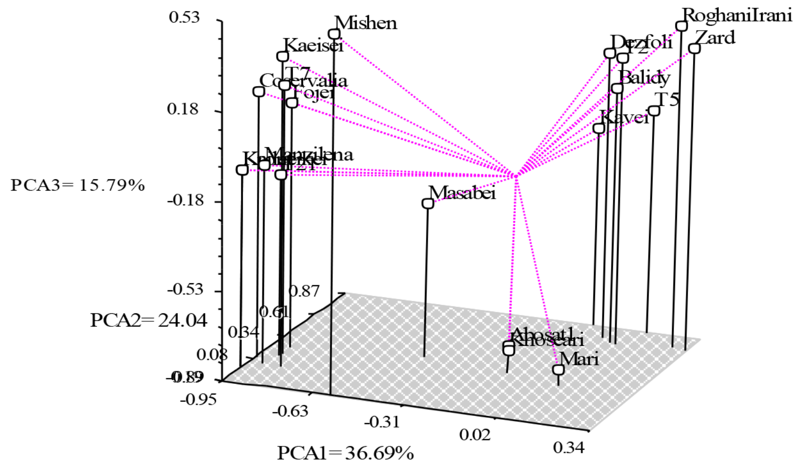
Zaher H, Boulouha B, Baaziz M, Sikaoui L, Gaboun F, Sripada M (2011). Morpho-logical and genetic diversity in olive (*Olea europaea* subsp. europaea L.) clones and varieties. Plant Omics Journal 4: 370-376.

Supplementary Table 1. Eigenvalues, percentage of variability explained by each coordinate, and accumulated variability for morphological data

Principal	Eigenvalue	% variability	Accumulated coordinate variance
1	7.973	28.474	28.474
2	5.958	21.279	49.752
3	3.229	11.532	61.284
4	2.598	9.279	70.563
5	1.915	6.839	77.402
6	1.640	5.856	83.258
7	1.166	4.163	87.421

Supplementary Table 2. Standardized canonical coefficients of the first four canonical discriminant functions (CDFs) of 28 morphological variables and relative % of variance explained.

Traits	Canonical variable			
	CDF1	CDF2	CDF3	CDF4
Fruit shape index	0.891	–0.175	0.164	–0.137
Fruit length	0.495	0.827	–0.154	–0.015
Fruit width	–0.292	0.875	–0.261	0.082
Weight of wet	0.028	0.910	–0.295	0.024
Fruit shape	0.83	–0.141	0.090	–0.231
Fruit symmetry (position A)	0.506	–0.400	0.053	–0.025
Fruit position of maximum diameter (B position)	0.444	–0.124	0.703	–0.141
Fruit apex shape (position A)	–0.454	–0.058	0.154	0.105
Fruit base shape (position A)	–0.294	0.038	–0.509	0.626
Location appearance of fruit color change	0.103	0.289	0.378	0.215
Fruit lenticels	0.044	0.604	–0.320	0.324
Lenticel size	–0.145	0.570	0.276	0.195
Stone shape index	0.962	0.127	–0.055	–0.097
Stone length	0.685	0.632	–0.091	–0.117
Stone width	–0.592	0.729	–0.076	0.053
Stone shape (position A)	0.944	0.065	–0.107	–0.244
Stone symmetry (position A)	0.387	–0.229	–0.054	0.005
Fruit symmetry (position B)	–0.301	0.085	0.115	0.035
Stone position of maximum diameter (B position)	0.062	–0.025	0.586	–0.489
Stone apex shape(A)	–0.285	0.059	0.108	0.772
Stone base shape (position A)	–0.224	–0.024	–0.266	0.739
Stone surface	–0.477	0.437	–0.050	0.621
Stone distribution of grooves	0.515	–0.024	–0.687	–0.153
Leaf length	–0.069	–0.314	0.782	–0.038
Leaf width	0.041	–0.086	0.110	–0.015
Leaf shape (length/width)	–0.069	–0.260	0.620	–0.051
Bending leaf	–0.163	–0.515	–0.243	–0.453
Leaf shape index	0.016	–0.306	0.670	–0.208



Supplementary Figure 1. Principal coordinates map for the first, second, and third principal coordinates estimated for SSR markers data using genetic similarity matrix for 20 olive trees.