Libyan Journal of Science & Technology 9:2 (2019) 129–141

Faculty of Science - University of Benghazi

Libyan Journal of Science & Technology

journal home page: www.sc.uob.edu.ly/pages/page/77

Characterization and identification of Libyan olive diversity using microsatellite markers

Salem Abdul Sadeg*[,a], Gayle Volk[c], Christopher Richards[c], and Harrison Hughes[b]

[a]*Biology Department Elmergib University, Masallatah, Libya,*

[b]*Horticulture & Landscape Architecture, Colorado State University, Fort Collins, Colorado*

[c]*National center for genetic resource preservation, Fort Collins, Colorado.*

Highlights

- **Combination of morphological traits and molecular data were highly useful to separate closely related genotypes within Libyan olive landraces.**
- **The denominations of homonyms and synonyms or mislabeling were more frequently within landraces than other cultivated and wild types.**
- **The wild types were more closely related to the introduced genotypes than Libyan olive landraces.**

ARTICLE INFO

Article history:
Received 10 June 2017
Revised 03 March 2019
Accepted 10 March 2019
Available online 16 June 2019

Keywords:
SSR Markers, Olive tree, Microsatellite markers, and Genetic diversity.

* Corresponding author:
E-mail address: salem.elleaga@yahoo.com
S. Abdul

ABSTRACT

Ten microsatellite markers were used to differentiate and evaluate the relationships among a total of 91 olive genotypes (39 local cultivated, 36 introduced cultivars and 16 wild types) collected in Libya. A total of 109 alleles were identified, with the number of alleles per locus ranging from 4 to 20 alleles. Three loci (UDO43, DCA16 and GAPU101) had the most alleles across all loci with 20, 18 and 16, respectively. The wild types and introduced cultivars had greater numbers of alleles than the local cultivars. Six cases of duplicated genotypes, two cases of synonymy, and thirteen homonyms that were genetically distinct were observed in the Libyan collection. UPGMA clustering classified the accessions into two main distinct groups. The first group consisted of local genotypes and the second group included introduced and wild type accessions. Admixture analysis also clearly distinguished between local ancient landraces and wild genotypes. In general, using molecular data enables to separate the Libyan olive accessions based on their origin but not fruit use.

1. Introduction

Libyan olives (*Olea europaea* subsp. *europaea* var. *sativa* or *sylvestris*) have traditionally been evaluated by leaf, fruit, and seed morphological as well as phonological characteristics. It has been difficult to properly manage and conserve olive germplasm because of the problems associated with clearly distinguishing among cultivars. Further complicating identification of cultivars is the observation that wild populations have likely introgressed with locally adapted cultivars. There are more than 100 named olive cultivars are grown along the coastal region of Libya. Some of these cultivars are likely to be identical due to the historical renaming of material. This has led to the perception that numerous cultivars exist when in fact, they are actually synonyms or homonyms Morphological differences associated with specific environmental effects have also lead to a mistaken identification of the cultivars. The level of knowledge about cultivar origin, selection and molecular variability is limited because the identification of Libyan olive accessions has previously been based on phenotypic traits. Recently, morphological descriptions have improved, and are now considered to be complementary tools to molecular marker, aiding in olive cultivar identification. Using both morphological and molecular descriptors of genotypes within other crops (Corrado *et al*., 2009). This combination of techniques leads to more robust results (Leon *et al*., 2005). To date, SSR markers have not been used in combination with morphological data to evaluate and improve the collection of Libyan olive accessions as a genetic resource. In this work, SSR markers were used to differentiate and classify Libyan olive accessions.

2. Materials and methods

2.1 Collection sites and plant materials

Accessions were classified into three categories: 42 local cultivated varieties, 41 introduced cultivars of *Olea europaea* and 16 wild *Olea europaea var. sylvestris*. Leaf tissue was collected in 2009 and 2012. Most of the local cultivars (Libyan landraces) were collected from orchards of Masallatah city while the introduced cultivars were collected from Tharouna and Gharian government collections as well as from farmers in the Zaltin and Tripoli regions. The wild type accessions were collected from four different sites (S1, S2, S3 and S4) in the Green Mountain region (Fig. 1). Leaf samples were collected then immediately stored in containers with dry ice to prevent DNA degradation. They were then transferred to the National Medical Research Center in Tripoli–where they were washed with double distilled water and freeze-dried. Samples were then transported to the Horticulture Laboratory at Colorado State University in Fort Collins, CO. USA where they were stored at –80°C until DNA use.

2.2 Processing samples

Total genomic DNA was extracted from 100-200 mg lyophilized of tissue using the method of large-scale CTAB extraction was performed according to (Mace *et al*., 2003). This protocol was a modification of the CTAB procedure for obtaining purified genomic DNA. Twelve sets of primer pairs were selected (Table 1) because of their high resolution in discriminating polymorphism previous use in the identification of olive genotypes (Ercisli *et al*., 2011; Ercisli *et al*., 2012; Sefc *et al*., 2000; Baldoni *et al*., 2009; Sarri *et al*., 2006;

© 2019 University of Benghazi. All rights reserved. ISSN 2663-1407; National Library of Libya, Legal number: 390/2018

Carriero *et al*., 2002; Cipriani *et al*., 2002; Belaj *et al*., 2003; De La Rosa *et al*., 2002). These were multiplexed using multiplex manager 1.2 software (Guichoux *et al*., 2011) to minimize overlap among the markers and to maximize similarity in the annealing temperature of each primer combination to reduce the variation and a total number of PCR reactions. Each cycle of multiplex PCR amplification was performed with combinations of three different primers labeled with specific fluorescent dyes that incorporated during multiplex PCR amplification giving a specific color tag to each PCR product (Table 1).

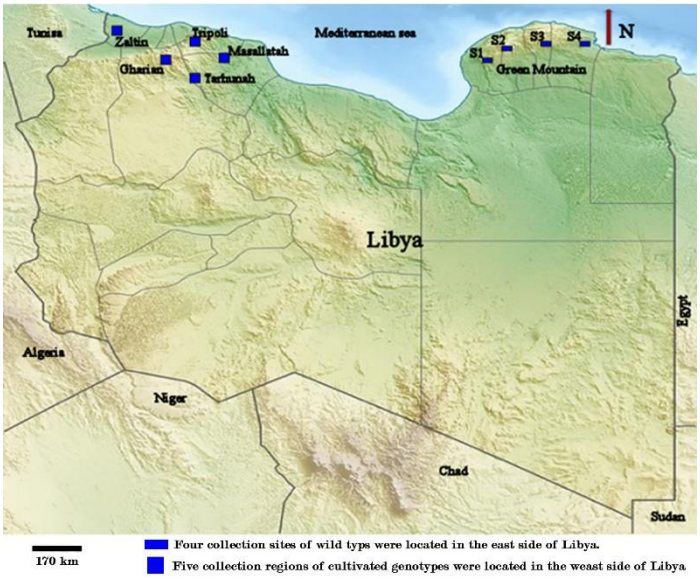

Fig. 1. Map of Libya that illustrates the collection sites of cultivated and wild olive.

Table 1

Locus names, Forward primers including nucleotide sequences and company references.

Locus Name	Forward dye label	Company	Primer sequence labeled with fluorescent probe (5' –3')
EMO-90-F	56-FAM	IDT	5'-/56-FAM/CAT CCG GAT TTC TTG CTT TT-3'
EMO-90-R	PIGtail	IDT	5'-GTT TCT T/AG CGA ATG TAG CTT TGC ATG T-3'
DCA3-F	56-FAM	IDT	5'-/56-FAM/CCC AAG CGG AGG TGT ATA TTG TTA C-3'
DCA3-R	PIGtail	IDT	5'-GTT TCT T/TG CTT TTG TCG TGT TTG AGA TGT TG-3'
DCA14-F	56-FAM	IDT	5'-/56-FAM/AAT TTT TTA ATG CAC TAT AAT TTA C-3'
DCA14-R	PIGtail	IDT	5'-GTT TCT T/TT GAG GTC TCT ATA TCT CCC AGG GG-3'
GAPU101-F	56-FAM	IDT	5'-/56-FAM/CAT GAA AGG AGG GGG ACA TA-3'
GAPU101-R	PIGtail	IDT	5'-GTT TCT T/GG CAC TTG TTG TGC AGA TTG-3'
DCA18-F	VIC	AB	5'-/VIC/AAG AAA GAA AAA GGC AGA ATT AAG C-3'
DCA18-R	PIGtail	IDT	5'-GTT TCT T/GT TTT CGT CTC TCT ACA TAA GTG AC-3'
DCA16-F	VIC	AB	5'-/VIC/TTA GGT GGG ATT CTG TAG ATG GTT G-3'
DCA16-R	PIGtail	IDT	5'-GTT TCT T/TT TTA GGT GAG TTC ATA GAA TTA GC-3'
DCA5-F	VIC	AB	5'-/VIC/AAC AAA TCC CAT ACG AAC TGC C-3'
DCA5-R	PIGtail	IDT	5'-GTT TCT T/CG TGT TGC TGT GAA GAA AAT CG-3'
DCA17-F	VIC	AB	5'-/VIC/GAT CAA ATT CTA CCA AAA ATA TA-3'
DCA17-R	PIGtail	IDT	5'-GTT TCT T/TA AAT TTT TGG CAC GTA GTA TTG G-3'
GAPU103A-F	PET	AB	5'-/PET/TGA ATT TAA CTT TAA ACC CAC ACA-3'
GAPU103A-R	PIGtail	IDT	5'-GTT TCT T/GC ATC GCT CGA TTT TTA TCC-3'
GAPU71B-F	PET	AB	5'-/PET/GAT CAA AGG AAG AAG GGG ATA AA-3'
GAPU71B-R	PIGtail	IDT	5'-GTT TCT T/AC AAC AAA TCC GTA CGC TTG-3'
UDO-043-F	PET	AB	5'-/PET/TCG GCT TTA CAA CCC ATT TC-3'
UDO-043-R	PIGtail	IDT	5'-GTT TCT T/TG CCA ATT ATG GGG CTA ACT-3'
DCA9-F	PET	AB	5'-/PET/AAT CAA AGT CTT CCT TCT CAT TTC G-3'
DCA9-R	PIGtail	IDT	5'-GTT TCT T/GA TCC TTC CAA AAG TAT AAC CTC TC-3'

Primers EMO90-F, DCA3-F, DCA14-F, and GAPU101-F were labeled with fluorescent dye (56-FAM) attached to the 5'-end of oligonucleotides from Integrated DNA Technologies (IDT) (IDT, Coralville, IA). The forward primers DCA18-F, DCA16-F, DCA5-F, and DCA17-F were attached with a green fluorescent dye (VIC) while GAPU103A-F, GAPU71B-F, UDO-043-F, and DCA9-F were attached with a fluorescent dye (PET) (both labeled groups were synthesized by Applied Biosystems (AB) (Foster City, CA). The reverse primers for all sets of 12 primer pairs were unlabeled and were obtained from Integrated DNA Technologies (IDT).

A small-tailed oligonucleotide or PIG-tail sequence (GTTTCTT) was added to all the unlabeled reverse primers to promote specific priming, full adenylation and reduce stutter bands (Brownstein *et al.*, 1996). PCR amplifications were carried out in a final volume of 10µL in 2 mL 8-strip PCR tubes with 2 µM. The solution mix for PCR reactions consisted of the following: 2.0 µL of (20 ng/µL) genomic DNA; 3 µL of (Type-it microsatellite PCR –Maste mix; QIAGEN, USA); 2.0 µL of (2.0 µM) primer mix; and 3.0 µL of deionized water. All amplifications of multiplex PCR were performed in a 96-well thermocycler (Applied Biosystems, USA) under the following conditions of touchdown annealing temperature profile (Viljoen *et al.*, 2005): 2 min at 94°C; 10 cycles of 45 sec at 94°C, 1 min at 65°C (annealing temperature was reduced 1°C after every cycle), and 1 min and 30 sec at 72°C; 35 cycles of 45 s at 94°C, 1 min at 55°C, and 1 min and 30 s at 72°C; and a final extension step of 5 min at 72°C. The touchdown procedure was used to reduce non-specific priming during PCR amplification.

After successful amplification of the target region of isolated DNA, PCR samples were combined with LIZ 600 internal size standards. Fragment analyses were performed on an Applied Biosystems 3130 xL. The fragment data were scored using 'GeneMapper' software v.3.7 to size and genotype the alleles. Once allele sizes were determined (allele calling), the data set was formatted such that it could be converted to the various formats required by the software packages (Convert program Version) (Glaubitz, 2004).

2.3 Analytical methods

2.3.1 Quality control

Quality control was performed using a set of procedures to ensure the integrity, stability and consistency of SSR results. All amplifications of PCR for each sample three times. Negative and positive standard controls were applied. Quality was evaluated prior to exporting the results of the genotype samples as matrix data. Genotypes that have the same gene fragment to minimize the error estimation of genotyping. Filtering loci set to eliminate markers that have a missing data across all genotypes.

2.3.2 Population genetic analyses

Descriptive statistics were performed using FSTAT software version 2.9.3.2 (Goudet, 2002) and GDA software version 1.1. (Observed alleles, observed fragment size, private alleles, the probability of identity and power of discrimination) were estimated for each individual locus (Table 3) (probability of identity, power of discrimination, allele richness, expected heterozygosity, observed heterozygosity and population inbreeding coefficient).

2.4 Diversity and differentiation

2.4.1 Estimation of population structure and diversity

To estimate the dissimilarity or similarity of genetic data based on their populations or type of genotype. The pairwise distance matrix of SSR data was implemented as a (.txt) input file of allelic data in DARwin software v 5.0.158 (Raman *et al.*, 2014). The constructed tree from *DARwin* software applied into the *Fig Tree* software v1.4.0 (Rambaut, 2012) to describe the relationship among olive samples using genetic distance as a tree based on (UPGMA) with the support of bootstrapped dissimilarities number of (1000) to assess the uncertainty of the tree structure.

2.4.2 Estimation of partition by assignment

Structure analysis was used to estimate genetic data to assign genotypes to specific groups without any prior information. The probability of membership into 1-4 K groups was determined by multiple runs (10 times) using STRUCTURE software Version 2.3.4 by (Pritchard *et al.*, 2003). The STRUCTURE HARVESTER program (Earl, 2012) collects results generated by STRUCTURE program. This method allows assessment and visualizes the likelihood scores of multiple values of K, to evaluate the most likely level of genetic group subdivision. The probability of identity (IP) for each locus and all SSR loci set (accumulated IP) was calculated by means of the CLUster Matching and Permutation Program (CLUMPP) version 1.1.2 (Jakobsson and Rosenberg, 2007). This program assigns individuals on the basis of optimal membership coefficients within clusters. Molecular data were combined together with morphological data of stable phenotypic traits that were blocked by results of structure assignment of molecular data to evaluate the relationship between phenotypic and genotypic data.

3. Results

A matrix of 12 SSR primers by 99 individuals (Table 1) was used to evaluate the genetic relationships among genotypes of local cultivated, introduced cultivars and wild types. As a result of filtering loci and genotypes that have missing data, allelic data of DCA17 and DCA9 were removed from the dataset due to high failure rate. Eight duplicated accessions, based on their identical genotypes, were also excluded (Table 2). Consequently, a total of 10 SSR loci and 91 genotypes (39 local, 36 introduced and 16 wild) remained in the genetic data matrix.

3.1 Identification of duplicated genotypes

Ten SSRs loci (Table 3) were used to determine if duplicate olive cultivar samples were present in the dataset. Twelve genotypes (6 pairs) had the same names and were genetically identical as true duplicates (Table 2 and Fig. 2). Two sets of cultivars had different names but identical genotypes and were therefore considered to be (synonyms) (Table 2 and Fig. 2). One cultivar from each of these eight pairs was excluded from further analyses. A review of their morphological data and associated images indicated similarity in phenotypic traits (Fig. 3).

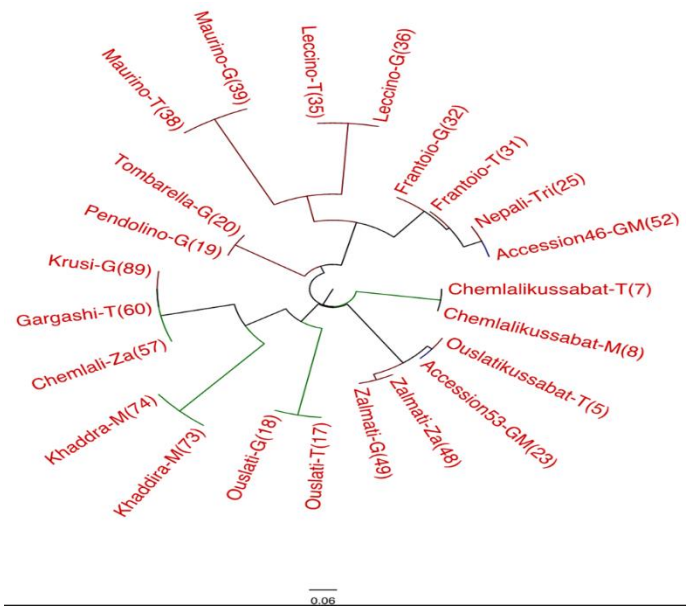

Fig. 2. A Neighbor-joining tree of 23 duplicated olive genotypes; each tip represents a single individual genotype with all pairs of duplicated genotypes similar.

Fig. 3. Phenotypic traits of the duplicated olive genotypes that illustrate similarity of genotypes.

Table 2

The same fragment sizes were considered to be duplicated or synonyms.

Variety name	Local location	Relationship	Bootstrap values (1000)
Chemlkussabat	Tharouna	Duplicated	99
Chemlkussabat	Mesalata	Duplicated	99
Khaddira	Mesalata	Synonyms	100
Khaddra	Mesalata	Synonyms	100
Ouslati	Tharouna	Duplicated	100
Ouslati	Gharian	Duplicated	100
Leccino	Tharouna	Duplicated	100
Leccino	Gharian	Duplicated	100

Variety name	Local location	Relationship	Bootstrap values (1000)
Maurino	Tharouna	Duplicated	100
Maurino	Gharian	Duplicated	100
Zalmati	Zaltin	Duplicated	59
Zalmati	Gharian	Duplicated	59
Frantoio	Tharouna	Duplicated	96
Frantoio	Gharian	Duplicated	96
Chemlali	Zaltin	Synonyms	98
Gargashi	Tharouna	Synonyms	98

3.2 Descriptive statistics of loci

A total of 109 alleles were identified, and the number of alleles per locus ranged from 4 alleles at the DCA5 locus to 20 alleles at the UDO043 locus, with an average of approximately 11 alleles per locus (Table 3). The combined discrimination power for all 10 loci was calculated with an average of (0.70) indicating that there is a moderate to high discrimination of the markers that were used, so there is a high probability that two individuals have different genotypes for each locus. The average probability of identity for all loci was low, indicating that there is a low (0.30) probability of accessions matching by chance.

Table 3

Descriptive statistics of 10 loci based on genetic data from 91 individual olive genotypes collected in Libya.

Locus	Sample size	Observed alleles (A)	Observed fragment size	Private alleles	Probability of identity (PI)	Power of discrimination (PD)
DCA14	90	10	168-188	3	0.22	0.78
DCA16	85	18	121-193	10	0.24	0.76
DCA18	92	10	154-180	3	0.20	0.80
DCA3	84	9	229-252	4	0.49	0.51
DCA5	83	4	194-206	1	0.85	0.15
EMO90	92	5	180-193	0	0.30	0.70
GAPU101	81	16	164-215	6	0.12	0.88
GAPU103A	88	11	134-189	4	0.21	0.79
GAPU71B	92	6	117-140	1	0.23	0.77
UDO043	69	20	154-227	9	0.10	0.90
All	85.6	10.9	161-198	4.1	0.30	0.70

3.3 Descriptive statistics of populations

Descriptive analysis of populations using GDA analysis (Table 4) revealed a higher inbreeding coefficient in the wild population (0.36) than the two sets of individuals, introduced (0.23) and local (0.24). The private allele frequency in the wild types was relatively higher than the other two populations. However, the discrimination power (PD) of private alleles in local and introduced genotypes was relatively high (0.99 and 0.98) respectively (Table 4). Results from the descriptive statistics of these populations provided insights into observed and expected heterozygosity. The value of expected heterozygosity (He) was higher than the value of observed heterozygosity for all three sets of individuals (Table 4) indicating there is more chance of heterozygosity at each population and they have some outbreeding resulting in disassortative mating and dissimilar traits.

Table 4

Descriptive statistics of three sets of individuals (Introduced, local and wild) collected from six locations in Libya

Sets of individuals	Sample size	Number of Private alleles	Probability of identity (PI)	Power of discrimination (PD)	Allele richness	He	Ho	Population inbreeding coefficient
Introduced	36	19	0.02	0.98	5.89	0.71	0.55	0.23
Local	39	4	0.002	0.99	4.88	0.68	0.52	0.24
Wild	16	18	0.13	0.87	5.88	0.64	0.41	0.36
Overall	30.33	13.67	0.05	0.95	5.55	0.68	0.49	0.28

z Six locations located as identified in Fig. 1.

In general, allelic richness was higher in wild and introduced genotypes (5.89 and 5.88) respectively than in local genotypes (4.88) (Table 4). There were more private alleles (observed once) in the introduced genotypes (19 private alleles), than in the wild (18 alleles) and local genotypes (4 alleles) (Table 4). Overall, all of the 41 private alleles were considered to be highly polymorphic across locations and could be used to assign individuals into a specific population based on their origins (Table 4). A total of 42 monomorphic alleles were estimated in all three different populations. These could not be used to assign any genotype to a specific population. Common alleles were most often observed in wild and introduced genotypes.

F-stats for the three sets of individuals (Introduced, local and wild) were estimated by performing a bootstrap analysis across loci to create 95% confidence intervals (Table 5). The pairwise Fst for the three sets of individuals were significantly different. Genetic differentiation of Fit, Fst and Fis was estimated by bootstrap test over all loci, and it was significant among all loci.

Table 5

Genetic differentiation as estimated by Fst with confidence intervals of 95% overall loci and three different locations.

Source	Fst	Fst confidence interval
Loci	0.025	−0.025-0.077
Sets of individuals	0.030	−0.030-0.080

3.4 Estimation of diversity and differentiation

3.4.1 Identification of mislabeled genotypes

Neighbor-joining relationships revealed that the 10 loci failed to distinguish a total of seven cultivars appeared to be similar when the molecular data were evaluated. These genotypes were Krusi, Pendolino, Tombarella, Ouslatikussabat, Accession53, Nepal and Aceession46. However, all seven cultivars had missing data for two loci (Table 6). A review of their morphological data and associated images (Fig. 4) indicated large differences in phenotypic traits across all of these cultivars.

Table 6

The seven cultivars had missing data that were considered to be mislabeled genotypes.

POP = Introduced	DCA18	DCA18	UDO043	UDO043	GAPU101	GAPU101	DCA3	DCA3	DCA5	DCA5
KrusiG	168	168	227	227	?	?	240	240	202	202
GargashiT	168	168	?	?	187	193	240	240	202	202
PendolinoG	174	174	204	204	189	203	240	240	202	202
TombarellaG	174	174	?	?	?	?	240	240	202	202
Ac#53	174	174	168	168	189	195	240	240	202	202
OuslatikussabatT	174	174	168	168	189	195	?	?	?	?
Ac#46	?	?	?	?	181	195	240	240	202	202
NepalTri	174	174	177	177	181	195	?	?	?	?
POP = Introduced	DCA14	DCA14	GAPU103A	GAPU103A	DCA16	DCA16	GAPU71B	GAPU71B	EMO90	EMO90
KrusiG	182	186	159	159	147	160	124	127	184	184
GargashiT	182	186	159	159	147	160	124	127	184	184
PendolinoG	186	186	150	150	147	147	121	127	183	189
TombarellaG	186	186	150	150	147	147	121	127	183	189
Ac#53	168	186	159	159	?	?	?	?	183	184
OuslatikussabatT	168	186	159	159	147	183	121	140	183	184
Ac#46	178	186	162	174	147	147	121	140	183	189
NepalTri	178	186	162	174	147	147	121	140	183	189

3.4.2 Identification of homonyms genotypes

There were 13 samples that had the same cultivar names but did not have matching genotypes (Table 7 & Fig. 5). This suggests that most of them (Chemlali-M, Chemlalisfax-T, Chemlalisfax-G, Coratina-T, Coratina-G, Jabbugi-T, Jabbugi-M, Mbuti-T, Mbuti-M, Mignolo-T, Mignolo-G, Moraiolo-T, Moraiolo-G, Rasli-T, Rasli-M, Zaafrani-T, Zaafrani-M, Zarrasi-M and Zarrasi-T) considered to be homonyms and were given the same names by human error, some of the labeled cultivars were misidentified because they matched other cultivars (Gargashi-T match Chemlali –Za, 53% bootstrap). Whereas other four cultivars (Hammudi-M, Hammudi-T and Marrari-M, Marrari-T) considered being close clones and were different from each other by 3 and 2 different alleles respectively. Comparisons of morphology images of each duplicate pair of genotypes showed distinct differences and supported the genetic results of polymorphism (Fig. 6). The problems associated with cultivar identification likely in landrace types than in introduced cultivars or wild type olives.

Table 7

Duplicated cultivars were considered to be mislabeled or homonyms genotypes

Variety name	Local location	Relationship
Chemlali	Masallatah	Homonyms
Chemlali	Zaltin	Homonyms
Chemlalisfax	Gharian	Homonyms
Chemlalisfax	Tharouna	Homonyms
Coratina	Gharian	Homonyms
Coratina	Tharouna	Homonyms
Gargashi	Masallatah	Homonyms
Gargashi	Tharouna	Homonyms
Hammudi	Masallatah	Homonyms
Hammudi	Tharouna	Homonyms
Jabbugi	Masallatah	Homonyms
Jabbugi	Tharouna	Homonyms
Marrari	Masallatah	Homonyms
Marrari	Tharouna	Homonyms

Variety name	Local location	Relationship
Mbuti	Masallatah	Homonyms
Mbuti	Tharouna	Homonyms
Mignolo	Gharian	Homonyms
Mignolo	Tharouna	Homonyms
Moraiolo	Gharian	Homonyms
Moraiolo	Tharouna	Homonyms
Rasli	Masallatah	Homonyms
Rasli	Tharouna	Homonyms
Zaafrani	Masallatah	Homonyms
Zaafrani	Tharouna	Homonyms
Zarrasi	Masallatah	Homonyms
Zarrasi	Tharouna	Homonyms

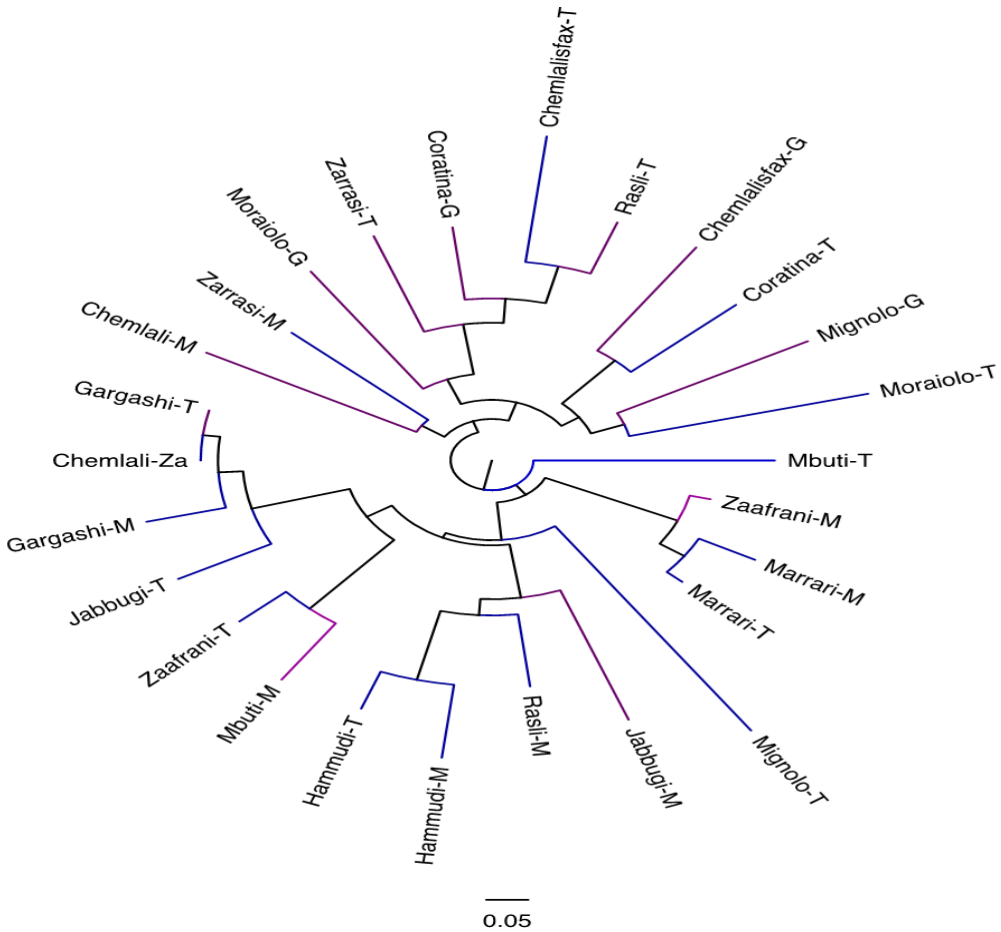

Fig. 5. Neighbor-joining tree of 13 duplicated pairs of olive genotypes; each tip represents a single individual accession with all pairs of duplicated genotypes different.

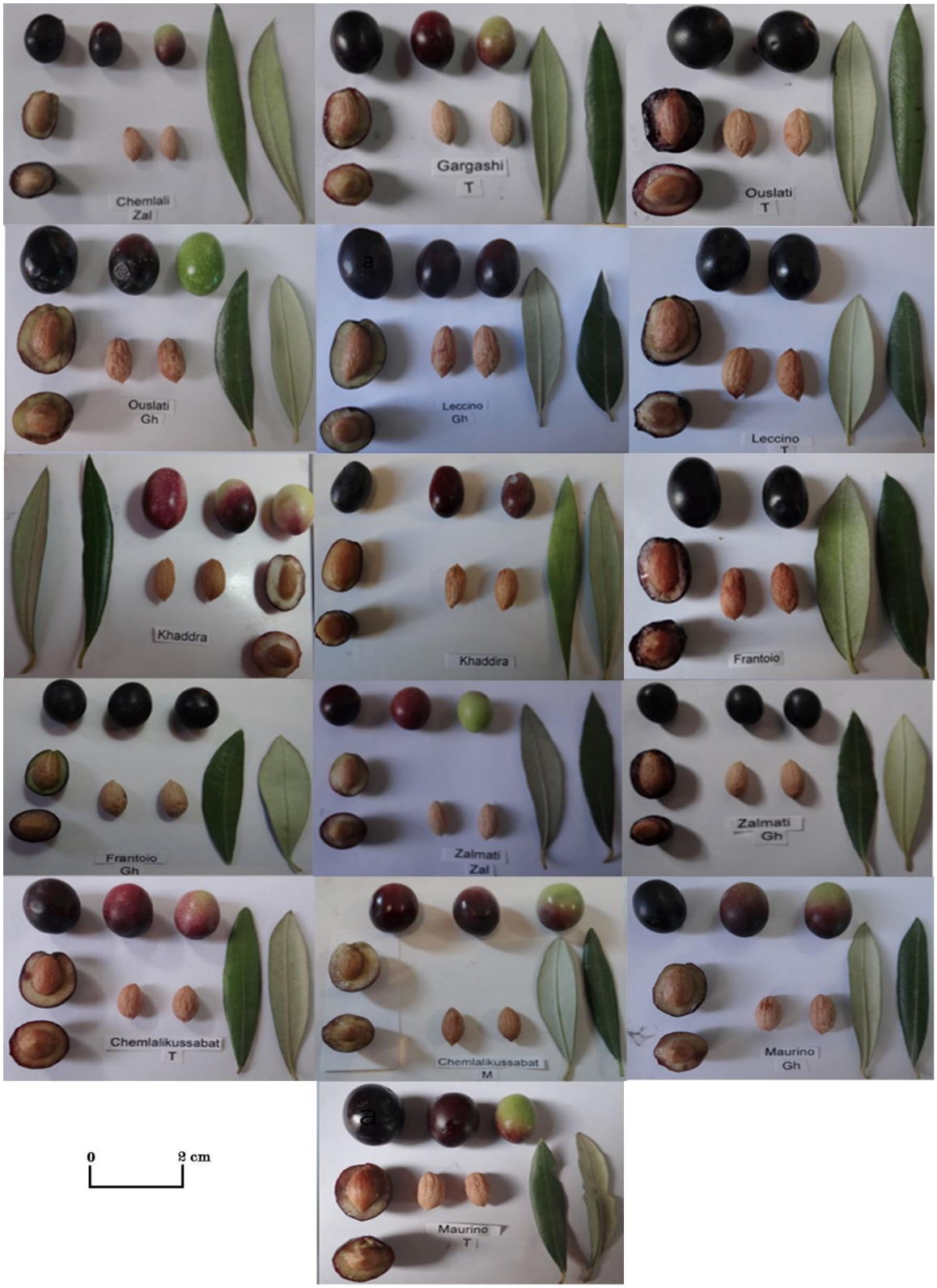

Fig. 6. Accessions identified by the same name (Homonyms accessions).

An UPGMA neighbor-joining tree (Fig. 7) was constructed to study the genetic relationships among the 91 different olive genotypes that were discriminated by the 10 SSR markers. Two primary clusters of individuals were identified (green color = landraces) and (intermixed color, red = introduced cultivars and blue = wild types). Most of the wild types were found within the intermixed wild and introduced genotypes (Fig. 7).

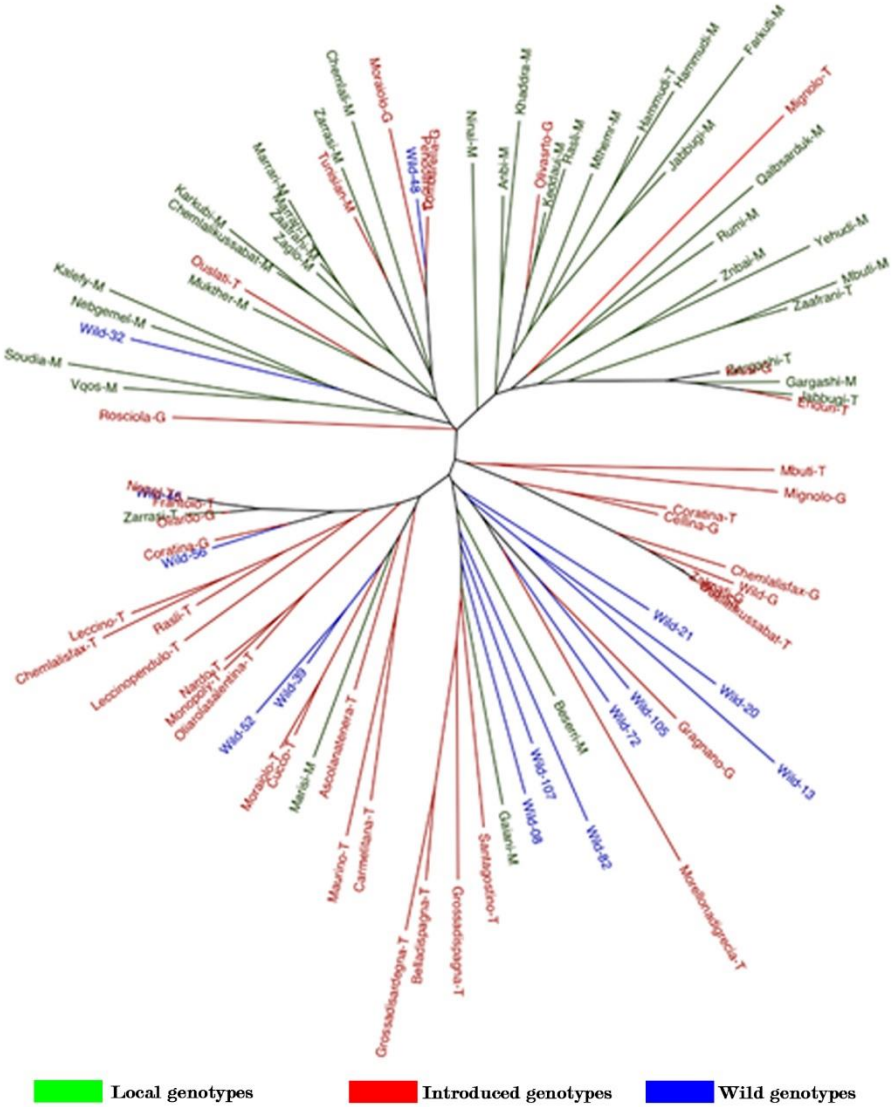

Fig. 7. Neighbor-joining tree of 91 individuals, each tip represents a single olive genotype and the colors of clades indicate the populations of origin (Local, introduced and wild).

3.5 Estimation of partition by assignment

Structure analysis using the admixture model without prior information was used to identify the genetic relationships of Libyan landraces, wild, and introduced cultivars. It was also used to differentiate individuals within each population. The most likely number of clusters inferred by structure software were at K=3. The local genotypes clustered together and two distinct sub-groups were identified. The first group consisted of the 20 most popular local genotypes (blue color) that are used mainly to produce olive oil. The second group consisted of 11 hybrid genotypes (blue and red color) between local and introduced cultivars (Fig. 8). These accessions are not widely grown and are not preferred for oil production. Those cultivars that were primarily local cultivars genetically were ancient ones grown in the Masallatah region where they are widely grown for their valuable oil characteristics. This group includes the main two cultivars Rasli and Gargashi that are used mainly for their oil production under extremely dry climates. There were six genotypes (ZarrasiM, ChemlaliM, MoraioloG, Ac#48, PendolinoG and TombarellaG) that were considered to be local genotypes in neighbor-joining tree cluster (Fig. 7) but based on the structure analysis were included in the introduced genotype grouping. This is perhaps best explained by saying that they are really introduced genotypes especially given the derivation of the names of 4 of them is not Arabic but Italian. In the case of ZarrasiM relative fruit size is similar to the introduced genotypes that have larger fruit size as compared to the smaller fruit of the local types. The wild and introduced accessions remained unchanged and were clustered the same as the UPGMA of the neighbor-joining tree (Fig. 7). They had an intermixed genetic background (red color) as shown in (Fig. 8). There were 13 genotypes that had a lot of admixture and mixed genetic background of all populations (Fig. 8). Most of these genotypes (Beserri-M, Oliarolasalentina-T, Santagostin-T, Mignolo-T, Gragnano-G, Ouslati-T, Nebgemel-M and Kalefy-M) were previously reported to be clustered as individual genotypes with Fig Tree cluster too (Fig. 7), also they have proportions of their membership in three different gene pools. Finally, the results from population structure analyses clearly distinguished the known ancient local cultivars, introduced cultivars and wild types into specific clusters associated with their origin (local, introduced and wild), but not always due to their use (oil, table and dual purpose) as reported in previous studies (Besnard *et al.*, 2001; Belaj *et al.*, 2010).

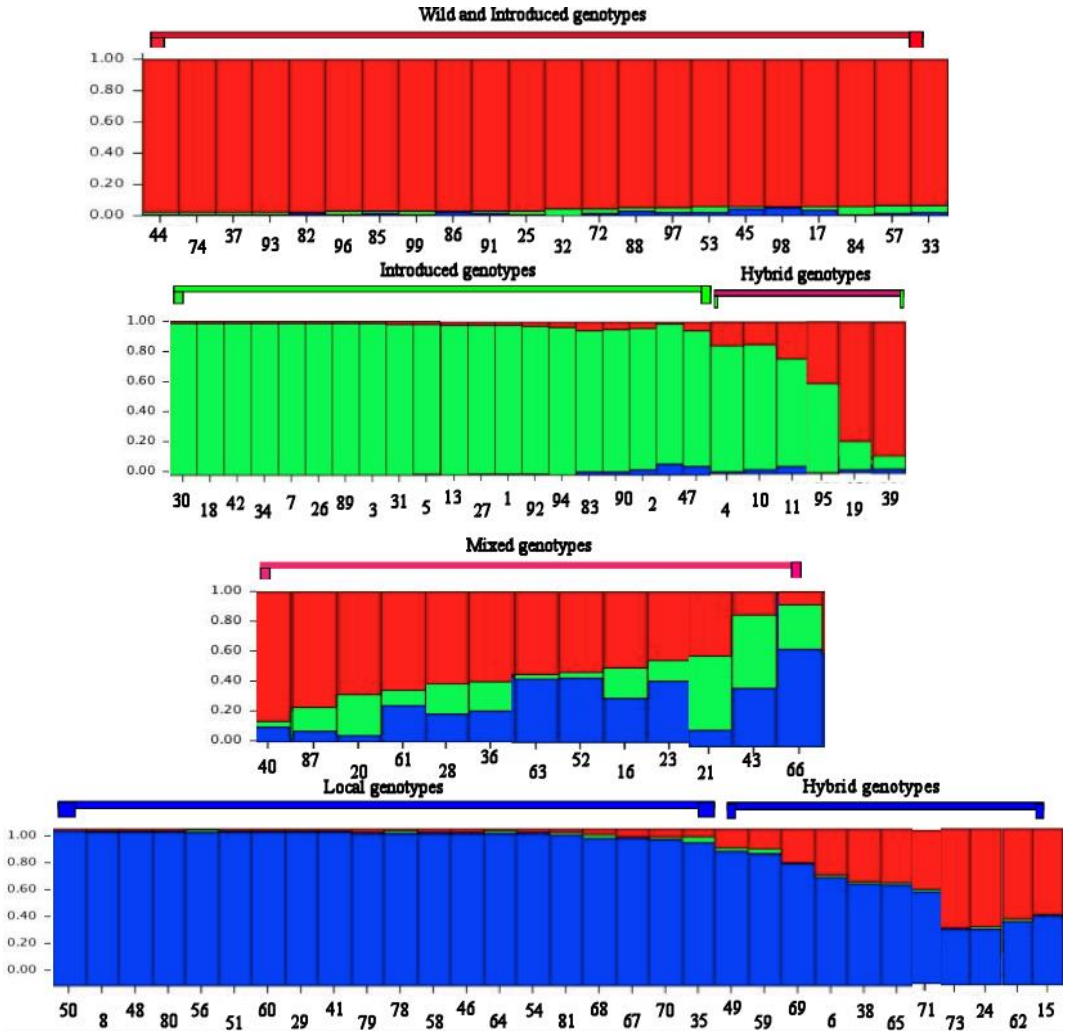

Fig. 8. Separation of the structure analysis into specific groups; intermixed group between introduce and wild genotypes (red color), Introduced genotypes (green color) and hybrid genotypes (mixed color) and local genotypes (blue color). Every single vertical strain is represented by an individual genotype.

3.6 Genotype-phenotype comparison

We sought to determine if independent stable phenotypic traits could be used to predicate the genetic classification of olive genotypes to verify if there is a strong correlation between the phenotypic and genotypic traits. Highly significant differentiation (P<0.0001***) (Fig. 9 A and Fig. 9 B) of stable phenotypic traits were observed when using the average q values of structure membership coefficient (1=local, 2=mixed and 3=introduced) or structurama partition assignment (1=mixed, 2=Introduced and 3=Landraces) respectively as a categorical data for all 90 genotypes based on the cultivar origin (introduced or local).

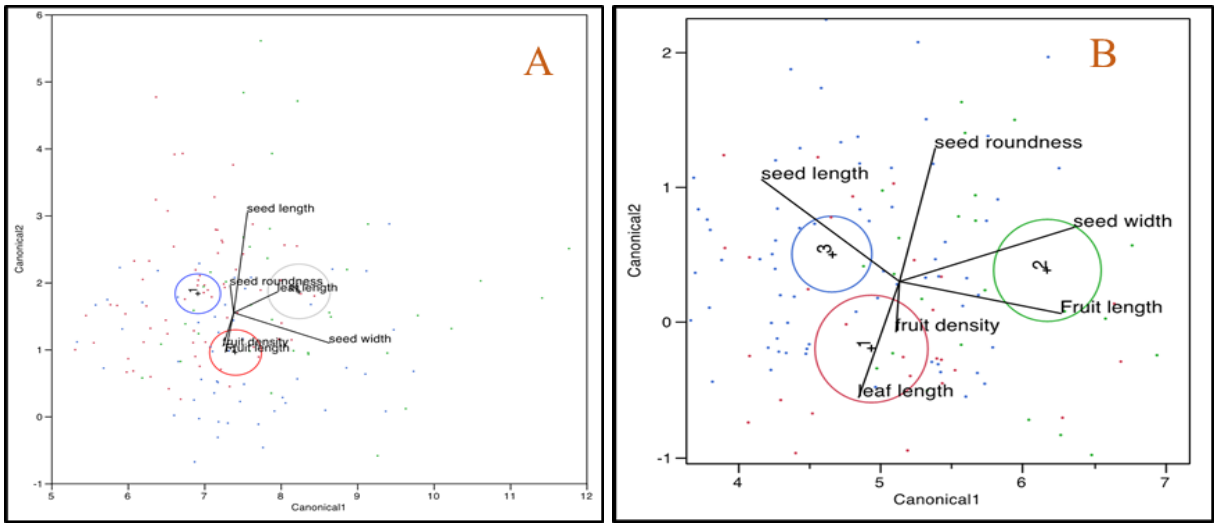

Fig. 9 Discriminant analysis was used to differentiate among all 90 genotypes based on membership q values of structure (p<0.0001***) A, and structurama assignment (p<0.0001***) B.

4. Discussion

The SSR markers (Table 1) used in this study were selected based on previously published reports (Baldoni *et al*., 2009; Erre *et al*., 2010; Dı´ez *et al*., 2011 and Ipek *et al*., 2012). The identification of duplicated, mislabeled or homonymes genotypes (Table 2, 6 and 7) respectively, found within the Libyan olive collection illustrates one of the most important problems associated with olive production in Libya. This misidentification may growers planting genotypes that are not those of yield potential in their specific area. A source of this misidentification may be due to phenotypic variation (Fig. 3) associated with environmental conditions when grown in diverse locations, to the description of the same genotype with different names. Rao *et al*. (2009) showed that synonyms and homonyms occur more frequently among landraces than in common cultivars. However, phenotypic data (fruit, seed and leaf) may be important in distinguishing different genotypes when molecular data indicates no differences due to missing or limited data. This is especially true when stable phenotype characteristics indicate differences between genotypes. Seven cultivars (Table 6) were determined to be identical based on the data from eight loci. However, this data was insufficient to discriminate all seven cultivars due to missing data of two additional loci. The combination of phenotypic traits (Fig. 4) clearly indicated that these cultivars were different.

The genetic descriptive analysis identified the most informative with a total of 20 alleles as similarly reported by Dı´ez *et al.* (2011). In general, loci that have many different alleles were preferred to distinguish between two different individuals. (*http://www.mathcs.citadel.edu*). The lowest probability of identity (PI) (0.1) was observed for locus UDO43 that was the most informative locus the highest discrimination power (0.90). The highest probability of identity (0.85) observed was for locus DCA5 that had the lowest power of discrimination (PD) (0.15) with (Table 3).

Overall, loci probability were generally low (0.30), particularly at loci that have a high allelic number as noted also in previous results (Roubos *et al*., 2010). Overall, the values observed for the expected and observed heterozygosity, for all three sets of individuals (0.68 and 0.49), respectively, were somewhat higher than reported by the authors using similar sets of SSR markers (Erre *et al*., 2010; Belaj *et al*., 2010; Muzzalupo *et al*., 2010; Baldoni *et al*., 2009; Zaher *et al*., 2011 and Erre *et al*., 2010). Reason for the number of alleles observed in the study could be due to the use of a large number of exotic genotypes.

Wild types have a higher inbreeding coefficient (0.36) than the two cultivated populations, introduced (0.23) and local (0.24). This may be the result of continued breeding of closely related individuals since the area in which the wild genotypes grow is far away from cultivated genotypes. In addition, it has the highest number of private alleles and the highest level of genetic diversity found in this area in spite of the low number of wild types. This may be useful information for the preservation traits of the wild type in the same genetic pool. The result is that the wild type may then be a source of some genes for potential improvement of local cultivars. Genetic diversity studies of the local ancient olive cultivars (Banilas *et al*., 2003 and Baldoni *et al*., 2006) have revealed that only a few of these landraces matched current olive cultivars grown today. These studies comparable to our results, which clearly indicate large differences observed in the Libyan collection.

Distinct groups of local landraces differed from introduced and wild genotypes as indicated in both the neighbor-joining tree (Fig. 7) and the admixture analysis (Fig. 8). This was also noted by (Zaher *et al*., 2011) distinct clustering of the landraces from the same region a unique genetic background and did not have matching genotypes form the other two sets of individuals. In contrast, early (Hannachi *et al*., 2010) that 'Roumi' could be a progeny of 'Chemlali', but our results from the dendrogram the major proportion of ancient Libyan local landraces did not match any other introduced or wild olive genotypes. The local Libyan cultivars may represent early stages of olive cultivation (Dı´ez *et al*., 2011 and Belaj *et al*., 2010) that remain as unexploited genetic diversity and therefore important germplasm resources Among the three sets of individuals (local, introduced and wild) that were assumed to be different not as different as expected. Neighbor-joining tree (Fig. 7) and STRUCTURE analysis (Fig. 8) demonstrated a strong between wild and introduced genotypes. The wild types were genetically more closely related to the introduced. This was unexpected since one would most commonly assume that the local cultivars were descended from the native wild types. However, samples of wild-32 and wild -48 were an exception they were phenotypically and genetically related to the landraces than the wild type. This may be due to errors of the propagation process. Therefore, the idea of Libyan ancient local cultivars maybe descendants from the wild types not supported by either neighbor-joining tree or the structural analyses. This is likely due to the result of gene flow based on geographical proximity over the years. Our results are comparable with previous studies (Hannachi *et al*., 2008 and Hannachi *et al*., 2010) that showed there are close genetic relationships between oleaster types and cultivated genotypes using SSR data with NJ method. Although some oleaster types were intermixed within cultivated genotypes, others only clustered from wild types alone.

Most of the wild type accessions were collected from the Eastern side of Libya (Fig. 1), which is closer to Europe from which introduced genotypes came to Libya in 1954 during the years of colonization by Italy. Dı´ez (2011) noted the exchange genetic material between North Africa and Europe took place during the Arab expansion through Andalusia between the eighth and fourteenth centuries. This offers the archaeological evidence to support the gene flow of olives with human migration. Wild olive genotypes are currently thought to have a common gene pool in the entire Mediterranean Basin (Kole, 2011). This may be why the wild Libyan accessions are closely related to the introduced lines from Europe. Several morphological traits can differentiate between wild and cultivated olive (Hannachi *et al*., 2008). Phenotypic traits not as informative as molecular data and limited in discriminatory power to evaluate the relatedness and the level of genetic similarity (Corrado *et al*., 2009 and Hannachi *et al*., 2008). In addition, Rao *et al.* (2009) reported that biometry values alone were unable to differentiate between similar genotypes that were evaluated by morphological traits.

It seems, there is a strong correlation of comparison between the genotype and phenotype data (Fig. 9 A and Fig. 9 B) that were based on independent phenotypic stable traits and blocked by structure membership coefficient (1=local, 2=mixed and 3=introduced) or structurama partition assignment (1=mixed, 2=introduced and 3=local) (Fig. 9 A and Fig. 9 B) respectively. The results showed that stable phenotypic data could be used the same as genetic data to assign each individual to a specific group of cultivars based on their origin (local, introduced or wild). The resemblance between molecular and morphological relationships within olive varieties expected when there is a little effect of genetic and environment interaction observed. Our results are relevant to the most recent olive. Recently, both morphological and molecular aspects have been combined to clarify the identity of genotypes within other crops (Corrado *et al*., 2009; Hannachi *et al*., 2010; Dı´ez *et al*., 2011 and Belaj *et al*., 2012).

5. Conclusion

The study of local ancient cultivars and wild types of the Libyan collection is increasingly important in order to conserve those genotypes as a potential genetic resource; they may have valuable genes that could provide a novel and useful phenotypic traits for advanced plant breeding. This study provides useful information a general molecular database of Libyan olive cultivars. There is a high heterozygosity within the Libyan collection studied, which identified all genotypes with limited similarity. The current set of 10 SSR loci amplified the corresponding microsatellite fragments in all 91 genotypes; also, it can be used to genotype the Libyan olive collection and to assign each individual into a genetic relatedness group. In this study, molecular data led to the clear identification

of 91 distinct genotypes (39 local, 36 introduced and 16 wild) out of the 99 accessions included in this study, also it revealed the existence of a high level of genetic variability among Libyan collection. It is interesting that changes of the denominations are more frequently within landraces than other cultivated and wild types. Identification of additional new candidate loci with the use of a reference sample could lead to a more robust molecular database, which could be used to characterize the Libya olive collection. This may then be used to optimize the management strategy of the Libyan olive germplasm. The combination based on morphological traits and molecular data were highly useful to separate closely related genotypes and facilitate genetic differentiation among olive genotypes.

Reference

Baldoni, L., Cultrera, N.G., Mariotti, R., Ricciolini, C., Arcioni, S., Vendramin, G.G., Buonamici, A., Porceddu, A., Sarri, V., Ojeda, M.A., Trujillo, I., Rallo, I., Belaj, A., Perri, E., Salimonti, A., Muzzalupo,I., Casagrande,A., Lain,O., Messina,R., and Testolin, R. (2009) 'A consensus list of microsatellite markers for olive genotyping ', *Mol Breeding*, 24, pp.213–231.

Baldoni, L., Tosti, N., Ricciolini, C., Belaj, A., Arcioni, S., Pannell, G., Germana, M. A., Mulas, M., and Porceddu, A. (2006) 'Genetic structure of wild and cultivated olives in the central Mediterranean Basin ', *Annals of Botany*, 98, pp.935–942.

Banilas, G., Minas, J., Gregoriou, C., Demoliou, C., Kourti, A., and Hatzopoulos, P. (2003) 'Genetic diversity among accessions of an ancient olive variety of Cyprus ', *Genome* 46, pp.370–376.

Belaj, A., del Carmen Dominguez-García, M., Atienza, S. G., Urdíroz, N. M., De la Rosa, R., Satovic, Z., and Del Río, C. (2012) 'Developing a core collection of olive (*Olea europaea* L.) based on molecular markers (DArTs, SSRs, and SNPs) and agronomic traits ', *Tree Genetics & Genomes*, 8(2), pp.365-378.

Belaj, A., Munoz-Diez, C., Baldoni, L., Satovic, Z., and Barranco, D. (2010). 'Genetic diversity and relationships of wild and cultivated olives at regional level in Spain ', *Scientia Horti culturae*, 124, pp.323–330.

Belaj, A., Satovic, Z., Cipriani, G., Baldoni, L., Testolin, R., Rallo, L., and Trujillo, I. (2003) 'Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive ', *TheorAppl Genet*. 107, pp.736–744.

Besnard, G., Baradat, P., and Berville. A. (2001) 'Genetic relationships in the olive (*Olea europaea* L.) reflect multifocal selection of cultivars ', *Theor. Appl. Genet.* , 102, pp.251–258.

Brownstein, M. J., Carpten, J. D., and Smith, J. R. (1996) 'Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping ', *Biotechniques*, 20(6), pp.1004 -6, 1008-10.

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

Cipriani, G., Marrazzo, M.T., Marconi, R., Cimato, A., and 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, pp.223–228.

Corrado, G., La Mura, M., Ambrosino, O., Pugliano, G., Varricchio, P., and Rao, R. (2009) 'Relationships of companion olive cultivars: comparative analysis of molecular and phenotypic data ', *Genome*, 52, pp.692–700.

De La Rosa, R., James, C. M., and Tobutt, K. R. (2002) 'Isolation and characterization of polymorphic microsatellites in olive (*Olea europaea* L.) and their transferability to other genera in the Oleaceae ', *Molecular Ecology Notes*, 2, pp .265–267.

D íez, C. M., Trujillo, I., Barrio, E., Belaj, A., Barranco, D., and Rallo, L. (2011) 'Centennial olive trees as a reservoir of genetic diversity. *Annals of Botany*, 108(5), pp.797-807.

Earl, D. A. (2012) 'STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method ', *Conservation Genetics Resources*, 4(2), pp.359-361.

Ercisli, S., Bencic, D., Ipek, A., Barut, E., and Liber, Z. (2012) 'Genetic relationships among olive (*Olea europaea* L.) cultivars native to Croatia and Turkey ', *Journal of Applied Botany and Food Quality*, 85(2), pp.144-149.

Ercisli, S., Ipek, A., and Barut, E. (2011) 'SSR marker-based DNA fingerprinting and cultivar identification of olives (*Olea europaea* L.) ', *Biochem Genet*. 49(9-10):555-61.

Erre, P., Chessa, I., Muñoz-Diez, C., Belaj, A., Rallo, L., and Trujillo, I. (2010) 'Genetic diversity and relationships between wild and cultivated olives (*Olea europaea* L.) in Sardinia as assessed by SSR markers ', *Genetic Resources and Crop Evolution*, 57(1), pp.41-54.

Glaubitz, J. C. (2004) 'Convert: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages ', *Molecular Ecology Notes*, 4(2), pp.309-310.

Goudet, J. (2002) 'Fstat Vision (1.2): A Computer Program to Calculate F-Statistics ', *Journal of Heredity*, 86, 485-486.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., and Petit, R. J. (2011) 'Current trends in microsatellite genotyping ', *Molecular ecology resources*, 11(4), pp.591-611.

Hannachi, H., Breton, C., Msallem, M., Ben El Hadj, S., El Gazzah,M., and Berville, A. (2010) 'Genetic relationships between cultivated and wild olive trees (*Olea europaea* L. var. *europaea* and var. *sylvestris*) based on nuclear and chloroplast SSR markers', *Natural Resources*, 1, pp. 95-103.

Hannachi, H., Breton, C., Msallem, M., Ben El Hadj, S., El Gazzah, M., and Berville, A. (2008) 'Differences between native and introduced olive cultivars as revealed by morphology of drupes, oil composition and SSR polymorphisms: A case study in Tunisia', *Scientia Horti culturae*, 116(3), pp.280–290.

http://*www.mathcs.citadel.edu/trautmand/stuff/dnapapers/little.htm*

Ipek, A., Barut, E., Gulen, H., and Ipek, M. (2012) 'Assessment of inter-and intra-cultivar variations in olive using SSR markers', *Scientia Agricola*, 69(5), pp.327-335.

Jakobsson, M., and Rosenberg, N. A. (2007) 'CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure', *Bioinformatics*, 23, pp.1801–1806.

Kole, C. (2011) 'Wild crop relatives: Genomic and breeding resources: Temperate fruits. Springer Heidelberg Dordrecht London, New York', *Vegetables*, pp.217-246.

Leon, L., Garrido, V. A., and Downey, G. (2005) 'Near infrared spectroscopy (NIRS) as a promising selection tool in olive breeding programs. *FRUTIC*, 05, pp.12-16.

Mace, E. S., Hutokshi K. Buhariwalla, H., and Crouch, J. H. (2003) 'A high-throughput DNA extraction protocol for tropical molecular breeding programs', *Plant Molecular Biology Reporter*, 21, pp.459a–459h.

Mariotti, R.; Cultrera, N. G. M.; Munoz-Diez, C.; Baldoni, L. and Rubini, A. (2010). 'Identification of new polymorphic regions and differentiation of cultivated olives (*Olea europaea* L.) through platome sequence comparison', *BMC Plant Biol*, 10, pp. 211.

Muzzalupo, I., Chiappetta, A., Benincasa, C., and Perri, E. (2010) 'Intra-cultivar variability of three major olive cultivars grown in different areas of central-southern Italy and studied using microsatellite markers', *Scientia Horticulturae*, 126(3), pp.324-329.

Pritchard, J. K., Wen, W., and Falush, D. (2003) 'Documentation for STRUCTURE software: version 2', http://*pritch.bsd.uchicago.edu*

Raman, R., Cowley, R. B., Raman, H., and Luckett, D. J. (2014) 'Analyses Using SSR and DArT Molecular Markers Reveal that Ethiopian Accessions of White Lupin (*Lupinus albus* L.) Represent a Unique Genepool', *Open Journal of Genetics*, 4, pp.87-98.

Rambaut, A. (2012) 'FigTree version 1.4. 0. Computer program distributed by the author', http://*tree. bio. ed. ac. uk/software/figtree*.

Rao, R., La Mura, M., Corrado, G., Ambrosino, O., Foroni, I., Perri, E., and 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(3), pp.261-266.

Roubos, K., Moustakas, M., and Aravanopoulos, F. A. (2010) 'Molecular identification of Greek olive (*Olea europaea*) cultivars based on microsatellite loci', *Genet. Mol. Res*, 9, pp.1865-1876.

Sarri, V., Baldoni, L., Porceddu, A., Cultrera, N.G.M., Contento, A., Frediani, M., Belaj, A., Trujillo, I., and Cionini, P.G. (2006) 'Microsatellite markers are powerful tools for discriminating among olive cultivars and assigning them to geographically defined populations', *Genome*, 49, pp.1606-1615.

Sefc, K. M., Lopes, S., Mendonca, D., Dos Santos, M. R., Machado, M. L. D., and Machado, A. D. (2000) Identification of microsatellite loci in olive (*Olea europaea*) and their characterizationin Italian and Iberian olive trees', *Mol. Ecol*. 9, pp.1171–1173.

Viljoen, G. J., Nel, L. H., and Crowther, J. R. (2005) 'Molecular diagnostic PCR handbook Dordrecht', *Springer*, pp. 13-14.

Zaher, H., Boulouha, B., Baaziz, M., Sikaoui, L., Gaboun, F., and Udupa, S. M. (2011) 'Morphological and genetic diversity in olive (*Olea europaea* subsp. *europaea* L.) clones and varieties', *Plant Omics Journal*, 4(7), pp.370-376.