UNIVERSITY OF LJUBLJANA
BIOTECHNICAL FACULTY

Aida DERVISHI

MOLECULAR CHARACTERIZATION AND ANALYSIS OF THE GENETIC VARIABILITY OF ALBANIAN OLIVES (*Olea europaea* L.)

DOCTORAL DISSERTATION

Ljubljana, 2015

UNIVERSITY OF LJUBLJANA
BIOTECHNICAL FACULTY

Aida DERVISHI

MOLECULAR CHARACTERIZATION AND ANALYSIS OF THE GENETIC VARIABILITY OF ALBANIAN OLIVES (*Olea europaea* L.)

DOCTORAL DISSERTATION

MOLEKULARNA KARAKTERIZACIJA IN ANALIZA GENETSKE RAZNOLIKOSTI OLJK (*Olea europaea* L.) NA OBMOČJU ALBANIJE

DOTORSKA DISERTACIJA

Ljubljana, 2015

The doctoral dissertation is the completion of the PhD Study of Biological and Biotechnical Sciences and relates to the field of Genetics. Experiments and analyses were performed at the laboratory of the Centre for Plant Biotechnology and Breeding, Biotechnical Faculty, Ljubljana.

On the basis of the decision of the Senate of the Biotechnical Faculty number 24. the commission of PhD studies of University of Ljubljana date 10. 09. 2009 confirmed that the PhD candidate fulfil the conditions to perform PhD studies in the PhD program of Biological and Biotechnical Sciences in the field of Genetics with thesis: Molecular characterization and analysis of the genetic variability of Albanian olives (*Olea europaea* L.). Prof. Dr. Branka Javornik was designated as the supervisor and Assoc. Prof. Dr. Nataša Štajner as co-supervisor (28.10.2014)

Commission for evaluation and defence:

President: Assoc. Prof. Dr. Jernej JAKŠE
University of Ljubljana, Biotechnical Faculty, Department of Agronomy

Member: Assist. Prof. Dr. Dunja BANDELJ
University of Primorska, Mathematics, Natural Sciences and Information Technologies Faculty, Department of Natural Sciences

Member: Assoc. Prof. Dr. Jelka ŠUŠTAR VOZLIČ
Agricultural Institute of Slovenia

I, the undersigned doctoral candidate declare that this doctoral dissertation is a result of my own research work and that the electronic and printed versions are identical. I am hereby non-paidly, non-exclusively, and spatially and timelessly unlimitedly transferring to University the right to store this authorial work in electronic version and to reproduce it, and the right to enable it publicly accessible on the web pages of Digital Library of Biotechnical faculty.

Aida Dervishi

KEY WORDS DOCUMENTATION

DN Dd
DC UDC 634.63:577.2:631.524(043.3)
CX genetic variability/*Olea europaea*/genotyping/microsatellite markers/EST-SSR/Albania
AU DERVISHI Aida
AA JAVORNIK, Branka (supervisor)/ŠTAJNER, Nataša (co-supervisor)
PP SI-1000 Ljubljana Jamnikarjeva 101
PB University of Ljubljana, Biotechnical Faculty, Postgraduate Study of Biological and Biotechnical Sciences, Field: Genetics
PY 2015
TI MOLECULAR CHARACTERIZATION AND ANALYSIS OF THE GENETIC VARIABILITY OF ALBANIAN OLIVES (*Olea europaea* L.)
DT Doctoral Dissertation
NO X, 113, [22] p., 19 tab., 15 fig., 6 ann., 186 ref.
LA En
AL en/sl
AB To study the genetic variability of Albanian olive germplasm, a set of 194 olive genotypes (130 local cultivars, 19 oleasters and 45 cultivars of foreign origin), collected in Albania, were genotyped with 26 microsatellite markers (14 g-SSR and 12 EST-SSR). The genetic diversity and relatedness among genotypes was investigated by genetic distance-based and model-based analysis. Microsatellite markers showed a high level of polymorphism, amplifying a total of 203 alleles with a mean of 7.8 alleles per locus and a mean PIC content of 0.630. A minimum number of 11 SSRs discriminated all analyzed olive genotypes. The analyses based on Dice's similarity coefficient and the UPGMA clustering method showed 12 main clusters, with an average similarity of 50.9 %, The Albanian olive genotypes clustered in one main group, separated from foreign cultivars, which was also supported by PCoA and the model-based method. A weak differentiation was observed between oleasters and cultivated olives. Weak genetic differentiation was observed based on the olive product end-use. AMOVA analysis of variation within groups defined according to their origin and breeding revealed a higher genetic variation within groups than that between groups. Several cases of synonymy, homonymy and intra-cultivar variation were observed, indicating high varietal disorder. In conclusion, microsatellite markers, especially those newly developed from ESTs, proved to be efficient in genetic diversity studies and assessment of genetic relationship among analyzed olive. A geographical structuring of olive genotypes was observed. Olive cultivars in Albania have an authocthonous origin. However, a possible exchange of plant material between Albania and neighboring regions cannot be excluded.

KLJUČNA DOKUMENTACIJSKA INFORMACIJA

ŠD Dd
DK UDK 634.63:577.2:631.524(043.3)
KG genetska variabilnost/*Olea europaea*/genotipizacija/mikrosatelitni markerji/EST-SSR/Albanija
AV DERVISHI, Aida
SA JAVORNIK, Branka (mentor)/ŠTAJNER, Nataša (somentor)
KZ SI-1000 Ljubljana Jamnikarjeva 101
ZA Univerza v Ljubljani, Biotehniška fakulteta, Podiplomski študij bioloških in biotehniških znanosti, področje genetika
LI 2015
IN MOLEKULARNA KARAKTERIZACIJA IN ANALIZA GENETSKE RAZNOLIKOSTI OLJK (*Olea europaea* L.) NA OBMOČJU ALBANIJE
TD Doktorska disertacija
OP X, 113, [22] str., 19 pregl., 15 sl., 6 pril., 186 vir.
IJ en
JI en/sl
AI Analiza genetske variabilnosti oljk s področja Albanije je bila narejena na 194-ih vzorcih, od tega 130-ih kultivarjih, 19-ih vzorcih divje oljke (*Olea europaea* L. *oleaster*) in 45-ih kultivarjih tujega izvora. Genetska analiza je bila narejena na osnovi 26 mikrosatelitnih markerjev (14 gSSR and 12 EST-SSR). Genetsko sorodnost med genotipi in delež variabilnosti smo proučili na osnovi klastrske, modelno-strukturne in PCoA analize. Z analiziranimi mikrosatelitnimi markerji smo odkrili visok nivo polimorfizma, saj smo skupno pomnožili 203 alele, v povprečju 7,8 alela na lokus pri PIC vrednosti 0,630. Vendar pa je analiza pokazala, da lahko že z 11-imi mikrosatelitnimi markerji dosežemo enako vrednost ločevanja vzorcev kot v primeru vseh analiziranih markerjev. S klastrsko analizo na osnovi koeficienta Dice in metode UPGMA smo dobili 12 ločenih skupin, v katere so se razvrstili vzorci pri povprečju podobnosti 50,9 %. Albanski genotipov so bili tako razvrščeni v eno glavno skupino, ki je ločena od tujih kultivarjev. Dobljeni rezultati so podprti tudi s PCoA im modelno-strukturno analizo. Diferenciacija med gojenimi in divjimi oljkami pa je zelo nizka in v primeru uporabne vrednosti produkta (oljčno olje ali oljke za vlaganje) vzorcev na osnovi genetske analize sploh nismo uspeli ločiti. Z AMOVA analizo posameznih skupin glede na izvor ali žlahtnjenje smo odkrili večjo raznolikost znotraj skupin kot pa med skupinami. Odkrili smo nekaj sinonimov, homonimov in znotraj-sortne variabilnosti, kar kaže na napake pri poimenovanju in identifikaciji. Mikrosatelitni markerji, posebno novejši, pridobljeni iz EST zaporedij, kažejo veliko uporabno vrednost v analizah genetske raznolikosti in odkrivanja genetske sorodnosti med kultiviranimi in divjimi oljkami. Z analizami smo odkrili tudi geografsko strukturiranje. Na osnovi dobljenih rezultatov lahko sklepamo, da imajo albanski kultivarji avtohtono poreklo, vendar pa je v preteklosti verjetno prišlo tudi do izmenjave rastlinskega materiala med Albanijo in nekaterimi sosednjimi državami.

TABLE OF CONTENTS

		p.
	Keywords documentation (KWD)	III
	Ključna dokumentacijska informacija	IV
	Table of contents	V
	Lists of tables	VII
	List of figures	VIII
	List of annexes	IX
	Abbrevations	X
1	**INTRODUCTION**	1
2	**REVIEW OF THE LITERATURE**	3
2.1	OLIVE CLASSIFICATION AND ORIGIN	3
2.2	OLIVE IN ALBANIA	4
2.3	OLIVE CHARACTERISATION METHODS	7
2.3.1	**Morphological identification**	7
2.3.2	**Biochemical methods**	8
2.3.3	**Molecular methods**	9
2.4	MICROSATELLITES DEVELOPMENT AND APPLICATION IN OLIVE	16
2.4.1	**Development of microsatellites**	16
2.4.2	**Application of microsatellite markers in *O. europaea* L.**	18
2.4.3	**Genetic parameters of microsatellite markers**	20
3	**AIM OF THE WORK**	23
4	**MATERIAL AND METHODS**	24
4.1	MATERIAL	24
4.1.1	**Plant material**	24
4.2	METHODS	27
4.2.1	**DNA isolation and quantification**	27
4.2.2	**SSR primer set**	28
4.2.3	**PCR optimisation and amplification**	31
4.2.4	**Data analysis**	33
5	**RESULTS AND DISCUSION**	36
5.1	MICROSATELLITE POLYMORPHISM AND IDENTITY ANALYSIS	36
5.2	COMPARATIVE ASSESSMENT OF GENOMIC AND EST DERIVED SSRs	45
5.3	GENETIC DIVERSITY OF OLEASTER AND CULTIVATED OLIVES	49
5.4	ANALYSIS OF IDENTITY	53
5.4.1	**Synonymy**	53
5.4.2	**Intra-cultivar variation, homonymy**	56
5.4.3	**Comparison of the Albanian olive sample set with the deposited olive genotypes**	59

5.5	PARENTAGE ANALYSIS	61
5.6	GENETIC RELATEDNESS	62
5.7	PRINCIPAL COORDINATE ANALYSIS ANDd AMOVA	73
5.8	ANALYSIS OF GENETIC STRUCTURE	81
5.9	CORE COLLECTION	84
6	**CONCLUSION**	86
7	**SUMMARY (POVZETEK)**	88
7.1	SUMMARY	88
7.2	POVZETEK	92
8	**REFERENCES**	99

ACKNOWLEDGEMENTS
ANNEXES

LIST OF TABLES

p.

Table 1: Place and origin of the olive genotypes collected in Albania 24

Table 2: Genomic primer sequences and microsatellite structures 29

Table 3: EST-SSR primer sequences and microsatellite structures 30

Table 4: Differences in PCR conditions applied for each microsatellite 32

Table 5: List of microsatellite loci and parameters of variability ranked according to their probability of identity values (PI) 40

Table 6: Genotype specific alleles 43

Table 7: Maximum number of genotypes discriminated by each of the loci combinations 44

Table 8: Mean genetic diversity parameters of different types of microsatellite loci; g-SSR, EST-SSR, di- trei- and tetra-nucleotide repeat motif microsatellites of different types of microsatellites used in the analysis 47

Table 9: Genetic diversity parameters found in different olive groups/populations: olive cultivars (current and ancient cultivars), oleasters and imported foreign olive 51

Table 10: Genetic diversity parameters found in different olive groups/populations: current, ancient olive cultivars and oleasters 52

Table 11: Groups of perfectly matching olive genotypes with the same allelic profile at 26 analysed microsatellite loci 54

Table 12: Genotypes considered near synonyms with difference ≤ 2 loci out of 26 microsatellite loci 55

Table 13: Cases of intra-cultivar variation; plants with the same name differing in 3-5 out of 26 microsatellite loci 56

Table 14: Cases of homonymy; plants with the same name differing in 6-12 out of 26 loci, which are grouped in the same cluster 57

Table 15: Cases of homonymy; plants with the same name differing in 14-24 out of 26 loci 58

Table 16: Genetic relationships of 183 olive genotypes clustered according to a UPGMA dendrogram 68

Table 17: Analysis of variance of oleasters, native and foreign cultivars 77

Table 18: Pairwaise population matrix of Nei's genetic distance between native olive cultivars, foreign cultivars and oleasters 77

Table 19: Pairwaise population matrix of Nei's genetic distance between native ancient, current cultivated olive cultivars , foreign cultivars and oleasters 78

LIST OF FIGURES

p.

Figure 1: Old olive mill of the 4th century BC and the White Olive of Tirana, with an estimated age 3000 years old

Figure 2: Olive fruit production (in tons), trend in Albania from 1961 to 2013 (FAOSTAT, 2013) 6

Figure 3: Geographical distribution of sample collection sites 25

Figure 4: Sttutering pattern reduction by applying long final elongation step at locus DCA11 37

Figure 5: Maximum number of genotypes discriminated as a function of the number of loci used 44

Figure 6: Allelic patterns across oleaster and cultivated olive genotypes and their expected heterozygosity 51

Figure 7: Dendrogram of 183 olive genotypes based on Dice coefficient and the UPGMA clustering method 67

Figure 8: Principal coordinate analysis of 183 olive genotypes based on SSR data. The first two principal coordinates explained 9.22 and 7.98 % of the variance, respectively 73

Figure 9: Principal coordinate analysis of olive genotypes based on their origin and breeding. The first two principal coordinates explained 9.22 and 7.98 % of the variance, respectively 74

Figure 10: Principal coordinate analysis of olive genotypes based on their origin (-A & -B) and breeding (-C) 76

Figure 11: PCoA plot of principal coordinate analysis of 183 native olive genotypes, defined according to their main geographical area of distribution/ origin (Adriatic vs. Ionic region). The first two principal coordinates explained 10.57 and 9.03 % of the variance, respectively 79

Figure 12: PCoA plot of principal coordinate analysis of 183 olive genotypes defined according to their fruit end-use. The first two principal coordinates explained 9.22 and 7.98 % of the variance, respectively 80

Figure 13: A- Mean Log likehood (Ln(K)+-s.d) averaged over 10 iterations for model based structure analysis (STRUCTURE); -B-Evanno test for delta K 82

Figure 14: Inference of population structure based on microsatellite data and Bayesian simulation 83

Figure 15: Genetic diversity as a function of the number of accessions included in the core collection 84

LIST OF ANNEXES

Annex A: Identification code, given name, origin and end-use of olive genotypes

Annex B: Microsatellite allelic profiles of olive genotypes for 26 microsatellites

Annex C: Allele frequencies for each of 26 microsatellite loci

Annex D: Graphic presentation of correlations between g-SSRs and EST-SSRs

Annex E: List of genotypes of *ex situ* and *in situ* core collection

Annex F Dendrogram of Albanian and olive genotypes from Slovenian collection based on Dice coefficient and the UPGMA clustering method

ABBREVATIONS

AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of Molecular Variance
CRA-OLI	CRA-Olive Growing and Oil Industry Research Centre, Italy
CTAB	Cetyl Trimethylammonium Bromide
DArT	Diversity Array Technology
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphates
EDTA	Ethylenediaminetetraacetic Acid
EST	Expressed Sequence Tag
FAOSTAT	Food and Agriculture Organization Corporate Statistical Database
Fnull	Null Allele
GenAlEx	Genetic Analysis in Excel
He	Expected Heterozygosity
Ho	Observed Heterozygosity
HWE	Hardy-Weinberg Equilibrium
IFAPA	Andalusian Institute of Agrarian and Fishing Research and Training
ISSR	Inter Simple Sequence Repeats
NaCl	Sodium Chloride
Ne	Effective Number of Alleles
No	Observed Number of Alleles
NTSYSpc	Numerical Taxonomy and Multivariate Analysis System
OWGB	Olive World Germplasm Bank, Marrakech
PCoA	Principal Coordinates Analysis
PCR	Polymerase Chain Reaction
PhiPT	Partitioning of among-group genetic variation
PI	Probability of Identity
PIC	Polymorphic Information Content
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SCAR	Sequence Characterised Amplified Region
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
TBE	Tris-Borate-EDTA
TE	Tris-EDTAbuffer
Tris	Tris (hidroksimetil) Aminometan
UPGMA	Unweight Pair Group Method with Arithmetic Average
UPOV	Union International pour Protection des Végetalés
WOGB	World Olive Germplasm Bank

1 INTRODUCTION

Olive (*Olea europaea* L.) is one of the most important fruit crops due to its worldwide use for oil production and table olives. It is one of the oldest cultivated fruit trees, with a very long history in the Mediterranean Basin. Olive is an emblematic tree crop, even in Albania, the germplasm of which has a very large genetic patrimony, represented by native cultivated olives, ancient, foreign and an abundant number of wild genotypes.

During the long history of cultivation of olives in Albania, some genetically different cultivars have been given the same name (homonymy) and, *vice versa*, some cultivars genetically identical or closely related are named differently (synonymy). Generally, generic denomination based on morphological characters, especially of the fruit, is the main cause of high varietal disorder in the country and olive collections. The existing olive germplasm thus needs to be properly identified and evaluated to ensure proper genetic resources management and preservation.

The Albanian government is currently promoting an increase of olive trees planted in the country, stressing the need for correct cultivar identification. Olive cultivars have been traditionally characterised and identified by morphologically based methods, which are known to be influenced by environmental factors and the developmental stage of the plant.

From the enormous number of molecular markers available, simple sequence repeats (SSR) are considered to be the most powerful for olive cultivar identification, genetic relationship assessment and parentage analysis, because they are co-dominant, highly polymorphic and give reproducible data that can be easily transferred among laboratories. A set of eleven microsatellites has been chosen as highly polymorphic and informative and these comprise a consensus list to be used in olive characterisation in order to provide comparable data and to construct the worldwide olive molecular database.

The present study was carried out to evaluate the genetic variability and infer the genetic structure of Albanian olive germplasm. Reference cultivars are introduced to enable data comparison with the worldwide olive genotypes deposited in the olive database and possible cases of synonymy, homonymy, intra-cultivar variability as well as planting misname or mislabeling and parentage kinship are revealed. Genetic relationships among olive genotypes and the variability partitioning among different geographical and breeding groups of olive are evaluated and visualized by dendrogram, by applying genetic distance analysis and model based clustering methods.

ESTs provide a fast and inexpensive development of SSR markers and they are especially useful for marker development because they represent coding regions of the genome and

EST derived microsatellites show high levels of transferability. They may also be linked to a particular trait, making these markers useful in marker assisted selection studies. The development of new EST-SSR markers is a valuable added tool in molecular breeding programmes of olive and cross-related species.

A new set of expressed sequence tag based microsatellite (EST-SSR) markers is introduced for the first time in this study and their efficiency in genotyping and estimation of genetic diversity of olive genotypes in comparison with genomic derived microsatellites is evaluated. The validation of microsatellite markers in genotyping of olive will enable the identification of a set of the minimum number of markers required to distinguish all olive genotypes and the proper identification of olive germplasm. The molecular data will also provide for the first time an inventory and a reference database of autochthonous olive genetic resources in Albania.

2 REVIEW OF THE LITERATURE

2.1 OLIVE CLASSIFICATION AND ORIGIN

Olea europaea L. is an evergreen tree widely grown in the Mediterranean Basin. It belongs to the medium-sized Oleaceae family, with around 600 species of 25 genera, distributed in all continents except the Antarctic (Wallander and Albert, 2000). *Olea* is one of the most economically important genera of the Oleaceae family. It includes *Olea europaea* L. species, cultivated for fruit and oil production. Based on morphological characteristics under the *Olea europaea* taxonomy, there are six known subspecies, which are also distributed in various geographical areas, such as the Mediterranean Basin (*subsp. europaea*); Africa, South Asia and the Arabian Peninsula (subsp. *cuspidate*); Sahara (subsp. *laperrini*); Morocco (subsp. *maroccana*); Madeira Islands (subsp. *cerasiformis*); and Canary Islands (subsp *guanchica*) (Green, 2002).

Botanists have reported two varieties of *Olea europaea* subsp. e*uropaea*: the cultivated form var. *sativa* and wild form var. *sylvestris* also called 'oleaster'. Both varieties are diploid species (2n = 46) (Green, 2002; Besnard *et al.*, 2008). Cultivated olive coexists in the Mediterranean countries with the oleasters and, in some cases, they are found as a complex of cultivated-wild-feral forms (Breton *et al.*, 2008). The distinction of these two forms was mainly based on morphological characteristics: on the size of the fruit, endocarp and the place in which they are found. True oleaster has been found in forest or land with apparently no relation to a cultivated area, whereas feral forms are found around current orchards or in old deserted orchads (Besnard and Bervillé, 2000).

It is assumed that olive domestication began in the Near East around 5500 to 5700 years ago, through a process of selection of wild olives that showed favourable traits. It is maintained by vegetative propagation (Zohary and Hoph, 1994). Breton *et al*. (2009), inferred the existence of nine domestication events in olive but the origins, which probably reflected the different reasons for use of domesticated oleasters, have been blurred by gene flow from oleaster and by human displacements.

Besnard and Bervillé (2000) analysed 121 olive cultivars and 300 oleasters collected from 27 populations from all around the Mediterranean Basin, using nuclear (8 RAPDs) and mitochondrial markers, in order to clarify the process of domestication of olive. The nuclear markers showed that feral oleasters cluster with the cultivars from which they are derived, whereas true oleasters cluster in distinct groups, related to those of cultivars. Mitochondrial variation showed that true oleasters and cultivars from the Near East display the same mitotype ME1 or ME2, evidencing a Near East origin of olive in oleaster. Mitotype MCK or MOM has been found in oleasters of the West. The presence of different mitotypes in oleasters of different regions of the Mediterranean reflects the complex olive

domestication, which apparently occurred in the Near East and then displaced to the West. Presumed crosses between wild and cultivated forms may have led to new cultivars around the Mediterranean countries (Besnard *et al*., 2001).

Breton *et al*. (2006) also concluded that cultivars might have a complex origin, deriving from different regions, due to gene flow occurring in oleaster, mediated by cultivars displaced by humans. Data obtained by SSR and chloroplast markers in 166 oleasters and 40 cultivars showed that oleaster genetic diversity is divided into seven regions, located in both eastern and western regions, which could coincide and overlay with glacial refuges.

Finding the location and timing of domestication have been the aim of studies by two research groups. Besnard *et al*. (2013) assessed plastid polymorphism and Baeysian molecular dating of 1263 oleasters and 534 cultivars in order to estimate the most recent common ancestor. They revealed three centers of main lineages of pre-Quaternary origin, the Near East, the Aegian area and the Strait of Gilbraltar, which are presumed to have led to the present huge olive diversity. Considering the diversity between wild and cultivated olive, as well as the geographic pattern of olive diversity, the authors suggested that the first domesticated gene pools had been northern Levant, from which olive spread with agriculture to the western Mediterranean. Newton *et al*. (2014) investigated the timing of the diffusion of varieties which were supposed to have been domesticated in the eastern Mediterranean and than introduced westwards. Based on morphometric comparison of olive stones from the Bronz Age, found in archaeological excavations in Syria with those of spontaneous and cultivated varieties of various orgins within the Mediterranean Basin, they concluded that oleasters divide into two geographical types. This division reflects not only ancient differentiation but also subsequent development of hybrids between wild and native or introduced olive cultivars. However, both wild and cultivated olives of the same geographic area may have similar morphotypes of stones because of the climate effect on this trait.

The contribution of oleaster on the evolution of cultivated olive is still a widely debated issue, which relates to the distinction between real oleasters and feral plants derived from natural dissemination of cultivars, since these two forms may show a similar appearance when grown in the same ecological sites (Baldoni and Belaj, 2009).

2.2 OLIVE IN ALBANIA

Olive (*Olea europaea* L.) is one of the most important and oldest cultivated plants in the Mediterranean Basin in general and in Albania. Olive has been grown in Albania since ancient times because the climatic and ecological conditions are favourable for olive cultivation. The cultivation of olive in Albania is thought to have begun approximately at

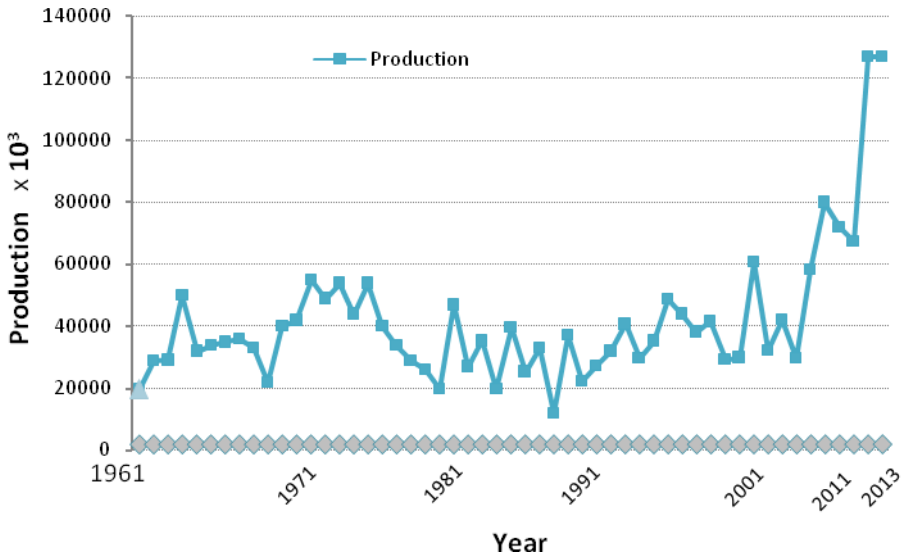
the same time as in neighbouring regions. The presence of olives in Albania has been historically demonstrated by records of olive oil trade through Apolonia through Vlora harbour in the south of Albania. There are records of olive oil exports to France by 300-150 BC (Ismaili, 2013). Finds at archaeological sites of olive stones dating to the 6th century AD (Kullaj, 2012), as well as many decorated storage vessels with olive trees are further proof of the existence of olive at that time. Olive oil production was mainly developed in olive growing regions, where many olive-related finds, such as old olive mills have been made (Meksi and Riza, 1974, *cit. by* Kullaj, 2012) in the Bylis region, which is still known for olive cultivation, which can be traced to the 13th-16th centuries.

Living proof of olive antiquity in Albania is provided by the enormous number of ancient cultivars and oleasters with an estimated age of up to 3000 years. Figure 1 shows one of the oldest olives, found in the central region of the country, as well as an ancient olive mill dating back to the 4thcentury BC.

From 1900-1945, 546 olive mills are recorded in Albania, mainly constructed in olive growing areas (Ismaili, 2013). The number of planted olive trees has increased through the years. In 1945, 1.5 million olive trees were recorded; this number increased to 4 million planted on 45 200 ha by 1990. After the 1990s, the economic and political changes had a bad impact on olive cultivation and plantations, which were greatly reduced (Kullaj, 2012). Due to the increased demand for olive oil, as well as table olives, the government initiated a major olive planting campaign in 2008, aiming at the planting up to 25 million olive trees over several years (Leoneti *et al*., 2009).The campaign is still on-going.

Olive is the most important oil-producing crop to date, planted in 8% of arable land (Kapaj A.M. and Kapaj I., 2012). There are actually 8 million olive trees planted on 26,700 ha (Veizi, *personal com.*). Orchards are mainly located in hilly zones (83 %), on plains (10 %) and in mountains (7 %). Around 72 % of olive orchards are found in the western coastal lowland, 24 % in the central region and only 3.4 % was found in the northern part of Albania. Olive fruit production in the country has increased from 17 million tons in 1961 to 125 million tons in 2013 (FAOSTAT, 2013). The olive fruit production trend in Albania from 1961-2013 is given in Figure 2.

Figure 1 Old olive mill of the 4th century BC and the White Olive of Tirana, with an estimated age 3000 years old (Ismaili *et al.*, 2013:16)
Slike 1: Star mlin za oljke in 'bela oljka' iz Tirane, katere starost je ocenjena na preko 3000 let (Ismaili in sod., 2013:16)



,Figure 2: Olive fruit production (in tons), trend in Albania from 1961 to 2013 (FAOSTAT, 2013)
Slike 2: Trend pridelave oljk v Albaniji v obdobju od leta 1961 do 2013 (FAOSTAT, 2013)

The Albanian olive sector is dominated by 7 main cultivars: Kalinjot, Kokërr Madh Berati, Kokërr Madh Elbasanit, Mixan, Ulli i Bardhë i Tiranës, Nisiot, Ulli i Hollë i Himarës, which amount to around 40 % of the total number of planted trees. The two leaders are Kalinjot and Kokërr Madh Berati, which cover 40 and 21% of total plantations, respectively (Kapaj A.M. and Kapaj I., 2012). Kalinjot has dual use, whereas Kokërr Madh Berati is used for table olives due to its large fruit size.

The most recent olive catalogue, based on morphological descriptors and oil content, describes 55 autochthonous olive genotypes, of which 45 are cultivated olives and 10 are oleasters (Ismaili, 2013).

The Albanian olive germplasm was preserved into two collections. The *ex situ* National Olive Collection was established in 1973 under the Olive Research Institute, in Shamogjin, Vlora. It corresponds to the major part of the national Albanian olive genotypes and some imported foreing olive cultivars. The *in situ* olive collection established under the Gene Bank of Albania, which consisted of ancient trees of Albanian olive genotypes with an estimated age up to 3000 years old.

2.3 OLIVE CHARACTERISATION METHODS

2.3.1 Morphological identification

Olive cultivars have been described and named since Roman times, based on various agronomic features and morphological characters, usually chosen in accordance with the researcher's selection priorities. The lack of a uniform description methodology, as well as the generic naming of cultivars, has led to increased homonymy and synonymy within a given olive growing area or between different areas (Ganino *et al.*, 2006).

The correct identification of olive cultivars is a fundamental step towards preserving the main cultivars and safeguarding minor olive genotypes, avoiding genetic erosion. Due to the known influence of climate and environmental factors on the appearance of morphological traits, the fact that the fruits are not the most efficient for cultivar differentiation, and the juvenile phase of the tree, morphological based identification led to an increase in misnaming cases (Besnard *et al.,* 2001a). In order to reduce this error rate, researchers have tended to include in their descriptive list a high number of qualitative and quantitative descriptors, as well as several agronomical features usually related to resistance to biotic and abiotic factors (Ganino *et al*., 2006).

A uniform list of 58 morphological and agronomical characters, such as plant vigour, leaves, fruit and endocarp size, inflorescence, time of fruit ripening etc., was proposed by the Union International pour Protection des Végetalés (UPOV). Information on the olive cultivar name, its given synonyms, area of origin and distribution, were added in a section of information called 'cultivar passport' together with a morphological and agronomical description in a template used to describe olive cultivars in different local or world catalogues, as well as in the world wide olive database (Barranco *et al*., 2000). The unification of morphological methodology enables a comparative evaluation of olive germplasm.

Some authors have used morphologic-based descriptors to assess genetic variability in olive (Kafazi and Muço, 1984; Trentacaste and Puertas, 2011; Ismaili *et al*., 2010, 2013; Ibtissem *et al*., 2014), while some other authors have evaluated the olive germplasm by combining morphological methodology and molecular markers, such as RAPD (Ozkaya *et al.*, 2006), AFLP (Bassi *et al*., 2002) and SSR (Hannachi *et al.*, 2008; Fendri *et al.*, 2010; Belaj *et al*., 2011; D'Imperio *et al.*, 2011; Rotondi *et al.*, 2003, 2011).

Despite the low cost of morphological analyses, molecular analysis remains a powerful tool for variability assessment within and between genotypes and for the detection of relationships related to geographical and environmental impacts on morphological traits. However, the molecular and morphological analyses are complementary tools for olive cultivar characterisation. Taken separately, the two techniques are incomplete because molecular data are not useful for cultivar identification by agronomists. In contrast, morphological data are not as useful without molecular data because of the variability due to environmental factors (D'Imperio *et al.*, 2011).

2.3.2 Biochemical methods

The biochemical approach is based on electrophoretic separation followed by specific staining of isozymes with differences in their conformation, size and charge, produced by the expression of gene/genes with nucleotide sequence alterations (Soltis and Soltis, 1989).

Isozymes were used to supplement the morphological-based descriptors due to advantages such as co-dominant nature, low cost and simple methodology. The biochemical markers applied to olive cultivar and oleaster identification were pollen or leaf proteins. Differences in the banding pattern of pollen isozymes were effective in identifying olive cultivars.

Only two out of 16 isozyme systems, malic and esterase isozymes, were effective in the identification of the majority of 27 olive cultivars collected in Greece (Pontikis *et al*., 1980). The usefulness of pollen isozymes in the detection of polymorphism among olive cultivars was confirmed by Trujillo *et al.* (1995). The different banding patterns yielded by only four isozymes, malic, phosphate isomeraze, esterase and aminopeptidase, identified 85 % of 155 analysed olive cultivars collected into 13 countries. The method was also effective for the identification of cultivars irrespective of their origin. However, isozymes were not able to detect intra-cultivar variation either in samples collected within the same or in different geographical areas.

Variation of ten leaves allozymes showed substantial differentiation of 75 oleaster populations from various habitats of Mediterranean, that were distinctly grouped into East and Western Mediterranean oleasters (Lumaret *et al.*, 2004).

The biochemical markers may reveal the polymorphism of only expressed genes, and are affected by the development stage of the plant. These limitations as well as the development of other types DNA based markers limited its application in population genetics.

2.3.3 Molecular methods

Molecular markers are DNA fragments found at specific location at the genome. They are used to identify a specific DNA sequence in a pool of unknown DNA. They are not influenced by environmental conditions and the developmental stage of the plant, which are considered to be the main advantages of their use. Several molecular markers are available and are considered useful in genetic diversity studies. Considering specific application, level of polymorphism, reproducibility, cost and the time constraints, researchers choose the molecular marker that best suits the aims of the study. Generally, molecular markers are classified as hybridisation-based markers and PCR-based markers (Kumar *et al*., 2009).

2.3.3.1 Hybridisation-based markers

2.3.3.1.1 Restriction Fragment Length Polymorphism (RFLP)

The RFPL (Restriction Fragment Length Polymorphism) approach was first introduced by Botstein *et al*. (1980) and used as the first DNA based marker in polymorphism detection. Polymorphism among individuals is expressed by the different length of fragments obtained by the digestion of DNA with a particular restriction enzyme. A mutation, insertion or deletion of one or more bases alters the cleavage site of the restriction enzyme, changing the length of the obtained restrictive fragments. These restrictive fragments, show different mobility on agarose gel, in relation to their size. They are visualized by a labelled probe indicating the existence of polymorphism.

RFLP markers are co-dominant and show high reproducibility, and they are preferentially used over morphological markers in the identification of olive origin domestication and to investigate relationships between oleasters and cultivated olive in the Mediterranean Basin. Besnard and Bervillé, (2000) thus distinguished three groups, including both oleaster and cultivated olive, by applying RFLP technology in the detection of mitochondrial polymorphism among 121 cultivars and 300 oleasters sampled in 27 populations around the Mediterranean Basin. The mitochondrial polymorphism displayed a clear–cut geographical differentiation of oleasters and cultivars from the eastern and western Mediterranean, which showed three different specific miotypes.

The RFLP approach has also been applied for the identification of phylogenetic relationships among cultivated and wild olives based on chloroplast polymorphism. In a total of 111 olive samples (66 cultivated and 45 wild), four chlorotypes were identified, defining cultivars and oleasters in three geographical areas, Western Mediterranean, Morocco and Balearic Islands (Amane *et al.*, 2000).

The RFLP approach has also been widely used to detect cytoplasmatic cpDNA polymorphism in olive by other authors (Breton *et al*., 2006; Lumaret, 2000; Amane *et al*., 1999, Besnard and Bervillé, 2002). However, some differences were observed among results obtained in polymorphism of cpDNA detected using classical RFLP based on hybridisation and the RFLP-PCR approach in *Olea europaea* complex. The latter method revealed more cpDNA polymorphism and led to more reliable results than the classical RFLP method, suggesting that the use of PCR-RFLP is both cheaper and more efficient than the classical RFPL (Besnard and Bervillé, 2002).

RFLP markers have obviously been applied efficiently in olive studies, overcoming the limitations related to morphological markers. However, they have been replaced by more sensitive and suitable PCR-based marker technologies.

2.3.3.1.2 Diversity Array Technology (DArT)

Diversity Array Technology (DArT) is a hybridisation-based DNA marker method for genetic variation assessment first introduced by Jaccoud *et al*. (2001). This marker system does not require prior knowledge of the DNA sequence and simultaneously types hundred of loci in a single assay that can be deployed in a single microarray platform. Genomic representation of individuals obtained after the amplification of restricted DNA fragments are then hybridized to the array and scored as present versus absent (Wittenberg *et al.*, 2005). DArT markers lack the ability to differentiate among homozygous and heterozygous individuals, since they are dominant. However, this technology provides a high through put, robust, low cost system that provides large genome coverage (Jaccoud *et al*., 2001).

The development of the first high throughput DArT marker set in olive was reported by Domínguez-García *et al*. (2012). They generated and tested two genomic representations using a *Pst*I/*Taq*I combination. The first was based on 87 selected olive cultivars from 16 olive growing countries while the second was from the DNA of parents of mapping populations ‘Picual’ and ‘Arbequina’. The developed DArT markers proved to be a powerful tool for olive germplasm evaluation since they permitted discrimination and provided important information on genetic relationships, clustering cultivars according to

their geographical areas of origin. Given their high through put and low cost, DArT markers also allowed the construction of genetic maps.

Atienza *et al*. (2013) confirmed that DArT markers could identify closely related varieties in the WOGB (Worldwide Olive Germplasm Bank) collection of Córdoba, synonyms as well as somatic accumulated mutations. Their transferability to wild olives was also proved, indicating a close-relationship between cultivated and wild olive gene-pools.

In combination with other marker systems, such as SSRs, SNPs and some agronomical traits, DArT technology was used to develop a core collection of the largest world olive germplasm in Córdoba, as well as to infer genetic structure among 361 olive accessions (Belaj *et al*., 2012).

2.3.3.2 PCR-based methods

Researchers have been assisted by the invention of the polymerase chain reaction method of DNA amplification in the provision of new, more advanced PCR-based marker systems to be used in detection of polymorphism. PCR (Polymerase Chain Reaction) based markers involve *in vitro* amplification of a particular DNA sequence through an enzimatic reaction enabled by the use of specific or abritrarily chosen oligonucleotidic primers (Kumar *et al.,* 2009). PCR ensures high throughput with a small amount of sample DNA.

2.3.3.2.1 Random Amplified Polymorphic DNA (RAPD)

The RAPD technique is based on PCR amplification of unknown regions of DNA by using short primers of an arbitrary nucleotide sequence of 10 bp. The amplified fragments are then separated by ethidium bromide stained agarose gel electrophoresis and visualized under UV light (Williams *et al.*, 1980). RAPD analysis of polymorphism includes insertion or deletion of small sequences of DNA, which may lead to a loss or change in the size of amplified fragments (Weising *et al*., 2005).

RAPD marker analysis has been shown to be more advantageous than RFLP analysis or isozyme characterisation. This technique was widely used in the 90s in olive genotyping because it requires a small amount of anonymous DNA, it is rapid, simple and low cost.

RAPD markers were usefully applied to distinguish olive cultivars; Bogani *et al*. (1994), described the first olive genotyping by means of five RAPDs. The low number of markers influenced their results; three of them were even reported as being polymorphic in screening of 11 olive cultivars. Promising results on the efficiency of RAPD markers for

identification of olive cultivars were reported by Fabbri *et al*. (1995). They screened 17 cultivars by 40 decameric RAPD primers, by which 47 polymorphic fragments were produced, useful for clustering olive cultivars according to their geographical origin and their fruit size.

Sans-Cortés *et al*. (2001) proved that RAPD technology was also useful in discriminating closely related cultivars. They analysed the genetic relatedness of forty Spanish olive cultivars with 18 RAPD markers and constructed an UPGMA dendrogram that clustered cultivars according to their geographical region. In contrast to previous studies (Fabbri *et al*, 1995), no clustering of cultivars based on their agronomical traits was observed. The usefulness of RAPD technology for identification and discrimination even among accessions of an ancient olive variety was reported by Banilas *et al*. (2010).

RAPD technology in combination with other markers systems was used to construct the two first linkage genetic maps of olive. In combination with RFLP, AFLP and SSR markers, a genetic map of 'Leccino x Dolce Agoia' (De la Rosa *et al.*, 2003) was constructed and, in combination with SSR and SCAR markers, that of 'Frantoio x Kalamata' (Wu *et al.*, 2004).

RAPD markers were applied for cultivar identification and to assess genetic relationships either alone or in combination with other markers on 82 Spanish olive cultivars (Belaj *et al*., 2004b), 113 olive cultivars of different origin grown in France (Besnard *et al*., 2001b), 100 accessions from a French olive collection (Khadari *et al*., 2003), 19 Albanian cultivars in relation to other Mediterranean olive cultivars (Belaj *et al*., 2003a), the main traditional and introduced olive varieties grown in Israel (Wiesman *et al.*, 1998), 201 Portugese olive accessions in combination with ISSR markers (Gemas *et al*., 2004), 84 Tunisian olives, as well as in comparison with other Mediterranean olives (Zitoun *et al*., 2008),12 wild and 10 cultivated olives trees from the Aegean region (Sesli and Yeğenoğlu, 2010), 13 Jordanian olive cultivars (Brake *et al*., 2014), 216 olive cultivars from Italy and other countries of the Mediterranean Basin in combination with SSR markers (Muzzalupo and Perri, 2009), 101 Greek olive cultivars in combination with ISSRs and SSRs (Linos *et al*., 2014) and many other olive cultivars.

Some of the shortcomings of RAPD include their dominant nature, because they are not able to detect heterozygous individuals. They lack reproducibility, which makes data obtained by means of RAPD markers unsuitable to be transferred among laboratories. Despite their limitations, in combination with other markers systems they are used efficiently in olive molecular characterization.

2.3.3.2.2 Amplified Fragment Lenghth Polymorphism (AFLP)

Amplified Fragment Length Polymorphism (AFLP), first described by Vos *et al*. (1995), is a powerful technique for DNA fingerprinting. It involves three main steps: (i) restriction (ii) selective amplification and (iii) gel analysis of the amplified fragments. Total genomic DNA is digested with a combination of two digestion enzymes, rare cutting (*Eco*RI and *Pst*I) and frequent cutting (*Mse*I and *Taq*I). Double- stranded DNA nucleotide adapters are legated to restriction fragments and serve as primer binding sites for PCR amplification. Only fragments with nucleotides that flank the restriction sites that match the selective nucleotides added to the 3' end of the primers are amplified. Polymorphisms are identified by the presence or absence of fragments following denaturing polyacrylamide gel or capillary electrophoresis (Blears *et al.,* 1998). The dominant nature and subjectivity in fragment scoring are the two main short comings of the AFLP technique However, the AFLP approach can analyse multiple loci simultaneously in the same assay, without prior knowledge of the DNA sequence, producing high number of polymorphic fragments.

The AFLP technique has been effectively applied in several studies of inter- and intra-cultivar variability evaluation (Bandelj *et al*., 2004; Sanz–Cortés *et al.*, 2008), assessment of genetic relationship among cultivated and wild olive (Angiolillo *et al*., 1999) and among olive cultivars (Owen *et al*., 2005; Ercisli *et al*., 2009; Rao *et al*., 2009). The construction of olive genetic maps has also employed a high number of AFLP combinations, since they generate a vast number of polymorphic fragments (De la Rosa *et al*., 2003; El Aabidine *et al*., 2010; Khadari *et al*., 2010).

2.3.3.2.3 Sequence Characterised Amplified Region (SCAR)

SCAR markers are specific primers generated by isolating, re-amplifying, cloning and sequencing of a RAPD or AFLP fragment. They exhibit some advantages over RAPD and AFLP, especially for genetic mapping but their development is time and labour consuming, as well as expensive (Weising *et al.,* 2005). However, the development of SCARs by using direct sequencing provides an alternative for SCAR marker development, in less time and at less cost (Hernández *et al*., 2001).They demonstrated that the developed markers were useful for olive germplasm evaluation and marker assisted selection.

Though SCARs are less polymorphic that other types of markers, such as SSRs, they have proved to be useful in olive cultivar identification (Bautista *et al*., 2002; Bucsoni *et al*., 2006) and they were used in combination with other markers (RAPDs and SSRs) in linkage map construction of the cross 'Frantoio' x 'Kalamata' (Wu *et al*., 2004).

2.3.3.2.4 Simple Sequence Repeats (SSR)

Microsatellites or simple sequence repeat (SSRs) are tandemly repeated motifs of 1-6 bases found abundantly dispersed throughout the genome. Microsatellites are mainly found in non-coding regions of the genome, where the point mutation ratio is higher than in coding ones. Microsatellites are thus specific to the species from which they are developed. This is also the main limitation of microsatellites. They need to be developed for each species (Zane *et al*., 2002). However, the transferability of microsatellites within genera of the same family has proven to be successful.

Microsatellites are co-dominant markers; they can discriminate homozygous from heterozygous individuals. They are highly informative because of the variation in the repeat length, which can easily be detected by fluorescence labelling or by capillary electrophoresis.

The results of analysis may be affected by the presence of null alleles, defined as no amplification of the true allele due to a possible mutation to the primer site and of stutter bands, which may be a result of DNA slippage giving an additional band in the microsatellite pattern (Weising *et al*., 2005).

According to Weber (1990), microsatellites can be classified into three classes (i) perfect repeat sequences, in which the tandem repeats CA alternate without interruption; (ii) imperfect repeat sequences, in which the repeat CA is interrupted by no more than three consecutive non-repeat bases and (iii) compound repeat sequences interrupted by no more than three consecutive non-repeat bases from a run of >5 uninterrupted dinucleotide or longer repeat sequences other than (CA)n (Weber, 1990).

Microsatellites are highly polymorphic and widely applied in olive studies for cultivar identification, genetic relatedness, investigating population structuring, sorting out core collections etc.

2.3.3.2.5 Inter Simple Sequence Repeats (ISSR)

ISSR markers are fragments of 100-3000 bp located between adjacent oppositely oriented microsatellites regions, which are selectively amplified by using microsatellite core sequences as primers with selective nucleotides. The amplification products are visualized by gel electrophoresis and scored as absent or present (Bracci *et al.*, 2011). ISSR are dominant markers, they are specific as microsatellites but they do not need prior sequence information for their synthesis (Gomes *et al*., 2012).

Inter Simple Sequence Repeats are applied to olive for cultivar identification and genetic diversity studies (Gemas *et al*., 2004; Essadki *et al*., 2006; Martins-Lopes *et al*., 2007), as well as the investigation of clonal genetic variability (Martins-Lopes *et al*., 2009) and linkage map construction in combination with other marker systems (Khadari *et al*., 2010). RAPD and ISSR combined for olive cultivar identification are a powerful tool, enabling a high level of accurate characterisation of olive cultivars (Brake *et al*., 2014).

2.3.3.2.6 Single Nucleotide Polymorphism (SNP)

A Single Nucleotide Polymorphism is a small genetic variation in the genome that originates with a change of a single nucleotide, including deletion, insertion or substitution. SNPs are abundant in the olive genome. Analyses of only three genes in 11 diverse cultivars, revealed that the SNP frequency was 1 to 190 and 1 to 149 in exons and introns, respectively, suggesting that only a small number of cultivars is needed for effective SNP identification (Reale *et. al*., 2006). The development of SNPs requires prior sequence information. Once the location of SNPs is identified, appropriate primers can be designed and the technique can be automated, which is an added advantage of SNP technology (Kumar *et al.*, 2009).

Reale *et al*. (2006) reported the discovery of eight SNP markers, while Hakim *et al*. (2010) reported the development of another nine new SNPs in olive. The newly developed SNPs were judged to be useful in olive genetic diversity studies; their discriminatory power was comparable with that of microsatellites. However, the combination of microsatellites and SNPs was proposed to be used in juvenile phases of the plant in oil traceability studies. The development of SNPs has recently employed transcriptome sequencing, and they are considered to be valuable markers for future genetic diversity studies (Kaya *et al*., 2013).

Comparative analys is of three marker systems, RAPDs, AFLPs, SSRs, in terms of their discriminatory power and their effectiveness in establishing genetic relationships in olive studies was reported by Belaj *et al*. (2003b). This study revealed that the co-dominant nature of SSR markers permits the detection of a high number of alleles per locus and contributes to higher levels of expected heterozygosity being achieved than would be possible with RAPDs and AFLPs. All three systems were highly efficient in detecting genetic similarities in olive, while the co-dominant nature of SSRs and their high discriminatory power makes them a marker of choice in olive studies, especially in segregation studies and genome-mapping (Belaj *et al*., 2003b).

2.4 MICROSATELLITES DEVELOPMENT AND APPLICATION IN OLIVE

2.4.1 Development of microsatellites

Microsatellites have become markers of choice for genetics analyses due to their advantages over other marker systems. One of the major limitations of the use of microsatellites is their specificity. Specific microsatellite loci need to be developed for each particular species. Traditionally, microsatellite markers have been developed in olive based on the screening of genomic or enriched genomic library. The latter technology significantly increased the yield of developed microsatellites (Zane *et al*., 2002). Microsatellites are widely used in olive characterisation, and the majority of them have been developed by screening genomic and enriched genomic libraries (Sefc *et. al*., 2000; Rallo *et al*., 2000; Carriero *et al.,* 2002; Cipriani *et al*., 2002; De la Rosa *et al*., 2002; Díaz *et al*., 2006; Sabino-Gil *et al*., 2006). Only one author (De la Rosa *et al*., 2013) has reported the development of four EST derived microsatellites in olive.

Sefc *et al*. (2000) developed the first set of microsatellite markers in olive. From the screening of a size selected genomic library of a Portuguese olive collection into three different regions for (GA/CA)n repeats, 28 primers in total were designed. Tested on 48 Italian and Iberian olive cultivars, fifteen microsatellite pairs, which were designated ssrOeUA-DCA (number), successfully genotyped the olive cultivars by amplifying from 5-15 alleles. Their heterozygosities values ranged from He=0.545-0.859; Ho=0.283-0.979. Rallo *et al*. (2000) reported the development of 13 microsatellite loci by screening a (GA)n enriched library. The amplification test on 46 cultivars of the Olive Germplasm Bank in Córdoba was successful in 5 microsatellite loci, designated IAS-oli (number). The heterozygosity ranged from 0.46-0.71.

A total of 42 microsatellite primer pairs were developed by highly enriching the genomic libraries for (GA/CT)n repeats of six olive cultivars; Maitica, Cima di Melfi, Olgliarola del Bradano, Ogliarola del Vulture, Lecciono and Kalamata (Carriero *et al*., 2002). Only 20 developed SSRs were tested on the 20 accessions of these six olive cultivars. The newly developed microsatellite markers, called GAPU (numbers) amplified on average 5.7 alleles per locus. They resulted in a useful genetic relationship assessment, differentiating the olive cultivars according to their geographic origin.

A series of 30 microsatellite markers designated UDO (numbers) were developed by Cipriani *et al*. (2002) and tested on 12 olive cultivars. The number of amplified alleles ranged from 1-5, the average heterozygosity was 0.55, while the power of discrimination was 0.60. This set of primers was useful in revealing intra-cultivar variability.
Seven polymorphic microsatellites were developed from a (GA)n and (CA)n enriched genomic library by De la Rosa *et al*. (2002). The microsatellite markers, called EMO

(numbers), were tested on 23 olive cultivars. They amplified 2-9 alleles and displayed high values of heterozygosity; He=0-0.739; Ho=0.159-0.811. The developed primers were even suitable for genotyping another five species of the Oleaceae family.

Twelve new microsatellite markers were developed by Díaz *et al*. (2006) from the enriched genomic library of Arbequina for GA, GT and ACT repeats. Of 24 microsatellites pair designed, only 12 were successfully amplified in 51 olive cultivars used in testing them, giving 68 alleles. The number of alleles ranged from 1-13, while He=0-0.83. This set of primers, designated IAS-oli (number) proved to be useful in genetic relatedness assessments; they differentiated the analysed cultivars according to their geographical origin.

Sabino-Gil *et al*. (2006) developed 19 microsatellites from the enriched genomic library of the cultivar Lezzo, of which only 12 were informative and useful for olive genotyping. They were tested on 33 olive cultivars from different Mediterranean countries, amplifying 6.75 alleles per locus, ranging from 2-14 alleles, while heterozygosity values were He=0.417-0.895; Ho=0.219-0.813.

The development of microsatellites from expressed sequence tags is becoming a preferred alternative strategy to traditional methods of development of SSR markers from genomic DNA. The vast number of expressed sequence tags in public databases facilitates and reduces the cost of new SSR development.

The development of the first eight hexa-nucleotide microsatellite primers from expressed sequence tags (EST) in olive was reported by De la Rosa *et al*. (2013). One olive breeding cultivar named Sikitita, a wild olive genotype and 12 of their progenies were used to test the amplification quality and the level of polymorphism, as well as the capacity for paternity testing. All eight EST developed markers generated reproducible and polymorphic products, displaying 2-7 alleles per locus and a high level of heterozygosities (Ho=0.38-1.00; He=0.49-0.85). Only four of the tested microsatellites, based on the displayed variability parameters on respective parents, were used in paternity testing. All four tested microsatellites were useful in revealing the paternity of breeding progenies.

ESTs provide a rapid and low cost alternative for SSR marker development. These sequences may also be linked to a specific trait, so the developed microsatellites may be successfully used in marker assisted selection studies. Additionally, EST derived microsatellites show high levels of cross-taxon transferability (Ellis and Burke, 2007).

Researchers make appropriate choices for primer selection, based on the aim of the study as well as the genetic parameters displayed by the available microsatellite markers. Two research studies examined the large number of available olive microsatellites in order to

select those most effective for olive characterisation and to select a set of consensus markers to be used world wide in olive studies (Baldoni *et al*., 2009; Doveri *et al*., 2008). Unification of the microsatellite primer set for olive genotyping would enable data comparison among laboratories, as well as the construction of a worldwide olive database. Baldoni *et al*. (2009) recommended a list of 11 markers for wide use DCA3, DCA5, DCA9, DCA14, DCA16, DCA18, GAPU71B, GAPU101, GAPU103A, UDO043, and EMO90. This consensus list of microsatellite markers and the allelic references provided in the study provide a solid platform for olive genotyping by different labs, enabling inter-lab data comparison and the construction of an SSR database of olive genotypes that would be a great help in true-to-type cultivar identification (Jakše *et al*., 2013).

2.4.2 Application of microsatellite markers in *O. europaea* L.

Microsatellite markers have found wide application in olive studies because of several important advantages, such as co-dominant inheritance, high abundance, enormous extent of allelic diversity and the ease of assessing size variation by PCR with pairs of flanking primers (Weising *et al*., 2005).

Microsatellite markers have been used in various studies in olive, such as cultivar identification, establishing relationships among cultivars, as well as among cultivars and oleasters, investigating the structure of populations, inferring the population genetic structure, sorting core collections etc.

The level of intra-cultivar variability is usually high in long lived trees that are vegetatively propagated, as olive is. Worldwide olive databases have collected information about 1250 cultivar denominations coming from 100 collections from 54 olive growing countries. Correct identification and revealing possible cases of homonymy and synonymy is important for germplasm management and preserving olive genetic resources. By applying microsatellite markers, a number of cases of homonyms, synonyms, as well as high intra-cultivar variability, have been identified in the olive germplasm of Italian cultivars (Bracci *et al*., 2009), Tunisian cultivars (Rekin *et al*., 2008; Fendri *et al*., 2010), Southern Italian cultivars (Muzzalupo *et al*., 2009b) and 70 olive cultivars from Central-Southern Italy (Muzzalupo *et al*., 2010).

The usefulness of microsatellite markers as a powerful tool for olive cultivar identification has been confirmed by several studies in various olive germplasm collections, such as identification of 19 olive cultivars from Slovenian olive collection (Bandelj *et al.*, 2002), 84 accessions of national Sfax collection of Tunisia (Fendri *et al*., 2010), 20 Tunisian cultivars (Rekik *et al*., 2008) etc.

Microsatellite markers have been applied successfully in the assessment of genetic diversity and relationships among wild olives (Belaj *et al*., 2011), among cultivated and wild olives from Spain (Belaj *et al*., 2010), Sardinia (Erre *et al*., 2010) and Tunisia (Hannachi *et al*., 2008). Genetic variation observed among wild olives by microsatellite markers was high, indicating the efficency of microsatellite markers in detecting genetic diversity and relationships among wild olive genotypes (Belaj *et al.*, 2011).

The usefulness of microsatellite markers has been confirmed as a powerful tool not only for studying variation between varieties but also for characterizing intraspecific variations among cultivated olive accessions (Muzzalupo *et al*., 2009b). Genetic relationships among cultivated olives of different regions or within native and introduced cultivated olives has been assessed by microsatellite markers, which assigned cultivars according to their area of origin. Sarri *et al*. (2006) reported that only 6 microsatellite markers correctly assigned according to their origin 75.4 % (1275) of olive cultivars from 12 olive growing Mediterranean countries (538 cultivars from Italy, 183 Spain, 88 France, 52 Greece, 45 Turkey and a few cultivars from other countries, such as Croatia, Syria, Egypt, Israel, Algeria, Tunisia and Morocco). Bracci *et al*. (2009) reported the genetic relationship assessment of 40 Mediterranean olive cultivars and 23 accessions of 16 Ligurian cultivars by 12 microsatellite markers, where by the majority of Italian cultivars were differently clustered from other Mediterranean ones. Clear structuring of variability relative to the geographic origin of cultivars was obtained by the use of 11 SSRs in 211 olive cultivars in Southern Italy (Muzzalupo *et al.*, 2009a) and by the application of 4 SSRs in 25 olive cultivars from Extremadura, Spain (Delgado-Martinrez *et al*., 2012).

Microsatellites are also preferentially used in inferring the structure of populations. Microsatellites weakly differentiated olive cultivars (51) and wild olives (107) collected in the three main Spanish olive growing regions, dividing them into four main gene pools (Belaj *et al.*, 2010). In contrast, Erre *et al*. (2010) reported assignment of wild (21) and cultivated olives (57) from Sardinia into two different clusters. Belaj *et al.* (2007) assumed that the population structure of wild olive (171 individuals) from the northeastern Mediterranean partially reflected the evolutionary history of populations. A degree of admixture was observed in all populations, making it difficult to identify a clear-cut genetic boundary between areas of southern, north-eastern Spain and islands, and mainland Italy, containing either genuine wild (genuinely wild forms present in natural areas) and feral olive germplasm (wild looking seedlings of cultivated clones or products of hybridization between oleasters and cultivars escaped cultivation).

Do Val *et al*. (2012) reported a certain extent of genetic differentiation according to their end-use, of 60 Brasilian olive cultivars, indicating a relationship between their genetic make-up and agronomical traits, such as the size of the fruit and the percentage of the oil.

Using 12 SSRs, Haouane *et al.* (2011) observed structuring into three main gene pools corresponding to geographical areas; western, eastern and central Mediterranean basin, of 561 olive genotypes of OWGB, (World Olive Germplasm Bank) of Morocco, which represented 14 olive growing countries.

The conservation of cultivated plants by the establishment of a core collection is essential for the optimal management and use of their genetic resouces (Houoane *et al.,* 2011). Two world germplasm banks, in Spain and in Marrakech, have been constructed, including a huge number of olive cultivars from all over the world. The construction of a core collection is recommended for optimizing an olive germplasm collection. Several researches have sorted core collection of large and important collections. Haouane *et al.* (2011) developed a core collection of 67 olive accessions from a total of 505 accessions from 14 Mediterranean countries planted at OWGB in Marrakech (Morocco). Only 15 % (59) of cultivars from the collection of WOGB Córdoba Spain, coming from 21 different Mediterranean countries, were necessary to capture 236 alleles displayed by the WOGB collection (Díez *et al.,* 2012). Muzzalupo *et al.* (2014) proposed an Italian olive core collection, capturing all 81 detected alleles of 489 olive cultivars of Italian germplasm CRA-OLI. Only 5 % of olive accessions were sufficient to construct the core collection. According to Belaj *et al*. (2012), a core collection containing 10-19 % of the total collection size was considered optimal to retain the bulk of genetic diversity found in the IFAPA (Andalusian Institute of Agrarian and Fishing Research and Training), a germplasm collection of 361 olive accessions.

Molecular data obtained by the use of microsatellite markers have allowed researchers to construct molecular catalogues of their regional olive germplasm, as well as to compare data between them. A worldwide olive database based on 15 microsatellite molecular data has been constructed but the amount of data deposited there is still small considering the huge number of cultivars that have been analysed by means of microsatellites.

2.4.3 Genetic parameters of microsatellite markers

The average number of alleles per locus is calculated by dividing the total number of alleles by the number of markers used in the analys is. The average number of alleles is one of the most important genetic diversity parameters. It depends on the level of polymorphism that each marker or the employed set of markers reveals and by the heterogeneity of the genotyping material.

The average allele number per locus detected in various gene pools has been reported as follows: for 489 Italian genotypes (7.6) (Muzzalupo *et al*., 2014), for 84 Tunisian accessions (8) (Fendri *et al*., 2010), for 48 cultivars from Iranian olive collections (9)

(Omrani-Sabbaghi *et al*., 2007), for 10 Turkish cultivars (4.57) (Ercisli *et al*., 2011), for 60 Brasilian accessions (6) (do Val *et al*., 2012), 27 Istrian accessions (6.75) (Poljuha *et al*., 2008), 19 cultivars from Slovenian olive collection (6.8) (Bandelj *et al*., 2004), 118 cultivars from the main Mediterranean olive-cultivating countries (13.2) (Sarri *et al*., 2006) and 104 Greek accessions (13.5) (Roubous *et al*., 2010).

The assessment of polymorphism of olive microsatellites of the expected heterozygosity of each microsatellite was calculated according to the formula of Nei (1979)

$$H_e = 1 - \Sigma(pi)^2 \quad \text{¨(1)}$$

Where He is the probability that two alleles of the same locus will be different when chosen at random; p_i the frequency of the i^{th} allele at one locus (Rallo *et al*., 2000).

The expected heterozigosity is a suitable indicator of the informative potential of a marker in cultivar identification, the higher the value of He, the higher will be the information provided by this marker (Díaz *et al*., 2006). This value also reflects the level of diversity within species (Gomes *et al.*, 2009). The higher the value of He, the broader is the genetic diversity within the analysed samples. In olive, the expected heterozygosity usually shows lower values than the observed heterozygosity, which may be due to the high selection pressure over olives towards traits of interest; i.e., in Slovenian olive collection [Ho (0.769) > He (0.705)] (Bandelj et al., 2004), Iranian cultivars [Ho (0.660) > He (0.594)] (Noormohammadi *et al.*, 2007), Istrian cultivars [Ho (0.828)> He (0.739)] (Poljuha *et al.*, 2008), Tunisian cultivars [Ho (0.77) > He (0.68)] (Fendri *et al*., 2010), Spanish olives [Ho (0.740) > He (0.698)] (Díez *et al*., 2011).

The polymorphism information content (PIC) for each marker is another parameter that provides marker evaluation. It is calculated according to the Botstein *et al.* (1980) equation

$$PIC = 1 - \Sigma p_i^2 - \Sigma\Sigma 2p_i^2 p_j^2 \quad ...(2)$$

where pi and pj are the frequencies of the *i*th and *j*th allele in the examined genotypes Loci with a high number of alleles and a PIC value near to 1 are the most informative (Botstein *et al*., 1980). According to Bandelj *et al*. (2004), microsatellites that show PIC > 0.5 are considered to be highly informative, while microsatellite markers that show PIC > 0.7 are classified as potential markers for genetic mapping. However, the PIC value is also dependant on the number of alleles detected. A comparison of PIC values of each previously described marker allows a researcher to choose a specific marker for a particular study.

In order to evaluate the efficiency of the set of SSRs to discriminate among olive cultivars, the cumulated probability of identity (PI) is calculated over the probability observed for each locus.

$$PI = \Sigma p_i^4 - \Sigma\Sigma 2(2p_i p_j)^2 \quad ...(3)$$

The usefulness of the microsatellites for cultivar identification is estimated by the power of discrimination D*j* as described by Tessier *et al*., (1999)

$$Dj = 1\text{-}Cj = \sum_{i=1}^{I} pi \frac{Npi}{N-i} \quad ...(4)$$

where the p_i is the frequence of the i^{th} pattern, N the genotypes and I is the total number of patterns produced by the j^{th} assay unit. A comparative study between two possible indices of diversity, power of discrimination and resolving power, revealed a higher suitability of power of discrimination than resolving power, to predict the ability of an SSR marker to distinguish among genotypes, suggesting that indices based on genotypic frequency are more suitable than those based on allelic frequency (Alba *et al*., 2009). The microsatellite markers UDO43 and DCA16, with higher discriminative power among 16 SSRs used to identify 30 cultivars from southern Italy, resulted in high discriminatory power and were chosen as the identification key of this set of genotypes (Alba *et al*., 2009).

3 AIM OF THE WORK

Albanian olive germplasm has a very large genetic patrimony, represented by native currently cultivated and ancient olives, foreign cultivars, and an abundant number of wild genotypes. Albanian olive germplasm evaluation has to date been done based on morphological and agronomical traits, which are known to be influenced by environmental factors. Only a limited number of cultivars have been analysed by 16 RAPD markers or by only two microsatellite loci.

Microsatellite characterization will be carried out in order to enable proper identification of olive germplasm, as the key to biodiversity preservation. Possible cases of synonymy, homonymy, intra-cultivar variability, as well as planting misnaming or mislabeling, and parentage kinship will be revealed.

Genetic relationships among olive genotypes and variability partititioning among different geographical and breeding olive groups, as well as genetic structuring of olive germplasm will be evaluated and vizualized by dendrogram and plots by applying genetic distance analysis and model based clustering methods.

In addition, the olive cultivars from Slovenian collection will be included in the analysis as references to enable comparison with microsatellite allelic profiles of world wide olive genotypes deposited in olive databases. This comparison will highlight unknown cases of synonyms or homonyms. Genetic relatedness analyses between native and foreign olive genotypes will give an insight into the origin of Albanian olive cultivars.

Molecular data will provide essential information for building the first molecular database of olive germplasm in Albania and to sort a core collection that accurately represents the entire genetic diversity in a minimum number of entities, providing proper management of the olive germplasm.

The validation of microsatellite markers in genotyping of olive genotypes will enable the identification of the minimum number of markers required to distingush all olive genotypes, which can be used in further studies or in rapid germplasm evaluation.

Assessment of the potential use of EST-derived microsatellite markers developed for the first time in this study, in genotyping and estimation of genetic diversity of olive genotypes in comparison with genomic derived microsatellites applied on the same sample set, will provide valuable information about the rational use of these categories of markers in further studies on olive.

4 MATERIAL AND METHODS

4.1 MATERIAL

4.1.1 Plant material

The Albanian *ex situ* National Olive Collection was established in 1973 under the Olive Research Institute, in Shamogjin, Vlora. The main objective of this institute is the identification and conservation of olive cultivars with better adaptability in different climates and environments, as well as those cultivars showing traits of economic value, usually linked to oil quality and production. Some introduced foreign olive cultivars have also been deposited in this collection, mainly Italian ones that have shown the best adaptation to Albanian environmental and climatic conditions. An *in situ* olive collection was recently established under the Gene Bank of Albania, with the aim of preserving autochthonous olive genetic diversity. In order to include the whole Albanian olive germplasm, plant material was sampled from both *in situ* and *ex situ* germplasm collections.

Leaf samples collected from young shoots from the upper part of each tree were stored at 4°C until DNA isolation.

Molecular analyses were performed on a set of 194 accessions of *Olea europaea* L. collected in Albania, of which 130 were local olive cultivars, 19 oleasters and 45 cultivars of foreign origin (35 Italian, 5 Spanish, 3 Greek, 1 French, 1 USA). Information about the place and genetic origin of the olive samples collected in Albania is presented in Table 1 and Figure 3.

Table 1: Place and origin of the olive genotypes collected in Albania
Preglednica 1: Mesto in izvor genotipov oljk v Albaniji

Site of sample collection	**Albanian origin**		**Foreign origin**	**Total No. of samples***
	O. europaea var. sativa	*O. europaea var. sylvestris*	*O. europaea var. sativa*	
Exsitu	66	9	45	120
Insitu	64	10	0	74
Total No.	**130**	**19**	**45**	**194**

* Nineteen reference olive cultivars are not included in the table

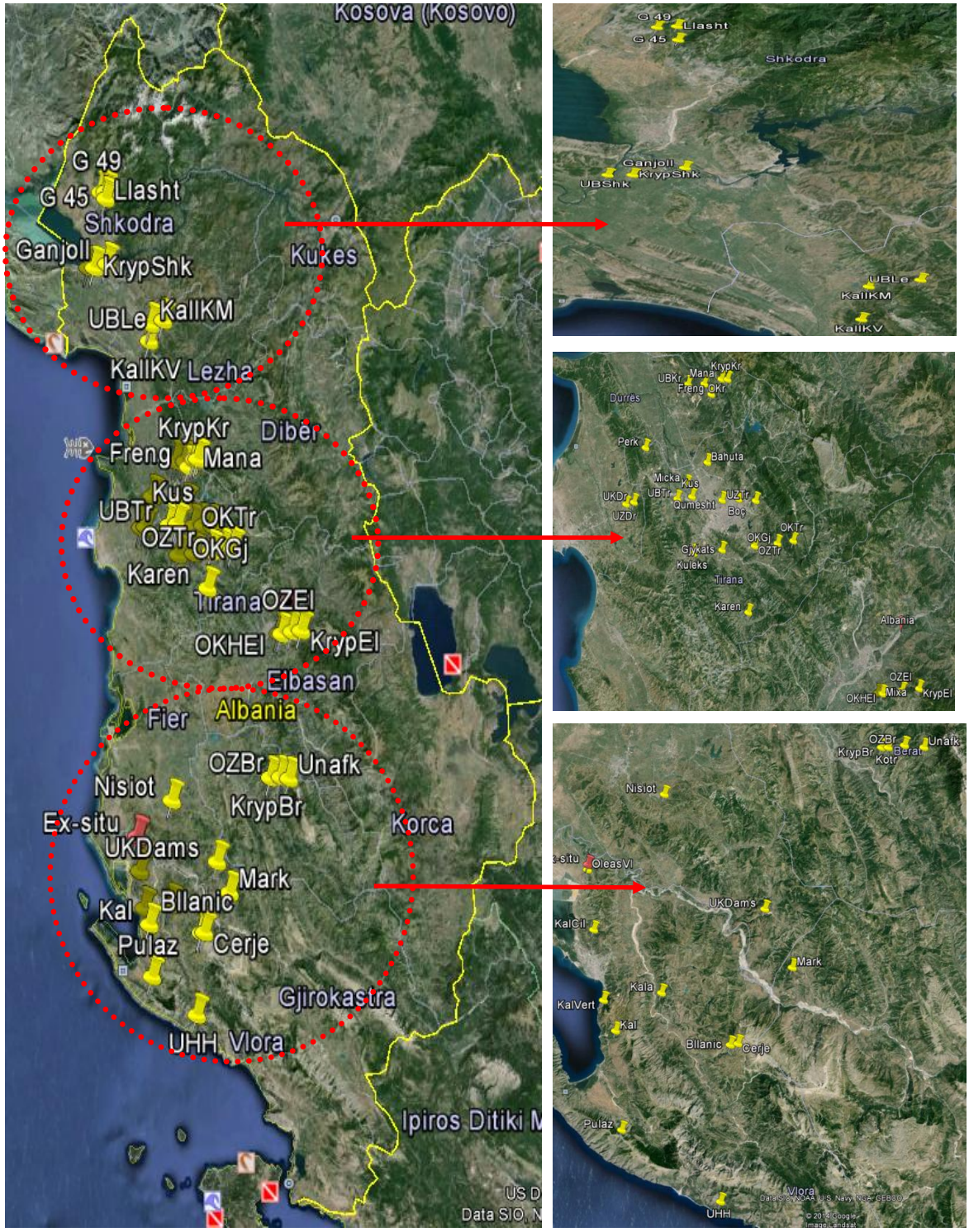


Figure 3: Geographical distribution of sample collection sites
Slike 3: Geografske prikaz mest vzorčenja

A group of 74 olive trees represent very old plants, with an estimated age of 500-3000 years, maintained *in situ* in their natural ecosystem. The ancient olive genotypes are morphologically characterized and considered to be potential olive trees to construct an Albanian *in situ* olive germplasm collection under the auspices of the Gene Bank of Albania.

Nineteen of the samples are a wild form of *Olea europaea var. sylvestris,* known also as 'oleaster'. They are used mainly as rootstock and as pollen donor for many cultivated olive cultivars (Table 1).

In addition, 19 reference cultivars were included in the analysis to enable the standardisation of amplified allele sizes and to carry out a comparison with world cultivars deposited at the Olive Germplasm Database (www.oleadb.it) and also with 306 olive genotypes of the WOGBC analysed by Trujillo *et al.* (2014) that belonged to 17 olive growing countries (171 Spanish, 35 Syrian, 22 Italian, 15 Greek, 15 Turkish, 10 Albanian, 7 Croatian, 6 French, 6 Portuguese, 3 Egyptian, 2 Tunisian, 3 Iranian, 3 Mexican, 2 USA, 1 Cyprian, 1 Lebanonese and 3 from Israel). Ten Albanian olive cultivars included in the WOGBC were previously analysed by means of RAPD markers by Belaj *et al.* (2003a).

The reference cultivars Arbequina, Ascolana Tenera, Buga, Cipressino, Črnica, Frantoio, Istarska Belica, Itrana, Leccino, Leccione, Maurino, Noccellara del Belice, Pendolino, Picholine, Santa Caterina, Štorta, Unknown 1 denominated Želenjak, Unknown 2 denominated Samo, Unknown 3 denominated Athena, had been previously studied by Bandelj *et al*. (2002) and Poljuha *et al*. (2008). The DNA of reference cultivars was provided by the Centre of Plant Biotechnology and Breeding, Biotechnical Faculty, University of Ljubljana.

The data of all sample sets, name, origin and the relevant product end-use, are presented in Annex A. Each accession is also designated according to an identification code, as a short indication of the given name, the roman number of different sample trees with the same name, and the acronym of the collection location (I=*In situ*; E=*Ex situ*; R=Repository; Ref = Reference cultivars).

4.2 METHODS

4.2.1 DNA isolation and quantification

4.2.1.1 DNA isolation

Total DNA was extracted from fresh young leaves following the CTAB method (Kump and Javornik, 1996). To extract the total DNA, 120 mg of fresh leaves were ground to powder with a pestle and mortar under liquid nitrogen. The samples were homogenized into 1000µl pre-warmed CTAB extraction buffer (2 % cetyltrimethylammonium bromide [wt/vol], 100 mM Tris-HCl pH 8.0, 20 mM EDTA pH 8.0 1.4 M NaCl and 0.2% [vol/vol] β-mercaptoethanol) and incubated at 68°C for 1-1.5 hours, mixing gently by tube inversion every 30 minutes. After incubation, 500 µl solution of phenol-chlorophorm-isoamyl alcohol (25:24:1), pH 8.0 was added and the samples were mixed well to get a suspension, and then centrifuged for 15 minutes at 11,000 rpm. The aqueous phase (500 µl) was transferred into new tubes, to which a 500 µl solution of chlorophorm-isoamyl alcohol (24:1) was added during the course of DNA purification. The samples were centrifuged at 11,000 rpm for 15 minutes, then approximately 500 µl of supernatant was removed to new micro-centrifuge tubes and extracted with 50 µl of 3 M sodium acetate pH 5.2, and 500 µl ice-cold 2-propanol. To improve the DNA precipitation, the samples were stored at -20°C for 30 minutes. The resulting DNA pellet was washed twice in the following step of centrifugation at 11,000 rpm for 15 minutes, by adding 500 µl ethanol 70 %. The final centrifugation step of samples at 11,000 rpm for 15 minutes was followed by pipetting ethanol removal, and the samples were then dried at room temperature and re-suspended in 60 µl of TE buffer (1M Tris-HCl Ph 8.0; 0.5 M EDTA pH 8.0).

The isolated DNA quality was checked by running samples on 0.8 % agarose gel in 5x TBE buffer, stained with ethidium bromide (0.5µg/ml). Samples were stored at - 20°C

4.2.1.2 DNA quantification

DNA concentration was measured based on fluorescent detection, on a DyNA Quant^{TM} 200 fluorometer (Amersham Biosciences), according to the manufacturer's instructions. In the fluorescent detection method, a single measurement establishes the DNA concentration relative to the standard DNA of a known concentration of 1000 ng/ml diluted in a working solution of 10 X TNE (0.2 N NaCl; 10 mM Tris-Cl; 1mM EDTA pH 7.4), to which a concentration of 0.1 mg/ml fluorescent dye H33258SS (Hoechst) was added.

The binding of fluorescent dye H33258 specifically to adjacent AT base pairs of DNA from outside the helix in the minor groove, shifts the maximum wavelength of the

fluorescence output of H33258, which can be measured by a DyNA Quant™ 200 fluorometer, allowing highly accurate DNA measurement.

The DNA samples were adjusted to a final concentration of 5 ng/µl and further analysed.

4.2.2 SSR primer set

The sample set of 194 olive samples was analysed with a set of 26 microsatellite markers, of which 14 were di-nucleotide microsatellite motifs originating from genomic libraries and 12 were new primers developed from Expressed Sequence Tags (EST-SSR) containing tri and tetra- nucleotide repeat motifs. The microsatellites used, their names, structure and primer sequences are listed in Table 2 and Table 3.

4.2.2.1 Genomic SSR

Fourteen genomic derived microsatellite markers used for genotyping were selected as the most effective in characterising the olive cultivars, due to the high degree of polymorphism (Polymorphic Information Content -PIC) reported in previous studies, as well as clear and repeatable amplified DNA fragments. SSR markers, ssrOeUA-DCA3, ssrOeUA-DCA5, ssrOeUA-DCA9, ssrOeUA-DCA11, ssrOeUA-DCA14, ssrOeUA-DCA15, ssrOeUA-DCA16, ssrOeUA-DCA18 were developed by Sefc *et al*. (2000), GAPU 59, GAPU89, GAPU71B and GAPU101 were identified by Carriero *et al*. (2002), while UDO99-24 and EMO90 were described by Cipriani *et al*. (2002) and by De la Rosa *et al*. (2002), respectively. Loci hereinafter are designated DCA, UDO and EMO with appropriate numbers.

Nine of these microsatellite markers (DCA3, DCA5 DCA9, DCA14, DCA16, DCA18, GAPU71B, GAPU101 and EMO90) were included in the consensus list of markers selected as the most informative and reliable by Baldoni *et al.* (2009). The use of these consensus markers will enable comparison with the results of olive genotyping from neighbouring countries, as well as introducing the allelic profiles of our olive cultivars into the olive database.

Table 2: Genomic primer sequences and microsatellite structures
Preglednica 2: Nukleotidno zaporedje začetnih oligonukleotidov in struktura mikrosatelitov

Source	Microsatellite Locus	Primer sequence (F-forward; R-reverse; 5'→3')	Repeat motif
Sefc *et al.*, (2000)	ssrOeUA-DCA03	F-CCCAAGCGGAGGTGTATATTGTTAC R-TGCTTTTGTCGTGTTTGAGATGTTG	$(GA)_{19}$
	ssrOeUA-DCA05	F-AACAAATCCCATACGAACTGCC R-CGTGTTGCTGTGAAGAAAATCG	$(GA)_{15}$
	ssrOeUA-DCA09	F-AATCAAAGTCTTCCTTCTCATTTCG R-GATCCTTCCAAAAGTATAACCTCTC	$(GA)_{23}$
	ssrOeUA-DCA11	F-GATCAAACTACTGCACGAGAGAG R-TTGTCTCAGTGAACCCTTAAACC	$(GA)_{26}(GGGA)_4$
	ssrOeUA-DCA14	F-AATTTTTTAATGCACTATAATTTTAC R-TTGAGGTCTCTATATCTCCCAGGGG	$(CA)_{18}A_6(TAA)_7$
	ssrOeUA-DCA15	F-GATCTTGTCTGTATATCCACAC R-TATACCTTTTCCATCTTGACGC	$(CA)_3G(AC)_{14}$
	ssrOeUA-DCA16	F-TTAGGTGGGATTCTGTAGATGGTGG R-TTTTAGGTGAGTTCAGAATTAGC	$(GT)_{13}(GA)_{29}$
	ssrOeUA-DCA18	F-AAGAAAGAAAAAGGCAGAATTAAGC R-GTTTTCGTCTCTCTACATAAGTGAC	$(CA)_4CT(CA)_3(GA)_{19}$
Carriero *et al.*, (2002)	GAPU59	F-CCCTGCTTTGGTCTTGCTAA R-CAAAGGTGCACTTTCTCTCG	$(CT)_9$
	GAPU89	F-GATCATTCCACACACGAGAG R-AACACATGCCCACAAACTGA	$(AG)_{16}(G)_3(GA)_9$
	GAPU71B	F-GATCAAAGGAAGAAGGGGATAAA R-ACAACAAATCCGTACGCTTG	$GA(AG)_6(AAG)_8$
	GAPU101	F-CACTGAAAGGAGGGGGACATA R-GGCACTTGTTGTGCAGATTG	$(GA)_8(G)_3(AG)_3$
Cipriani *et al.*, (2002)	UDO99-024	F-GGATTTATTAAAAGCAAAACATACAAA R-CAATAACAAATGAGCATGATAAGACA	$(CA)_{11}(TA)_2(CA)_4$
De la Rosa *et al.*, (2002)	EMO90	F-CATCCGGATTTCTTGCTTT R-AGCGAATGTAGCTTTGCATGT	$(CA)_{10}$

4.2.2.2 EST-SSR derived microsatellite primers

The EST sequences used for primer designation were obtained from the Centre of Plant Biotechnology and Breeding, Biotechnical Faculty, University of Ljubljana. Of a total of 625 olive EST sequences, those containing tri- and tetra- nucleotide repeat motif were selected to be used for the development of primers. The selected sequences had a number of repetitions of the motif ranging from 7 to 13 for tri-nucleotide motifs and from 5 to 13 for tetra-nucleotide motifs.

The primers were designed flanking the motifs with the aid of Primer3 software (Rozen and Skaletsy, 2000) applying specific selection criteria to ensure the amplification of specific and high quality PCR products. The main parameters used in primer design, were: a melting temperature of the primer (T_m) in the range of 57°C-62°C, with an optimum of 60°C; a CG content in the range of 40 % to 80 %, to avoid the chance of self-complementarity and dimer formation; a primer length between 18 and 25 nucleotides, with 20 bp as the optimum length; and PCR product size ranging from 150-250 bp.

The new EST-derived primers were synthesized by Integrated DNA Technologies. Their sequences, motif structure and expected product size are given in Table 3.

Table 3: EST-SSR primer sequences and microsatellite structures
Preglednica 3: Nukleotidno zaporedje EST-SSR začetnih oligonukleotidov in struktura mikrosatelitov

Microsatellite Locus	**Primer sequence (F-forward; R-reverse; 5'→3')**	**Repeat Motif**	**PCR product size**
SNB 03	F-CGGGGGCTGTCTTCTATTCT R-ACTGTTAATCAGGGCGACCA	(TTC)9	153
SNB 11	F-TTCCAAGCAGTCCAAATGGT R-CCCTCTTTCTTCAACAGTGG	(TTC)11	151
SNB 14	F-CCCAGAAAAAGGCATGAAGA R-TCGAAGGCTATCCCCTTATG	(TCT)10	218
SNB 19	F-GACCTCTTGTCTCATGCCAGA R-CCACAAACTAATAAGCAAAACTGAA	(AAG)9	215
SNB 20	F-GGCTCACCCTTTCAGTTTC R-CCCATGAATCAGCATCCATA	(TTC)9	176
SNB 22	F-GGATTGTGGCCGTTGTAAAG R-AATGGAAGAACTACCCGTTCC	(GAG)9	167
SiBi 03	F-GAGCTGCTGGGATAGGACTG R-CGATCCCTCACGTTCAAGTT	(TGTA)6	155
SiBi 04	F-TCCATTAAAACCTTGCAGTCAA R-TGACAATAACCCCATCAGCA	(ACAT)7	173
SiBi 05	F-GCATTTGGCACCTGAGTTTT R-CATTGATTGCGATGAGCACT	(GAAG)6	151
SiBi 07	F-CGACGGGAACGTTTACAGAT R-GTGATGGTGGTTGTGCAGTT	(ACAT)5	157
SiBi 11	F-CATTGATTGACCAGCGAATG R-GCAACGTCGAAAGATGGAAT	(AAAG)5	226
SiBi 19	F-CACAAGTCCGTGTCATTTGTAA R-TGTCGGCGTAAATTTGTGAG	(CAGG)4	218

4.2.3 PCR optimisation and amplification

4.2.3.1 Optimisation of PCR cycling profiles

The conditions for PCR reaction and cycling reported for the same primer set cannot be applied directly, since they usually depend on the chemicals and equipment used in the particular laboratory. Optimization of PCR involves testing different temperature conditions and cycling in order to get a specific, well separated and repeatable PCR product, which is very important for the accurate estimation of allele sizes.

Touchdown PCR is used because of tailing approach choosen to fluorescently label PCR product as described by Schuelke *et al*. (2000). It is used also as a strategy to increase the specificity of amplification. It runs with two steps of cycling at different annealing temperatures. The first set of cycles starts at a high annealing temperature to ensure the maximum specificity of the first primer-template hybridisation, followed by a gradual decrease in the annealing temperature to ensure specific amplification; then, in the second step, several cycles are performed at the lowest annealing temperature to provide high yields of the first product (Weising *et al.,* 2005).

The initial PCR cycling and thermal profile used for the optimisation of the PCR products of all microsatellite loci used in this study consisted of: one cycle of 94°C for 5 minutes, 5 touchdown cycles of 94°C for 30 s, 60°C with 1.0°C lower per cycle for 45 s, 72°C for 1.30 minutes, and then 30 cycles of 94°C for 30 s, 55 °C for 45 s, 72°C for 1.30 minutes. The final elongation step was performed at 72°C for 8 minutes.

Twenty loci (DCA5, DCA15, GAPU59, GAPU71B, GAPU89, GAPU101, UDO24, EMO 90, SNB and SiBi), with some minor modifications for some of them in annealing temperature or/and number of cycles, had good amplification and an easy to score pattern with this PCR profile.

The remaining six loci, DCA3, DCA9, DCA11, DCA14, DCA16, and DCA18, tend to produce slight to medium levels of stutter bands. Stutters differ in size by 2 bp and occur due to the slippage of DNA polymerase during the extension step of amplification (Weising *et al.*, 1995).

In our study, we made several attempts to reduce stuttering at these loci to provide a well-defined allelic pattern, such as:

1-decreasing the amount of DNA template from 20 ng to 5 ng, in order to proportionally decrease the amount of contaminants in the PCR reaction.

2-applying a short PCR (95°C for 5 min, 30 cycles of 94°C for 2 s, 55°C for 2 s, 72°c for 8s and final elongation step in 72°C for 8 min).

3-purifying of PCR product by reprecipitation with 1/5 volume 3M sodium acetate Ph=4.5 and isopropanol, followed by 95 % ethanol washing and redilution of purified PCR product in ddH_2O.

4-modifying the initial touchdown PCR conditions by increasing the time of the final elongation step from 72°C for 8 minutes to 65°C for 45 min.

4.2.3.2 PCR amplification

The PCR reactions were performed in 15 µl final volume containing 20 ng genomic DNA, 5x PCR buffer (Promega), 2 mM $MgCl_2$ (Promega), 0.2 mM of each of the dNTPs (Sigma), 0.2 µM of each primer, 0.25 µM fluorescently labelled M13 tail and 0.3 U Taq DNA Polymerase (Promega). The forward primer was labelled with 6-FAM, VIC, NED or PET fluorescent dye. Amplification was carried out in a GeneAmp 9600 thermal cycler (Applied Biosystem) using the temperature profiles given in Table 4 for each of the primers.

Table 4: Differences in PCR conditions applied for each microsatellite
Preglednica 4: Razlike v PCR pogojih za posamezen mikrosatelit

Microsatellite primer	**Touchdown T in C°/ Cycles / Final elongation step**
GAPU59, GAPU89, GAPU71B, GAPU101, UDO24, EMO90, DCA15	Touchdown cycles: 60°C -55°C (-1.0) / 30 cycles Final elongation: 72 °C-8 min
DCA09, DCA16, DCA18	Touchdown cycles :60°C -55°C (-1.0) / 30 cycle Final elongation: 65 °C-45 min
DCA05	Touchdown cycles: 62°C -57°C (-1.0) / 25 cycle Final elongation: 72 °C-8 min
DCA03, DCA11, DCA14	Touchdown cycles: 60°C -55°C (-1.0) /30 cycle Final elongation: 65 °C-45 min
SiBi03, SiBi04, SiBi05, SiBi07, SiBi11, SiBi19	Touchdown cycles : 58°C -53°C (-1.0)/ 30 cycle Final elongation: 72 °C-8 min
SNB03, SNB11, SNB14, SNB19, SNB20	Touchdown cycles: 60°C -55°C (-1.0) /30 cycle Final elongation: 72 °C-8 min
SNB 22	Touchdown cycles: 60°C -55°C (-1.0) / 35 cycle Final elongation: 72 °C-8 min

4.2.3.3 Electrophoretic analysis and fragment detection

The separation of amplified fragments was carried out with capillary electrophoresis using an ABI PRISM 3100 Genetic Analyser (Applied Biosystems), which relies on automated laser detection of fluorescently labelled DNA fragments.

To reduce the time and cost of PCR characterization, we used the fluorescent labelling method described by Schuelke, (2000). This method consists of the addition at the 5' end of the forward primer of a specific sequence (M13 tail) of 18 bp long (5'-TGTAAAACGACGGCCAGT-3') labelled with one of the fluorescent dyes 6-FAM, VIC, NED and PET. Different fluorescent labelling of PCR products enables analysis of four loci in the same capillary injection. Fragment scoring is based on different dyes of amplified loci, reducing the time and costs of the analysis. Four µl of four PCR products differently labelled were mixed together (4µl each). Hi-Di formamide (10.6 µl of) and size standard GeneScan-600 LIZ (0.5 µl) (Applied Biosystems) were mixed with 1 µl of the merged PCR product and then separated by capillary electrophoresis.

The sizes of alleles were determined against the internal standart GeneScan 600 LIZ applied with the samples, while allele naming was done using GeneMapper software version 4.0 (Applied Biosystems). Genotypes were manually scored to avoid errors attributed to the automated sizing, which usually increase in the case of stutters. Particular attention was paid during the size determination of fragments to potential genotyping error factors, such as allele dropout, which can lead to a decrease in sample heterozygosity, andstuttering patterns, which can hide the true allele peak. The resulting low frequency alleles occurring ≤ 5 times were double checked on the original pherogram and any genotyping errors were corrected accordingly. Genotypes showing only one allele were considered homozygous for the analysed locus.

4.2.4 Data analysis

The amplified fragments generated by analysing the sample set with 26 microsatellite markers were identified and sized using GeneMapper v. 4.0 software (Applied Biosystems). The obtained data were used as input in several data analysing softwares, according to their specific requirements.

The identification of genotypes that perfectly matched at all analysed loci, also referred to as perfect synonyms, was done using Identity 1.0 software (Wagner and Sefc, 1999). The synonyms were removed and a dataset of 183 genotypes showing unique DNA profiles was used as input with Cervus 3.0 software (Kalinowski *et al.*, 2007) to calculate the diversity parameters, such as: number of alleles per locus (No); the value of observed and

expected heterozygosity (H_o and H_e); Hardy-Weinberg equilibrium (HW); polymorphic information content (PIC); and frequency of null alleles (F_{null}) estimated through an iterative algorithm based on the diference between observed and expected heterozigosity. The probability of finding two identical genotypes (PI) was calculated both individually and for the total of combined loci by Identity 1.0 software (Wagner and Sefc , 1999). Ne (effective number of alleles) and the frequency of alleles (Freq ≤ 5) were calculated for each olive group defined by geographic origin and breeding using GenAlEx v 6.501 (Peakall and Smouse 2006; 2012).

The informativeness of the set of microsatellite markers used and their effectiveness in discriminating the olive genotypes was evaluated, based on the observed diversity indices in relation to their origin (genome-derived *vs* EST-derived microsatellites) and the length of their repeat motif (di-, tri- and tetra-nucleotidic microsatellites). The correlation among different Dice's based similarity matrixes obtained by each type of microsatellite of one marker and both markers together was performed using the the MXCOMP module of the Mantel test (Mantel, 1967), implemented on NTSYS v. 2.2 (Rohlf, 2005). The minimum number of markers required to distinguish all the observed genotypes was identified by AMaCAID script (Caroli *et al.,* 2011) designed in R program.

Pedigree analysis was performed using Identity 1.0 software (Wagner and Sefc, 1999). A maximum of two loci mistmatches was allowed for each trio on account of genotyping errors, the occurrence of null alleles and mutations.

Genetic relatedness among the olive accessions was evaluated by estimating DICE's similarity index (Dice, 1945), implemented in the statistical software NTSYS v 2.2 (Rohlf, 2005) on the data generated from all 26 microsatellite markers. Cluster analysis was performed based on the unweighted pair group method with an arithmetic mean (UPGMA) algorithm.

Principal coordinate analysis (PCoA) was performed based on a pairwase, individual-by-individual genetic distance matrix calculated for codominant data and also by intercomparison of the different olive groups/populations defined by their different breeding, geographic origin/main area of distribution (ancient, oleaster, cultivated and foreign olive genotypes; Ionic region and Adriatic region) and their product end use (table olive vs oil olive). Analysis of molecular variance (AMOVA) was carried out, to determine the relative partitioning of the total genetic variation among and within different groups of olive genotypes by using GenAlEx 6.5 (Peakall and Smouse 2006; 2012). The significance of the Φ_{PT} index was tested by 9999 permutations.

STRUCTURE version 2.3.4 software (Pritchard *et al*., 2000), with a model-based method, was applied to multilocus microsatellite data to infer the genetic structure. We used

STRUCTURE HARVESTER version 0.6.93 (Earl and von Holdt, 2012) for visualizing STRUCTURE output and the Evanno method for detecting the number of clusters of individuals (Evanno *et al*. 2005). Ten runs of STRUCTURE were performed by setting the number of clusters (K) ranging from 1 to 10. Each run consisted of a burn-in period of 200,000 steps followed by 10^6 MCMC (Monte Carlo Markov Chain) replicates, assuming an admixture model and correlated allele frequencies.

A core collection was sorted from the large database of individual molecular profiles of our study sample set, using the M strategy (Gouesnard *et al*., 2001; Escribano *et al*., 2008), applied with CoreFinder software. The algorithm for inferring the core collection was obtained heuristically.

5 RESULTS AND DISCUSION

5.1 MICROSATELLITE POLYMORPHISM AND IDENTITY ANALYSIS

Sample sets of 194 olive genotypes collected from *ex situ* and *in situ* collections in Albania were analysed with 26 microsatellite loci. The origin and location of collection of olive samples are presented in Table 1. The microsatellite allelic profiles of olive genotypes obtained in this investigation are given in Annex B.

The stutter patterns were observed in six out of 26 microsatellite loci (DCA03, DCA09, DCA11, DCA14, DCA16, and DCA18). The EST-SSR loci showed no stutter bands, due to their repeats being either tri- or tetra-nucleotidic origin.

Stutter patterns are considered to be the main cause of genotyping errors, due to the difficulty of correct interpretion and scoring. They may also impair the recognition or differentiation between homozygote and heterozygote (Arif *et al*., 2010). To minimize the effect of stutters in microsatellite genotyping, investigators usually try to optimize PCR conditions. Sefc e*t al.* (2000) reported a reduction of stutter bands by applying a final 10 minutes of incubation, allowing unambiguous allele designation. Omrani-Sabbaghi *et al.* (2007) reported that the application of stringent conditions using touchdown PCR minimized the stutter bands. The stutter profile observed by using sequencers was also greatly reduced in metaphore agarose.

In this work we used several protocols to reduce stutter bands. The first three attemps of decreasing amount of DNA template, applying short PCR strategy and of PCR product purification failed to improve the stuttering pattern. The application of a longer final elongation step of 45 minutes succeeded in reducing the stutter bands in the six loci, making the pattern clearer and therefore reducing the genotyping error rate, especially when the two true alleles differ by 2 bp (Figure 4).

The presence of stutter patterns of microsatellite loci in olive genotyping was also observed by some other authors (Sefc *et al*., 2000; Belaj *et al.*, 2004a; Omrani- Sabbaghi *et al*., 2007; Doveri *et al.*, 2008; Baldoni *et al.*, 2009).

*First electropherogram- amplification using touch down PCR, final elongation step 72°C for 8 minutes
*Second electropherogram- amplification using touch down PCR, final elongation step 65°C for 45 minutes

Figure 4: Sttutering pattern reduction by applying long final elongation step at locus DCA11
Slika 4: Zmanjšanje števila senčnih fragmentov na osnovi podaljšane faze PCR izdolževanja pri lokusu DCA11

Identity analysis revealed 183 unique genotypes from the 194 olive accessions analysed. The 183 unique microsatellite allelic profiles were further used to calculate several parameters of genetic diversity. The degree of polymorphism revealed by means of 26 microsatellite loci was evaluated based on parameters of variability that are presented in Table 5.

A total of 203 alleles were amplified across 183 unique genotypes by the 26 microsatellite markers (14 genomic and 12 EST-derived SSR). The range of allele sizes in base pairs varied from 136 bp at locus GAPU71B to 284 bp at locus DCA15. The number of alleles per locus ranged from 2 (SiBi 03 and SiBi 11) to 19 (DCA09), with an average of 7.8 alleles, revealing a high level of variability in our sample set.

The average allele number per locus detected in our analysis (7.8) was in the same range as that observed in 489 Italian olive varieties (7.6) (Muzzalupo *et al*., 2014), 84 Tunisian accessions (8) (Fendri *et al*., 2010), 30 cultivars from Southern Italy (8.8) (Alba *et al*., 2009), 48 cultivars from Iranian olive collection (9) (Ormani-Sabbaghi *et al*., 2007), 108 accessions from the Davis collection, USA (9.93) (Koehmstedt *et al.*, 2011) but higher than that reported for 10 Turkish cultivars (4.57) (Ercisli *et al*., 2011), 39 Italian accessions of Apulian germplasm, (5.6) (Muzzalupo *et al*. 2006), 211 Italian olive cultivars from Southern Italy (6.82) (Muzzalupo *et al.*, 2009a), 10 Iranian cultivars (5.6) (Noormohammadi *et al*., 2007), 60 Brasilian accessions (6) (do Val *et al*., 2012), 27 accession from Istria (6.75) (Poljuha *et al*., 2008), 19 cultivars from Slovenian olive

collection (6.8) (Bandelj *et al*., 2004) and 33 Tunisian accessions (7) (Abdelhamid *et al*., 2013).

The obtained average allele number per locus was lower than that reported for 142 Italian cultivars from Emilia-Romagna (10.2) (Beghé *et al.,* 2011), 73 olive trees, including wild, cultivated and ancient trees of Sardinia, Italy (10.46) (Erre *et al*., 2010), 46 Portugese cultivars (11) (Gomes *et al*., 2009), 77 olive cultivars from the two WOGB collections, Spain and CRA-OLI (12.2) (Baldoni *et al*., 2009), 505 accessions derived from 14 olive growing countries in the OWGB germplasm of Marrakech (12.5) (Haouane *et al*., 2011), 118 cultivars from the main Mediterranean olive-cultivating countries (13.2) (Sarri *et al*., 2006), 30 wild and 104 cultivated ancient olive trees of the Andalusian region, Spain (13.64) (Díez *et al*., 2011), 104 Greek accessions (13.5) (Roubous *et al*., 2010) and 158 samples of wild and cultivated olives from three olive growing areas in Spain (16.5) (Belaj *et al*., 2010). The high number of alleles obtained in the above mentioned studies may be due to the use of a large amount of highly diversified plant material (Belaj *et al*., 2007; Baldoni *et al*., 2009), as well as the high number of samples employed in the analysis. The lower mean number of alleles observed in our olive genotypes could be explained by the low degree of polymorphism revealed by some of the EST-SSR used, which is lower than the degree of polymorphism detected by genomic SSR used in previously reported.

The observed heterosygosity value (H_o), which represents the number of heterozygous individuals per locus, showed the lowest value at locus Sibi 03 (0.357) and the highest at locus DCA03 (0.939), with a mean value of 0.744.

The observed heterozigosity was higher than the expected heterozygosity across 19 loci, while it showed a slightly lower value than expected heterozygosity at seven loci out of 26, (DCA05, DCA09, DCA11, DCA16, GAPU59, UDO24 and SNB19). This observed heterozigosity deficiency may be related to the presence of null alleles, whose frequency values were positive at six of these loci.

The expected heterzygosity (He) showed the lowest value at locus SiBi 03 (0.294) and the highest at locus DCA09 (0.873), with a mean value of 0.678. This value was similar to those reported in other studies, such as Fendri *et al*. (2010) (0.680) and Diez *et al.* (2011) (0.698). The overall heterozigosity values in our olive genotypes were high, thus indicating the presence of broad genetic diversity, which is explained by high selection of cultivars showing traits of interest, such as the size of the fruit or oil content.

The presence of null alleles favours an increase of homozygosity over heterozygosity due to the absence of amplification of the second allele as a result of any mutation that may have occurred in the flanking region of primer linking. The occurrence of null alleles has already been described for the same microsatellite primers by other authors;

Noormohammadi *et al*. (2007) reported a positive value of null allele frequency at loci DCA05, DCA16, GAPU59 and UDO24, Baldoni *et al.* (2009) at locus DCA09, Lopes *et al*. (2004) at locus DCA11, and Muzzalupo *et al*. (2010) reported the occurrence of null alleles at locus GAPU59.

The average observed heterozygosity in our analysis was higher than the expected heterozygosity [Ho (0.744) > He (0.678)]. Similiarly, higher values of observed than expected heterozigosity were also observed in olive cultivars from Slovenian collection [Ho (0.769) > He (0.705)] (Bandelj *et al*., 2004), Iranian cultivars [Ho (0.660) > He (0.594)] (Noormohammadi *et al.*, 2007), Istrian cultivars [Ho (0.828)> He (0.739)] (Poljuha *et al.*, 2008), Tunisian cultivars [Ho (0.77) > He (0.68)] (Fendri *et al.*, 2010), 126 ancient Spanish olives [Ho (0.740) > He (0.698)] (Díez *et al.*, 2011), 77 accessions belonging to 25 olive cultivars in central-western Spain [Ho (0.780) > He (0.750)] (Delgado-Martinez *et al*., 2012), 65 accessions from Sicily [Ho (0.760) > He (0.748)] (Las Casas *et al*., 2014). The higher heterozygosity observed in the olive genotypes may be the result of high human selection pressure, which acts against homozygosity.

The different results related to polymorphism levels obtained in different studies could be explained by the fact that the number of cultivars and primers used were different. Moreover, these differences may reflect genetic differences within species (Gomes *et al.*, 2009).

Polymorphic information content (PIC) ranged from 0.250 to 0.859 at loci SiBi 03 and DCA 9, respectively, with an average of 0.630. The high PIC value found in most of the microsatellite loci indicated that approximately 73 % of loci (19) could be classified as highly informative, with PIC > 0.5. Around 46 % (12) loci, DCA03, DCA09, DCA16, DCA18, GAPU59, GAPU71B, GAPU101, UDO24, EMO90. SNB03, SNB11, SNB14, showed PIC > 0.7 and could be classified as potential markers for genetic mapping (Bandelj *et al*., 2004).

The PIC value indicates the level of information provided, as well as the usefulness of the microsatellite primers for olive genotyping, gene mapping, molecular breeding and germplasm evaluation. In this sense, the most suitable loci for the genetic characterisation of our olive genotypes were DCA03, DCA09, DCA18, SNB03 and SNB11, which showed polymorphism equal to or higher than 80 %. Based on two distinct discriminative indices, PIC value (0.873) and probability of identity (0.052), locus DCA09 was shown to be the most informative of the twenty six loci analysed, while the least informative locus, showing a low PIC value (0.250) and high PI value (0.628), was SiBi 03 (Table 5).

Table 5: List of microsatellite loci and parameters of variability ranked according to their probability of identity values (PI)
Preglednica 5: Seznam mikrosatelitnih lokusov in parametri variabilnosti razvrščeni glede na vrednosti enakosti genotipov (PI)

Locus	**No**	**Ho**	**He**	**PIC**	**F_{null}**	**HW**	**PI**
DCA 09	19	0.818	0.873	0.859	+0.0197	**	0.052
DCA 18	15	0.851	0.841	0.819	-0.0078	**	0.086
DCA 03	10	0.939	0.812	0.785	-0.0792	**	0.111
GAPU 101	9	0.894	0.807	0.780	-0.0580	**	0.112
SNB 03	11	0.896	0.788	0.762	-0.0814	**	0.119
SNB 11	8	0.932	0.790	0.755	-0.0872	**	0.144
DCA 16	13	0.771	0.777	0.739	+0.0051	**	0.156
UDO 24	11	0.583	0.761	0.725	+0.1236	**	0.162
GAPU 71B	6	0.918	0.770	0.730	-0.0907	**	0.168
EMO 90	5	0.810	0.762	0.721	-0.0329	**	0.175
GAPU 89	9	0.858	0.694	0.662	-0.1391	**	0.184
SNB 14	7	0.918	0.754	0.709	-0.1065	**	0.191
GAPU 59	5	0.567	0.747	0.702	+0.1230	**	0.196
SNB 20	9	0.821	0.704	0.665	-0.1029	**	0.199
DCA 05	10	0.672	0.681	0.646	-0.0131	NS	0.201
DCA14	8	0.674	0.665	0.632	-0.0375	**	0.208
DCA11	7	0.500	0.723	0.674	+0.1692	**	0.218
SNB 19	9	0.623	0.635	0.596	-0.0052	NS	0.246
SiBi 05	4	0.890	0.657	0.590	-0.1692	**	0.314
SiBi 04	4	0.691	0.560	0.499	-0.1123	**	0.374
DCA 15	6	0.541	0.533	0.474	-0.0023	NS	0.393
SiBi 07	5	0.835	0.557	0.465	-0.2099	**	0.471
SiBi 19	5	0.757	0.530	0.425	-0.1830	**	0.534
SNB 22	4	0.658	0.512	0.407	-0.1346	**	0.548
SiBi 11	2	0.574	0.410	0.325	-0.1672	**	0.600
SiBi 03	2	0.357	0.294	0.250	-0.0960	**	0.628
Mean	**7.8**	**0.744**	**0.678**	**0.630**	**-**	**-**	**5.19 x10^{-18*}**

N_o – number of alleles. H_o – observed heterozygosity. H_e – expected heterozygosity. PIC – polymorphic information content. F_{null} - estimated frequency of null alleles. HW – Hardy-Weinberg equilibrium and PI – probability of identity; *cumulative PI.** - statistically significant difference. $P < 0.001$; NS-not significant)

Only three of the 26 loci were in Hardy- Weinberg equilibrium, while the other loci show distortion from equilibrium. The deviation from equilibrium at loci DCA09, DCA11, DCA16, UDO24, and GAPU59 may be related to their positive values of null allele frequencies. The detection of null alleles in microsatellites, especially in outbred heterozygous species, is common (Rallo *et al*., 2000) and leads to increased homozygosity.

The Hardy-Weinberg equilibrium of most of the loci was not consistent among different populations. It was different when analysing it within different populations, such as wild, ancient, cultivated or foreign cultivars, indicating that the mixture of previously isolated populations maybe the reason for this distortion from Hardy-Weinberg equilibrium. As suggested by Belaj *et al*. (2007), the Wahlund effect, i.e., the structuring of populations in subunits within which mating is more probable, cannot be completely ignored in the case of *in situ* sampled trees that are geographically distant.

Five loci, DCA09, SNB11, SNB14, SiBi03 and SiBi11 showed a deviation from Hardy-Weinberg with all olive groups.This was probably due to the positive value of null frequency at locus DCA09, and the unequal distribution of alleles for the remaining loci, suggesting that attention should be paid in the parentage analysis interpretation at these loci. The allele frequencies for each of the 26 microsatellite loci are given at Annex C.

Probability of identity (PI) calculated for each locus ranged from 0.052 to 0.628 at DCA09 and SiBi03, respectively, with an accumulated probability of identical genotypes for all loci of 5.19×10^{-18}, indicating that the probability of two randomly sampled olive genotypes having the same allelic profile is extremely low. The higher values of probability of identity shown by some loci (SiBi07, SiBi19, SNB22, SiBi11, and SiBi03) may be attributed to the low number and unequal distribution of alleles. The effect on the variation of probability of identity between loci from an uneven allele frequency distribution was also reported by Lopes *et al.* (2004). At loci Sibi19 and SiBi07, two common alleles account for 96 % and 93 % of the allele frequencies, respectively, while the remaining 4 % and 7% were shared among the other three alleles of these loci. At locus SNB22, two alleles account for 97 % of allele frequencies and the other two account for only 3 % of allele frequencies. One of the two alleles at each of the loci, Sibi03, SiBi04 and SiBi05, showed up with a frequency of 82 % and the other only of 18 %. At locus SiBi11, the allele frequency was shared between two alleles, with 71 % and 29 % each.

The obtained PI value (5.19×10^{-18}), was lower than the PI obtained by the analysis of 92 accessions of 10 Iranian cultivars analysed by 13 SSR markers (1.2×10^{-9}) (Noormohammadi *et al*., 2007), 104 Greek accessions by 10 SSR (1.65×10^{-5}) (Roubos *et al*., 2010), 184 Tunisian accessions by 8 SSR (1.23×10^{-7}) (Fendri *et al.*, 2010), 24 Spanish cultivars (1.70×10^{-11}) by 10 SSR (Diaz *et al*., 2007), 19 cultivars from Slovenian olive collection by 14 SSR markers (1.433×10^{-11}) (Bandelj *et al.*, 2004), 142 Italian samples by 10 SSR (2.113×10^{-9}) (Beghe *et al*., 2011), 78 accessions of wild, cultivated and ancient trees from Sardinia by 13 SSR markers (5.274×10^{-12}) (Erre *et al*., 2010), 60 Brasilian cultivars by12 SSR (1.51×10^{-10}) (do Val *et al.*, 2012), 12 SSR, 126 ancient Spanish cultivars by 14 SSR markers (1.96×10^{-16}) (Díez *et al*., 2011), 505 accessions from 14 Mediterranean countries (2.55×10^{-14}) (Haouane *et al.*, 2011) and 66 olive cultivars from

three regions of Southern Italy (7.23×10^{-14}) (Marra *et al.*, 2013) by the use of 12 SSR markers.

The overall low probability of obtaining identical genotypes indicated that the set of markers used was notably effective for genotyping the sampled set of genotypes.

The observed frequency of 203 alleles ranged from 0.003 to 0.821, with a mean of 0.22. There were 92 rare alleles detected (freq <5 %), representing around 45 % of the total of 203 amplified alleles. The rare alleles per each locus are given in bold in Annex C. The highest number of rare alleles was detected at loci DCA09, with 12 rare alleles, and DCA16 and DCA18 with nine rare alleles, indicating the ability of these markers to identify rare alleles, which is extremely useful in cases of these alleles being associated with specific characteristics with agronomic value. However, at other loci, such as EMO90, SiBi03, Sibi05 and SiBi11, there were no rare alleles (Annex C).

Genotype specific alleles are considered to be those appearing only once, as revealed by means of MICROSAT software (Minch *et al*., 1997) based on their frequency. In total, 31 specific alleles were found in 24 genotypes, 29 alleles being specific to 22 Albanian olive genotypes (Table 6). Five out of 24 genotypes that had specific alleles, 'Carboncella R', 'Gordall Sevillana E', 'Kallmet Kokërr Vogël I', 'Olivastër Kokërr Hollë Elbasani E' and 'Unafka I', each had two specific alleles at two different loci. Only the cultivar 'Çerje' had the two specific alleles at the same locus (SiBi19). These alleles should be further evaluated and analysed, since they may be possible genotype specific markers, hence useful for further characterization of these genotypes. They may also represent a useful tool for molecular tracking of the genotype origin of oil (Rotondi *et al*., 2011).

The occurrence of three alleles amplified by one genotype was observed in three microsatellite loci, DCA09, DCA11 and GAPU71 B. It could be speculated that the third allele is due to a somatic mutation, also referred to as chimerism (Baali-Cherif and Besnard, 2005). The presence of three-allelic profiles with several olive cultivars was also observed by Poljuha *et al*. (2008), Soleri *et al.* (2010) and Bandelj *et al.* (2004).

Three alleles were found at loci DCA09 in five ancient cultivars. The allelic profiles with cultivars Kalinjot I and Kryps Elbasani I were 177: 187: 201; with Ulli I Bardhë I Shkodrës I 174: 187: 219, Vajs Peqini I 177: 201: 221 and with Pulazeqin I 177: 201:221; while one ancient cultivar, Kokërr Madh Berati I, showed three alleles 143: 158: 177 at locus DCA11.

The oleaster genotype Olivaster10 I showed a tri-allelic profile at both loci DCA09 (181: 187:219) and DCA11 (151-158-163). Another ancient cultivar, Gjykatës I, showed three alleles (139:141:145) at locus GAPU71B.

The presence of chimerism in old cultivars is fairly expected due to the accumulation of mutations.There are cases in which clones can be distinguished from each other by the presence of the third allele (Baali-Cherif and Besnard, 2005).

Table 6: Genotype specific alleles
Preglednica 6: Specifični aleli posameznih genotipov

Genotype	**Allele**	**Locus**	**Genotype**	**Allele**	**Locus**
Arbequina Ref*	223	UDO24	Micka I	193	DCA 09
Carboncella R**	213	DCA 05	Nisjot I	206	DCA 14
Carboncella R**	174	DCA 16	Nocellara del Belice E**	196	DCA 18
Črnica Ref*	210	SNB 19	Nolca E**	163	GAPU 89
Çerje I	230	SiBi 19	Ogliarola E**	225	GAPU 89
Çerje I	234	SiBi 19	Olivarstër 4- I	182	SNB 03
Gordal Sevillana E**	186	SiBi 04	Olivastër 10- I	280	DCA 15
Gordal Sevillana E**	174	SiBi 07	Olivastër Kokërr Hollë Elbasani E	217	DCA 05
Kala I	214	DCA 18	Olivastër Kokërr Hollë Elbasani E	180	DCA 16
Kallmet Kokërr Vogël I	200	SNB 19	Pulazeqin R	211	GAPU 101
Kallmet Kokërr Vogël I	167	SNB 20	Simona E**	200	DCA 18
Klon XII-E	180	SNB 11	Ulli I Bardhë Pobrati E	169	SNB 22
Klon XIII-E	213	UDO 24	Ulli I kuq II-I	183	DCA 11
Kushan I-I	158	SNB 03	Unafka I	217	DCA 09
Marks I	202	UDO 24	Unafka I	200	DCA 14
Mastoidis E**	279	DCA 15			

*Reference cultivars from Slovenia; ** Cultivars of foreign origin in Albania

The AMaCAID script within R software was used to evaluate the discriminative power of the markers through analysis of the maximisation of discriminated genotype markers as a function of loci (Figure 5). The maximum number of genotypes discriminated by each loci combination is given in Table 7.

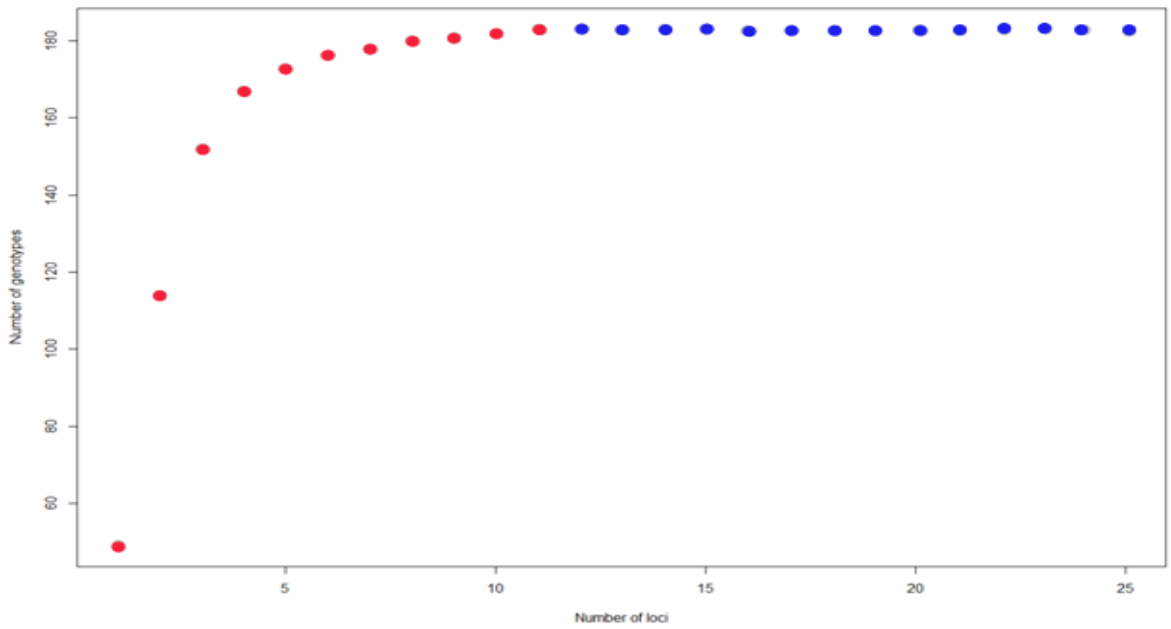


Figure 5: Maximum number of genotypes discriminated as a function of the number of loci used
Slika 5: Število genotipov, ki jih je moč ločiti glede na število analiziranih lokusov

Table 7: Maximum number of genotypes discriminated by each of the loci combinations
Preglednica 7: Največje število genotipov ločenih s posameznimi kombinacijami lokusov

No	**Loci combination**	**Discriminated genotypes**	**%**
1	DCA 09	49	26.7
2	DCA 09 + EMO 90	114	62.2
3	DCA 18 + GAPU 59 + EMO 90	152	83
4	DCA 18+ GAPU 59 + SNB 19 + SNB 22	162	88.5
5	DCA 18+ GAPU 59 + SNB 19 + SNB 22 + SiBi 04	173	94.5
6	DCA 18+ GAPU 59 + SNB 19 + SNB 20+ SNB 22 + SiBi 04	176	96.2
7	DCA 18+ GAPU 59 + UDO 24 + SNB 19 + SNB 20 +SNB 22 + SiBi 04	178	97.2
8	DCA 18 + GAPU 59 + UDO 24 + EMO 90 + SNB 19 + SNB 20 + SNB 22 + SiBi 04	180	98
9	DCA 09 + DCA 18+ GAPU 59 + GAPU 71B +UDO 24+ EMO 90 + SNB 19+ SNB 20+ SiBi 04 + SiBi 07	181	98.9
10	DCA 09 + DCA 18+ GAPU 59 + GAPU 71B+ UDO 24 + EMO 90 + SNB 19 + SNB 20+ SNB 22 + SiBi 04 + SiBi 07	182	99.5
11	DCA 09 + DCA 18+ GAPU 59 + GAPU 71B+ UDO 24 + EMO 90 + SNB 19 + SNB 20+ SiBi 04 + SiBi 07 + SiBi 19	183	100

A set of 11 markers, DCA09, DCA18, GAPU59, GAPU71, UDO24, EMO90, SNB19, SNB20, SiBi04, SiBi07 and SiBi19, discriminated all 183 olive genotypes in this study (Figure 5 and Table 7). This highly discriminative set of 11 markers consisted of six previously developed genomic SSR markers and five EST-SSR markers developed in this study. Five of the six g-SSR markers are also included in the consensus list of olive SSR markers (Baldoni *et al.,* 2009). The high discriminating capacity of the combined 11 loci was supported by the low probability of classifying two random accessions under an identical profile (PI =1.91 x 10^8). As suggested by other studies, the level of polymorphism displayed by this set of markers may be even higher without closely related genotypes (Haouane *et al*., 2011).

A combination of only three loci, DCA18, GAPU59 and EMO90, enabled discrimination of 83 % (152) of unique genotypes (Table 7). These three microsatellite markers may be very useful for rapid screening of olive gemplasm in a collection.

5.2 COMPARATIVE ASSESSMENT OF GENOMIC AND EST DERIVED SSR

The development of microsatellites (SSR-g) in olive has so far been mainly based on the sequencing of $(GA)_n$ and /or $(CA)_n$ enriched genomic libraries (Sefc *et. Al*., 2000; Rallo *et al*., 2000; Carriero *et al.,* 2002; Cipriani *et al*., 2002; De la Rosa *et al*., 2002; Díaz *et al*., 2006a; Sabino-Gil *et al*., 2006), through a costly and time consuming process.

Recent sequencing projects have generated and deposited a vast number of expressed sequence tags (EST) in public databases, providing the means of development of new microsatellites with lower costs and effort. The development of the first eight hexa-nucleotide microsatellite primers from expressed sequence tags (EST) in olive was reported by De la Rosa *et al*. (2013), four of them showed high discriminatory power and a proven usefulness in paternity testing comparable to that of four di-nucleotide SSR used on the same sample set.

A total of 625 olive EST sequences obtained at the Centre of Plant Biotechnology and Breeding, Biotechnical Faculty, University of Ljubljana were examined for the presence of tri- and tetra-nucleotide repeat motifs. Only 50 % of ESTs containing tri-nucleotide and 57.7 % of ESTs containing tetra-nucleotide motifs, out of 75 EST sequences that had tri- (30) and tetra-nucletide (45) repeat motifs, were suitable for primer design, due to an inappropriate flanking region of motifs.

Around 40 % and 23 % of designed tri-nucleotide and tetra-nucleotide repeat primers amplified the PCR product and gave a clear pattern, enabling easy scoring and allele size determination. The successful amplification rate of new developed EST-SSRs was

somewhat lower than the amplification rate success with g-SSRs as reported in studies of marker development by Sefc *et al.* (2000) (53.7 %), Rallo *et al.* (2000) (84.6 %), Carriero *et al.* (2002) (100 %), De la Rosa *et al.* (2002) (53.8 %), Cipriani *et al.* (2002) (83.3 %) and Díaz *et al.* (2006) (50 %).

In order to ascertain the effectiveness of EST-SSR markers, a comparison of 26 microsatellite markers (Table 2 and 3) was made using genetic diversity parameters: (Na, Ho, He, PIC, and PI). The mean values of diversity parameters were compared between the two classes of microsatellites (14 g-SSRs *vs* 12 EST-SSRs), as well as between different lengths of repeat motif (14 di-nucleotide, 6 tri-nucleotide and 6 tetra-nucleotide microsatellites). The results are given in Table 8.

The EST-SSR produced better quality allelic patterns and they were not as prone to stutters as di-nucleotide genomic-SSR, enabling better separation of alleles, although the application of a long final extension time also greatly reduced stutter on di-nucleotide loci. Several other authors have previously observed a better quality allelic pattern of EST-SSRs, with less stutter (Pashley *et al*., 2006; Woodheadet *al*., 2005; Gadaleta *et al.*, 2007; Simko, 2009).

The occurrence of null alleles was only observed in g-SSR, while their frequencies at EST-SSR loci had a negative value. Primers flanking EST-SSRs are derived from relatively conserved sequences of coding regions; it is therefore likely that null alleles will be less of a problem for EST-SSRs than for their anomymous counterparts (Ellis and Burke, 2007).

Genomic-SSRs amplified alleles in 98 % of genotypes, yielding 65.5 % (133) of a total of 203 alleles obtained by the twenty six microsatellite primer pairs. EST-SSRs amplified alleles in 94 % of genotypes, yielding 34.5 % (70) of the total alleles. Among markers with different lengths of repeat motif, the average number of alleles per locus amplified by di-nucleotide markers (9.5) was in the same range as the mean number of alleles amplified by tri-nucleotide markers (8) but much higher than the average number of allele's amplified by tetra-nucleotide markers (3.7). The low level of polymorphism detected by tetra-nucleotide markers might be due to the lower ratio of mutations in repeat motifs with a long core unit.

The average number of alleles per locus was obviously higher in g-SSRs (9.5) than in EST-SSRs (5.3), showing a higher discriminative power of genomic microsatellites.

The lower level of polymorphism generally detected in EST-SSRs than in g-SSRs (Hu *et al.*, 2011) might be due to selection against variation in the conserved region of the EST-SSRs (Zhang *et al*., 2014), in order to maintain an expressed trait and also to the conservative region EST-SSRs represent. Although in *Jathropa curcas* L., Wen *et al*.

(2010) observed a higher mean allelic value with EST-SSRs than with g-SSRs, suggesting that this result may vary in different species

The higher level of discriminative power of genomic microsatellites over EST-SSRs in our study was expected due to the different length in repeat motif of these two types of microsatellites; g-SSR have di-nucleotide repeat motifs while EST-SSRs have tri-and tetra-nucleotide repeat motif, which are subjected to lower variability

Table 8: Mean genetic diversity parameters of different types of microsatellite loci; g-SSR, EST-SSR, di-trei- and tetra-nucleotide repeat motif microsatellites of different types of microsatellites used in the analysis
Preglednica 8: Povprečje parametrov genetske variabilnosti izračunanih za posamezne skupine mikrosatelitov; g-SSR, EST-SSR, di- in tetra-nukleotinih ponovitev uporabljenih za analizo 138 genotipov oljke

Parameters of variability		**Genomic SSR**	**EST-derived SSR**	**Dinucleotide repeat motif SSR**	**Trinucleotide repeat motif SSR**	**Tetranucleotide repeat motif SSR**
	Na	133	70	133	48	22
	Mean	**9.5**	**5.8**	**9.5**	**8**	**3.7**
Na	Min	5	2	5	4	2
	Max	19	11	19	11	5
	Mean	**0.742**	**0.746**	**0.742**	**0.808**	**0.684**
Ho	Min	0.500	0.357	0.500	0.623	0.357
	Max	0.939	0.932	0.939	0.932	0.890
	Mean	**0.746**	**0.599**	**0.746**	**0.697**	**0.501**
He	Min	0.533	0.294	0.533	0.512	0.294
	Max	0.873	0.790	0.873	0.790	0.657
	Mean	**0.710**	**0.537**	**0.710**	**0.649**	**0.425**
PIC	Min	0.474	0.250	0.474	0.407	0.250
	Max	0.859	0.762	0.859	0.762	0.590
	PI*	**5.28×10^{-12}**	**9.82×10^{-7}**	**5.28×10^{-12}**	**8.8×10^{-5}**	**1.11×10^{-2}**
PI	Min	0.052	0.119	0.052	0.119	0.314
	Max	0.393	0.628	0.393	0.548	0.628

N_a – number of alleles. H_o – observed heterozygosity. H_e – expected heterozygosity. PIC – polymorphic information content. PI – probability of identity; *cumulative PI.

Of a total of 92 rare alleles (freq <5%) with an average of 3.65 alleles per locus, 61alleles (66 %) were amplified by genomic SSRs and only 34 alleles (31 %) were amplified by EST-SSRs: 22 (24 %) by trinucleotide SSRs and 9 (9 %) by tetranucleotide SSRs (Annex C). There were 31 alleles that appeared only once; these are considered specific alleles because they are linked to a specific genotype (Table 6). The highest number of specific

alleles, 20 (64.5 %), was amplified by g-SSR loci, while 11 alleles (35.5 %) were amplified by EST-SSR loci. Among microsatellites with different lengths of repeat motif, those with a tetra-nucleotide repeat motif gave the lowest number of specific alleles 4 (13 %), while the number of specific alleles amplified from di- and tri-nucleotide microsatellites was 20 (65.4 %) and 7 (22.5 %), respectively.

In the g-SSRs the mean values of observed and expected heterozigosity were in range 0.742 and 0.746 respectively. In the EST-SSRs the mean value of observed heterozygosity (0.746) was higher than the mean value of expected heterozygosity (0. 599) (Table 8). The expected heterozigosity value of EST-SSRs has a mean value in the same range as the expected heterozigosity values obtained by the whole number of microsatellites developed in olive and in the majority of outcrossing species that are vegetatively propagated (Diaz *et al*., 2006). The higher values of observed over expected heterozigosity showed by EST-SSRs were due to the high selective pressure on these loci, while the selective pressure might have been lower in neutral g-SSRs loci and their observed over expected heterozygosity showed lower values.

The average polymorphism of the analysed loci was 71 %, 64.9 % and 42.5% for di-, tri- and tetranucleotide microsarelite, respectively. The lowest level of polymorphism was obtained by tetranucleotide microsatellites, while the two other microsatellite categories showed roughly the same level of polymorphism. The majority of primers showed a PIC value > 0.5 (Table 8), indicating their ability to assess molecular diversity and genetic relatedness. Both g-SSRs and EST-SSRs markers can therefore be usefully applied in olive fingerprinting analysis.

The probability of obtaining identical genotypes was lower with g-SSRs (5.28×10^{-12}) than with EST-SSRs (9.82×10^{-7}). The probability of obtaining an identical genotype was much lower in dinucleotide loci (5.28×10^{-12}) than with trinucleotide (8.8×10^{-5}) and tetranucleotide (1.11×10^{-2}) microsatellites.

The goodness of fit between the similarity matrices produced by g-SSRs and EST-SSRs was analysed using the MXCOMP module of the Mantel test (Mantel, 1967) implemented in NTSYS v. 2.2 (Rohlf, 2005). The correlation between the genetic similarity matrices of g-SSRs and EST-SSRs was moderate but still significant ($r = 0.522$, $P < 0.005$). The goodness of fit of similarity matrices of g-SSR and of the two markers, g-SSRs and EST-SSRs, is higher ($r = 0.915$, $P < 0.005$) than the correlation between EST-SSRs based similarity matrices and the two markers, g-SSRs and EST-SSRs ($r = .809$, $P < 0.005$). The high mantel test correlations indicate that both types of markers are effective for genetic diversity study of olive. The data generated by g-SSRs resulted more congruent with the combined data of two types of markers. The combined use of both types of markers, g-

SSRs and EST-SSRs provides information from different regions of genome, ensuring better genetic diversity assessments.

Dice similarity coefficients were used to construct dendrograms based on the UPGMA algorithm, using a dataset of g-SSRs, EST-SSRs and combined markers. The average Dice's similarity obtained by g-SSR (0.414) was lower than the average similarity revealed by EST-SSR markers (0.628). There were no significant differences in overall grouping of olive genotypes; some groups consistently clustered the same in both dendrograms. Graphic presentations of correlations are given at Annex D.

Both categories of markers developed from genomic libraries, as well as those derived from expressed sequence tags, were suitable for the analysis of our sample set. Genomic microsatellites showed overall higher levels of polymorphism than EST-derived microsatellites. Nevertheless, the latter were still sufficiently informative and showed a slightly higher level of heterozigosity than genomic derived SSR. Microsatellites loci with di-nucleotide and tri-nucleotide repeat motifs showed higher degree of polymorphism than the microsatellites with the tetra-nucleotide repeat motif. The high values of polymorphism obtained in this study indicate that the EST-SSRs, especially those with a tri-nucleotide repeat motif, could efficiently be applied in analyses of genetic diversity in olive.

5.3 GENETIC DIVERSITY OF OLEASTER AND CULTIVATED OLIVES

The coexistence of cultivated olives with their wild relatives, the oleasters, in Mediterranean countries, has been inferred from archeological and paleobotanical findings (Terral *et al*., 2004). In some areas, cultivar-feral-wild olives live together within a radius of a few meters and they are not easily distinguished (Breton *et al*., 2008). The diversity of feral populations, due to the derivation of a new seedling by sexual reproduction, grown without any agricultural aid, could give rise to varieties with superior traits for cultivation (Breton *et al*., 2008). The oleasters could contribute to the explanation of domestication processes (Hannachi *et al*., 2008).

In Albania, feral and wild olives (oleasters) have also been found in isolated areas and in association with cultivated olives, for which they are thought to be pollinators; their estimated age ranges from 200-1000 years (Ismaili, 2013; Frezzoti, 1930, *cit. by* Belaj *et al*., 2003a).

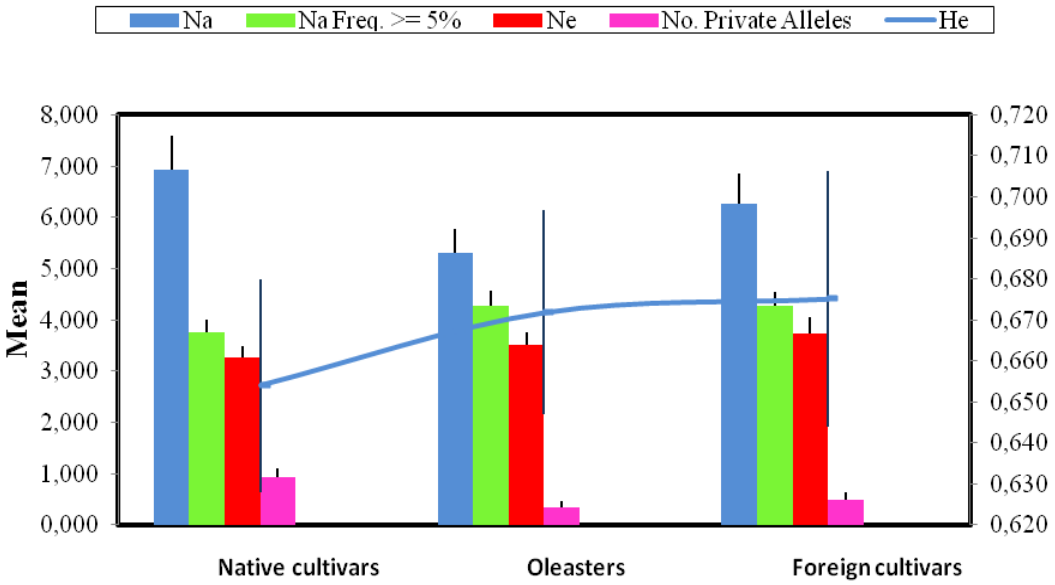
We analyzed 26 microsatellite loci for 183 unique olive genotypes; 18 oleaster genotypes, 120 native olive cultivars (64 current cultivars and 56 ancient cultivars) and 45 foreign olive cultivars. The parameters of diversity were calculated for each group (Tables 9-10 and Figure 6).

The average allele number per loci obtained for the 120 native cultivated trees (6.92) and for foreign cultivars (6.2) was higher than the average number of alleles amplified in oleaster genotypes (5.31). The number of alleles per locus detected in oleaster genotypes was in the same range as that detected in 48 genotypes of wild olive in Spain (5.62) (Belaj *et al*., 2011) but lower than the average number of alleles detected in a study of 21 wild olive individuals in the region of Sardinia (8) (Erre *et al.,* 2010). The mean number of effective alleles (Ne) was higher in oleaster genotypes (Table 9).

In total, 46 private alleles, unique alleles for a single population, were detected in three olive groups. The highest number of private alleles was found in native cultivated genotypes (current and ancient cultivars), 24 (52 %), whereas 13 (28.3 %) private alleles belonged to foreign olive cultivars and only 9 (19.6 %) private alleles were observed in oleaster genotypes (Table 9 and Figure 6).

In agreement with the previous studies that compared genetic diversity in wild and native cultivated olive groups (Belaj *et al.,* 2010; Erre *et al*., 2010; Haouane *et al.,* 2011), the mean expected heterozigosity value was higher in oleaster genotypes (He=0. 692) than in cultivated olives (He=0.657), whereas the observed heterozigosity values were higher in cultivated olives (Ho=0.760) than in oleaster genotypes (Ho=0.745). However, the level of heterozigosity showed by oleaster genotypes was similar to previous studies carried out in wild olives (Breton *et al*., 2006; Belaj *et al*., 2007; 2011, Erre *et al*., 2010). The mean observed heterozigosity in oleaster genotypes (Ho=0.745) was higher than expected (He=0. 692). The same was observed in the cultivated group in which the mean observed heterozigosity was higher (Ho=0.760) than expected (He=0.657) (Table 9 and Figure 6).

The lower diversity detected in oleaster genotypes compared with native cultivated olives in our study could be due to the low number of oleasters (18) in comparison to cultivated olives (120). It might also be due to the possibility that some of the sampled oleasters represent feral forms.



Na – no of different alleles, Na Freq ≥5% -no. of different alleles with a frequency ≥ 5%; Ne –no.of effective alleles=1/ (sum pi^2), No. of private alleles, He expected heterozygosity = 1-Sum pi ^2

Figure 6: Allelic patterns across oleaster and cultivated olive genotypes and their expected heterozygosity
Slika 6: Parametri variabilnosti izračunani za skupino divjih in skupino kultiviranih oljk in njihova pričakovana heterozigotnost

Table 9: Genetic diversity parameters found in different olive groups/populations: olive cultivars (current and ancient cultivars), oleasters and imported foreign olivecultivas
Preglednica 9: Parametri genetske variabilnosti izračunani za različne skupine/populacije: kultivarjev (sodobni in starodavni), divje oljke in tuje kultivarje oljke

Genetic diversity parameters		**Cultivars (120)**	**Oleaster (18)**	**Foreign (45)**
No	**Total**	**180**	**138**	**163**
No	**Mean**	**6.92**	**5.31**	**6.2**
	Min	2	2	2
	Max	18	11	14
Na (freq≥5%)	Mean	3.769	4.269	4.269
Npr	**Mean**	**0.923**	**0.346**	**0.500**
Ne	Mean	3.28	3.51	3.72
Ho	**Mean**	**0.760**	**0.745**	**0.703**
	Min	0.403	0.333	0.156
	Max	0.949	1	0.932
He	**Mean**	**0.657**	**0.692**	**0.683**
	Min	0.323	0.413	0.145
	Max	0.821	0.868	0.899

N_o – number of alleles. Na-number of alleles with frequency ≥5%,H_o– observed heterozygosity, Ne-number of effective alleles, Npr-number of private alleles per population, H_e – unibased expected heterozygosity =(2N/(2N-1))*He, *In paranthesis the number of accessions per each group

The genetic parameters calculated for native olive genotypes (64 current, 56 ancient cultivars and 18 oleasters), are presented in Table 10.

The average number of alleles detected in currently cultivated olives and ancient olives were within the same range, 6.04 and 6.08, respectively but higher than the average number of alleles detected in oleasters, while the average number of effective alleles was higher in oleasters (3.51) than in the other two groups of olive cultivars (Table 10).

Forty-one private alleles were detected in native current cultivars, ancient ones and oleasters. The number of private alleles is much higher in ancient 19 (46.3 %) than in current 12 (29.3 %) cultivars and oleaster 10 (24.4 %), suggesting that some alleles may be lost through the intensive human selection pressure on the *ex-situ* collection towards cultivars with specific desired qualities.

The group of ancient olive cultivars showed higher observed than expected heterozigosity. Observed heterozigosity was higher in ancient cultivars (Ho=0.817) than in current cultivated ones (Ho=0.710). However, the observed heterozygosity in both the currently cultivated olives and the ancient ones was higher than the expected ones.

Table 10: Genetic diversity parameters found in different olive groups/populations: current, ancient olive cultivars and oleasters
Preglednica 10: Parametri genetske variabilnosti izračunani za različne skupine/populacije: kultivarje in divje oljke

Genetic diversity parameters		**Ancient (56)**	**Current (64)**	**Oleaster (18)**
No	**Total**	**158**	**157**	**138**
No	**Mean**	**6.08**	**6.04**	**5.31**
	Min	2	2	2
	Max	13	16	11
Na (freq≥5%)	Mean	3.500	3.769	4.269
Npr	**Mean**	**0.731**	**0.462**	**0.346**
Ne	Mean	3.18	3.08	3.51
Ho	**Mean**	**0.817**	**0.710**	**0.745**
	Min	0.446	0.047	0.333
	Max	0.982	0.969	1
He	**Mean**	**0.657**	**0.624**	**0.692**
	Min	0.438	0.046	0.413
	Max	0.831	0.819	0.868

N_o – number of alleles. Na-number of alleles with frequency ≥5%, H_o – observed heterozygosity, Ne-number of effective alleles, Npr-number of private alleles per population, H_e – unibased expected heterozygosity = (2N/(2N-1))*He, *In paranthesis the number of accessions per each group

5.4 ANALYSIS OF IDENTITY

The names of olive cultivars mainly refer to certain morphological characteristics, such as the color, the size or the end-use of the fruit; for example, 'Kokërr Vogël' (Small fruit), 'Ulli i Bardhë' (White Olive), 'Olivastër Cilindrike' (Cylindrical Oleaster), 'Vajs' (Oily) and 'Kryps' (Salty) for table or canned olives. Some other olive names are related to their putative origin or main cultivation area, for example 'Ulli i Kuq i Damsit' (Red Olive of Dams), 'Ulli i Zi i Tiranës' (Black Olive of Tirana), etc.

The generic denomination of olive cultivars has been reported to be the main cause of varietal disorder and the presence of many cases of synonymy and homonymy in many olive growing countries and olive collections, such as Istrian olive cultivars (Poljuha *et al.*,2008), Iranian olive germplasm (Noormohammadi *et al.*, 2007); Italian olive germplasm (Muzzalupo *et al*., 2009a, 2014), Tunisian cultivars (Fendri *et al*., 2010), Brazilian olive germplasm (do Val *et al.*, 2012), Tunisian germplasm (Rekik *et al.*, 2008) Iberian cultivars (Fendri *et al*., 2014), and in the WOGB collection of Cordoba (Trujillo *et al*., 2014).

Identification of possible cases of synonyms and homonyms will enable proper management of olive germplasm. The analysis of identity of 194 olive genotypes by means of 26 microsatellite loci allowed the detection of 183 unique genotypes. The overall probability that two unrelated individuals drawn randomly from the sample set can have an identical genotype at 26 microsatellite loci was very low ($PI = 5.19 \times 10^{-18}$).

5.4.1 Synonymy

Identity analysis of samples based on perfectly matched microsatellite allelic profiles at 26 analysed loci revealed six groups of perfect synonyms. None of these synonymic groups have been previously reported. The cases of synonymic groups are presented in Table 11.

Identity analysis revealed the identity of two unknown ancient olive cultivars coded G48 I and G49 I, collected *in situ* in the north of Albania. The genetic profile of cultivar G48 I perfectly matched that of cultivars Ganjoll I = Ulli I Zi Karpan I = Ulli i Zi i Tiranës III-I = Ulli i Bardhë i Krujës I (Ganjoll I= G48= Black Olive of Karpan = Black Olive of Tirana III-I = White Olive of Kruja I).

All cultivars of the first synonymic group belong to the black olive form except the cultivar Ulli i Bardhë i Krujës I (White Olive of Kruja I), which is obviously a case of misllabelling or mistaken since its allelic profile differ from the profiles of the two other clones, Ulli i Bardhë i Krujës I-E and Ulli i Bardhë i Krujës II-E at 16 and 17 loci over 26 analysed,

respectively. The identity of the unknown cultivar G49 I was revealed and it had the same allelic profile as Llashtrak I-I. Both of these cultivars were collected in the same region in northern Albania.

Table 11: Groups of perfectly matching olive genotypes with the same allelic profile at 26 analysed microsatellite loci
Preglednica 11: Skupine genotipov, ki imajo enake alelne profile na vseh 26-ih analiziranih mikrosatelitnih lokusih

Groups of perfect synonyms	**No**
Ganjoll I = G48 I = Ulli I Zi iKarpan I = Ulli I Zi I Tiranës III-I = Ulli I Bardhë I Krujës I	1
G49 I' = Llashtrak I-I	2
Frëng I-I = Kushan Kokërr Madh II-E = Kushan Kokërr Mesëm I-E	3
Kushan Kokërr Madh III-E = Kushan Kokërr Mesëm II-E	4
Llashtrak II-I = Kokërr Madh Elbasani I-I	5
Ulli I Bardhë I Tiranës III-I = Ulli I Zi I Tiranës II-I	6

The cultivars of the third and fourth group of synonymy were considered an intra-cultivar variation of the known olive cultivar 'Kushan', due to their different size of the fruit. On the basis of fruit size these considered variations of catalogued cultivar 'Kushan' were named Kushan Kokërr Madh II-E/ III-E (Big Fruit Kushan II-E/III-E) and Kushan Kokërr Mesëm I-E/ II-E (Medium Fruit Kushan I-E/II-E). These cases of synonymy are evidence that variation in morphological traits within the same cultivar is not always based on allelic pattern differences.

The identical allelic profiles of two ancient cultivars in group six collected in the central region of Albania, Ulli i Bardhë i Tiranës III-I = Ulli i Zi i Tiranës II-I, (White Olive of Tirana III-I= Black Olive of Tirana II-I) were not predictable.

Near synonyms included cultivar pairs whose microsatellite allelic profile differed at ≤ 2 loci. These slight differences might be due to somatic mutations, reported as a frequent event in long-lived plants that are vegetatively propagated, especially in olive (Cipriani *et al.*, 2002; Lopes *et al.*, 2004; Mantia *et al.*, 2005 and Muzzalupo *et al.*, 2009b, 2014). Alternatively, as suggested by Baldoni *et al.* (2009), small allele differences of 2 bp may be genotyping errors due to stutter bands.

Nine cases of near synonymy were revealed (Table 12), 7 groups of which were ancient individuals with different names, indicating a mutational origin of these small allelic differences accumulated in ancient olives due to their very old age. The mutations expressed on a distinct phenotype may have been the reason why the different trees of the

same cultivar were differently denominated. However, further evaluation of all cases of synonymy should be done, in order to rule out the possibility of misnomers and to confirm their true-to-type identity. The groups of near synonyms included genotypes differing in ≤ 2 loci. An exception was made for two cultivars, Gjykatës I and Boç I, in group number 2, and genotypes Ulli I Bardhë and Tiranës V-I in the third group, which differ from some of the genotypes in their near synonymic group at 3-4 loci (Table 12). Examination of their allelic profile indicated that the difference observed for these genotypes was due to 2 bp differences at loci DCA 14 and DCA 18, which showed a stutter pattern, although possible mutation at these loci is certainly a possibility given the great age of these trees.

There were no observed cases of synonymy or near synonymy between native olive genotypes and foreign imported cultivars. They were all characterized as distinct genotypes.

Table 12: Genotypes considered near synonyms with difference ≤ 2 loci out of 26 microsatellite loci
Preglednica 12: Seznam genotipov (sinonimov), ki se med seboj razlikujejo na ≤ 2 lokusih od 26-ih analiziranih

Groups of near synonyms	**No**
Bahuta I ≠ (Ulli I Bardhë I Tiranës III-I = Ulli I Zi I Tiranës II-I)	1
Boç I ≠ Gjykatës I ≠ Krypsi I Krujës I-I ≠Krypsi I Shkodrës I ≠ Ulli I Bardhë I Shkodrës I ≠ (Llashtrak I-I = G49)	2
Frëng II-I ≠ Mixan I ≠ Karen I ≠ Ulli I Bardhë I Tiranës II-I ≠ Ulli I Bardhë I Tiranës V-I	3
Kushan Kokërr Mesëm I-E ≠ Ulli I Bardhë I Tiranës IV-I	4
Kuleks I ≠ Perk I ≠ (Ganjoll I = G48 = Ulli I Zi I Karpan I = Ulli I Zi I Tiranës III-I =Ulli I Bardhë I Krujës I)	5
Klon VII-E ≠ Kokërr Madh Berati I-R	6
Kokërr Madh Berati I ≠ Kokërr Madh Elbasani II-I	7
Olivastër e Kuqe e Tiranës I ≠ Olivastër e Zezë e Tiranës I	8
Unafka E ≠ Unafka R	9

In line with previous studies based on morphological descriptors, in the reported case of synonymy between Boç and Gjykatës (Osmani, 1995; Ismaili and Çeloaliaj, 1995; Thomaj and Panajoti, 2003) the two differed only at 2 loci. However the reported pairs of synonyms of Nisiot and Ulli i Hollë Himare, and Ulli i Bardhë i Tiranës and Qumështor (Thomaj and Panajoti, 2003) did not find support in our results. Many cases of possible synonymy were observed in this study for the first time (Table 12).

5.4.2 Intra-cultivar variation, homonymy

Some cultivars are represented by more than two to three samples in the *ex situ* collection or were collected at different locations *in situ* to confirm their identity and true to typeness, as well as to investigate possible cases of homonymy.

Only in one case did two olive accessions with the same name match perfectly in all analysed loci; the accessions Olivastër e Zezë Elbasani I-E and Olivastër e Zezë Elbasani II-E (Black Oleaster of Elbasan I-E & II-E) showed a perfect match at 26 loci, while they differ from the third accession Olivastër e Zezë Elbasani I, collected *in situ* in 21 out 26 loci, so it is obviously not a case of true–to-typeness. Registered in the *ex situ* collection as intra-cultivar variations, these two perfectly matched oleasters are a case of duplication that needs to be examined in order to provide better collection management.

A high level of intra-cultivar variability, (accessions with the same name that showed different allelic profiles at several loci), was observed in our sample set. Observed cases of intra-cultivar variability and homonymy are presented in three classes in Table 13-15

Table 13: Cases of intra-cultivar variation; plants with the same name differing in 3-5 out of 26 microsatellite loci
Preglednica 13: Primeri znotraj sortne variabilnosti; rastline z enakim imenom, ki se razlikujejo na do 3-5 lokusih od 26-ih analiziranih

Inta-cultivar variation; plants with different allelic profiles at 3-5 loci	**No**
Lukova E ≠Lukova R	1
Olivastër e Kuqe e Tiranës I-E ≠ Olivastër e Kuqe e Tiranës II-E ≠ Olivastër e Kuqe e Tiranës III-E	2
Ulli I Kuq I-E ≠ Ulli I Kuq II-E ≠ Ulli I kuq UBT –R	3
Ulli I Hollë I Himarës I-E ≠ Ulli I Hollë I Himarës II-E ≠ Ulli I Hollë I Himarës III-E	4
Kushan Kokërr Madh I-E ≠ Kushan Kokërr Madh III-E ≠ (Kushan Kokërr Mesëm I-E =Kushan Kokërr Madh II-E= Frëng I-I)	5
Managjel I-E ≠ Managjel II-E	6
Ulli I Zi I Tiranës I-I ≠ (Ulli I Zi I Tiranës II-I =Ulli I Bardhë I Tiranës III-I)	7
Unafka E ≠ Unafka R	8
Ulli I Bardhë I Tiranës II-I ≠ Ulli I Bardhë I Tiranës V-I	9
Olivastër ILB I-E ≠ Olivastër ILB II-E	10
Mixan I-E ≠ Mixan II-E	11
Kalinjot Vertikal-R ≠ Kalinjot –E	12
Kokërr Madh Berati I-R ≠ Kokërr Madh Berati II-R	13
*Kokërr Madh Berati I ≠ Kryps Berati I	14
Manzanilla E ≠ Manzanilla R	15
Ulli I Bardhë I Krujës I-E ≠ Ulli I Bardhë I Krujës II-E	16

* Kokërr Madh Berati and Kryps Berati are two names known that refer to the same cultivar

Differences in up to 15 % of the analysed alleles between cultivars were considered to be cases of intra-varietal polymorphism (Lopes *at al.,* 2004), while differences in ≥14 loci between accessions can only occur by sexual reproduction (Cipriani *et al.*, 2002) and are considered to be cases of homonymy or misnaming. Alternatively, intra-cultivar diversity could probably be explained by occasional outcrossing events that may have occurred spontaneously between clones and feral forms that are found in the same region (Omrani-Sabbaghi *et al*., 2007). The existence of intra-cultivar variation in olive based on microsatellite markers has been reported by many authors (Cipriani *et al.*, 2002; Lopes *et al.,* 2004; Omrani-Sabbaghi *et al*., 2007; Baldoni *et al*., 2009; Noormohammadi *et al*., 2009; Muzzalupo *et al.,* 2010, 2014; Beghé *et al*. 2011; Koehmstedt *et al.,* 2011).

Table 14: Cases of homonymy; plants with the same name differing in 6-12 out of 26 loci, which are grouped in the same cluster
Preglednica 14: Primeri odkritih homonimov; rastline z istim imenom, ki se razlikujejo na 6-12 lokusih od 26-ih in ki so združene v isti skupini

Homonyms; plants with different allelic profiles at 6-12 loci	**No**
Ulli I Hollë I Himarës I-E, Ulli I Hollë I Himarës II-E,Ulli I Hollë I Himarës III-E ≠ Ulli I Hollë I Himarës IV-E	1
Ulli I Hollë I Himarës I-E, Ulli I Hollë I Himarës II-E,Ulli I Hollë I Himarës III-E ≠ Ulli I Hollë I Himarës I	2
Halneiqis E ≠ Halneiqis R	3
Krypsi I Krujës I-I ≠ Krypsi I Krujës II-I ≠Krypsi I Krujës R	4
Kalinjot E, Kalinjot Vertikal R ≠Kalinjot Vertikal I	5
Ulli I Bardhë I Tiranës IV-I ≠ Ulli I Bardhë I Tiranës V-I	6
Llashtrak I-I ≠ (Llashtrak II-I =Kokërr Madh Elbasani I)	7
Cucurelia E ≠ Cucurelia R	8
Ulli I Zi I Tiranës I-I, (Ulli I Zi I Tiranës II-I=Ulli I Bardhë I Tiranës III-I) ≠ Ulli I Zi I-E	9
(Ulli I Bardhë I Tiranës II-I = Ulli I Bardhë I Tiranës III-I) ≠ Ulli I Bardhë I Tiranës IV-I	10
Olivastër e Kuqe e Tiranës I-E,Olivastër e Kuqe e Tiranës II-E, Olivastër e Kuqe e Tiranës III-E ≠ Olivastër e Kuqe e Tiranës I	11
Mixan I-E, Mixan II-E ≠ Mixan I	12
Boç I ≠Boç R	13
Kushan I-I ≠ Kushan II-I	14
Kalinjot I ≠ Kalinjot R	15

Table 15: Cases of homonymy; plants with the same name differing in 14-24 out of 26 loci
Preglednica 15: Primeri odkritih homonimov; rastline z istim imenom, ki se razlikujejo na 14-24 lokusih od 26-ih analiziranih

Homonyms; plants with different allelic profiles at 14-24 loci	**No**
Vajs Peqini E ≠ Vajs Peqini I	1
Ulli i Zi I-E, Ulli i Zi i Tiranës I-I ≠ Ulli i Zi II-E	2
Ulli i Hollë i Himarës IV-E ≠ Ulli i Hollë i Himarës I	3
Ulli i Bardhë i Tiranës II-I,Ulli i Bardhë i Tiranës IV-I, Ulli i Bardhë i Tiranës V-I ≠ Ulli i Bardhë i Tiranës I-I	4
Frantoio E ≠ Frantoio R	5
Ulli i Zi i Tiranës II-E ≠ (Ulli i Zi i Tiranës II-I=Ulli i Bardhë i Tiranës III-I)	6
Kalinjot Vertikal I, Kalinjot E, Kalinjot Vertikal R , Kalinjot R ≠ Kalinjot Cilindrik I	7
Managjel Formë Veçantë I-E≠ Managjel II-E	8
Managjel II-E, Managjel Formë Veçantë I-E, Managjel I ≠ Managjel R	9
Managjel I-E, Managjel II-E, Managjel Formë Veçantë I-E ≠ Managjel I	10
Kokërr Madh Elbasani II-I ≠ Kryps Elbasani I	11
Koran33 E ≠ Koran69 E	12
Kallmet Kokërr Madh I , Kallmet Kokërr Vogël E ≠ Kallmet Kokërr Vogël I	13
Kallmet Kokërr Madh I, Kallmet Kokërr Vogël I ≠ Kallmet Kokërr Vogël E	14
Karen E ≠ Karen I	15
Kalinjot Vertikal I, Kalinjot E, Kalinjot Vertikal R ≠ Kalinjot I	16
Ulli i Zi I-E, Ulli i Zi i Tiranës II-E, Ulli i Zi i Tiranës I-I, (Ulli i Zi i Tiranës II-I=Ulli i Bardhë i Tiranës III-I) ≠ Ulli i Zi I	17
Coratina I-E ≠ Coratina II-R	18
Nivica E ≠ Nivica R	19
Ulli i Kuq I-E, Ulli i Kuq II-E, Ulli i Kuq UBT R, Ulli i Kuq I-I ≠ Ulli i Kuq II-I	20
Ulli i Kuq UBT R≠ Ulli i Kuq I-I	21
Kalinjot Cilindrik I ≠ Kalinjot I	22
Kalinjot E, Kalinjot Vertikal R, Kalinjot Vertikal I ≠ Kalinjot R	23
Kokërr Madh Berati R, Kokërr Madh Berati III-R ≠ Kokërr Madh Berati I	24
Kokërr Madh Berati I-R, Kokërr Madh Berati II-R ≠ Kokërr Madh Berati III-R	25
Kokërr Madh Berati I-R, Kokërr Madh Berati II-R, Kokërr Madh Berati III-R ≠Kryps Berati I	26
Olivastër Cilindindrike E ≠ Olivastër Cilindindrike I	27
(Olivastër e Zezë Elbasani I-E =Olivastër e Zezë Elbasani II-E) ≠ Olivastër e Zezë Elbasani I	28
Pulazeqin I ≠ Pulazeqin20 E	29
Boç I, Boç R ≠ Boç E	30
Ulli i Bardhë i Tiranës I-I, Ulli i Bardhë i Tiranës IV-I, Ulli i Bardhë i Tiranës II-I, Ulli i Bardhë i Tiranës V-I ≠ Ulli i Bardhë i Tiranës E	31
Kushan Kokërr Madh I-E, (Kushan Kokërr Mesëm I-E=Kushan Kokërr Madh II-E= Frëng I-I), Kushan Kokërr Madh III-E ≠ Kushan I-I	32
Kushan Kokërr Madh I-E, (Kushan Kokërr Mesëm I-E=Kushan Kokërr Madh II-E= Frëng I-I), Kushan Kokërr Madh III-E ≠Kushan II-I	33
Pulazeqin I, Pulazeqin20 E ≠ Pulazeqin R	34
Unafka E, Unafka R ≠ Unafka I	35

Sixteen instances of intra-cultivar variation, genotypes with the same name but differing in 3-5 loci out of 26, (from 3 to 5 alleles) were revealed. The differences at locus GAPU 59 were observed almost at all groups classified as intra-cultivar variability, at which accessions of the same cultivar were differently homozygous. The high value of null allele frequency observed at this locus suggests that these differences could be the result of non-amplification of one of the alleles at each of the accessions of the same cultivar. The cases of intra-cultivar-variability are presented in Table 13.

In 15 cases, accessions with the same name showed a different banding pattern at 6 to 12 microsatellite loci (Table 14). However, they are clustered together, showing a close genetic relationship between them. The clustering together of accessions showing intra-cultivar variability was previously reported by Omrani-Sabbaghi *et al*. (2007) in Iranian olive cultivars.

In view of differences at 14 out of 26 loci, the 35 cases of homonymy given in Table 15 are cases of unclear identity and possible misllabelling or misdenomination. The generic approach to cultivar naming and characterization based only on morphologic traits may be the cause of such a high number of homonymies, stressing the need for further evaluation to determine their identity.

Cases of homonymy in Albanian olive germplasm were previously reported by Belaj *et al*. (2003a). Differentiation between two cases of presumed homonyms Ulli i Bardhë i Tiranës/ Ulli i Bardhë i Beratit and Kallmet/Kallmet Kokërr Vogël was performed by using one or two RAPD markers, suggesting a high level of difference among them. Our results support the reported homonymy by Belaj *et al.* (2003a), because the microsatellite allelic profiles of three accessions of the cultivar 'Kallmet' (Kallmet Kokërr Madh I, Kallmet Kokërr Vogël I/E) in our study differed in 17 to 19 alleles over 26 microsatellite loci.

There were 65 clear cases of mislabelling or misdenomination revealed in our sample set. Six are cases of synonymy, when the same microsatellite profile was found in accessions with different names, 9 are cases of near synonymy and 50 are cases of homonymy. Homonymy cases differing in more than 5-14 loci were considered too variable to the cases of intra-cultivar variation.

5.4.3 Comparison of the Albanian olive sample set with the deposited olive genotypes

The unique genotypes (183) were compared with the microsatellite allelic profiles at 10 microsatellite loci, of 80 olive genotypes of 11 olive-growing countries deposited in the World Olive Database (www. oleadb.it), with the profiles at 26 loci of 19 olive cultivars

from Slovenian collection, for which DNA was provided by the Centre of Plant Biotechnology and Breeding, Biotechnical Faculty, University of Ljubljana and the profiles at 13 microsatellite loci of 306 olive genotypes from 17 olive-growing countries from the Worldwide Olive Germplasm Bank of Córdoba, published by Trujillo *et al.* (2014).

No cases of perfect matches were detected among the allelic profiles of the ten microsatellite loci (DCA 3, DCA 5, DCA 9, DCA 14, DCA 15, DCA 16, DCA 18, GAPU 71B, GAPU 101 and EMO 90) of our native sample genotypes and the genotypes from the olive database.

Four cases of near synonymy, differing ≤ 2 loci were found between Italian genotypes deposited in the World Olive Database and the Albanian genotypes. Frantoio differed by 2 alleles (one at locus DCA14 and the other at locus GAPU101), from 3 Albanian genotypes of Olivastër ILB I-E, Olivastër ILB II-E and Ulli i Bardhë i Tiranës E; Leccino differed by 2 alleles (one at locus DCA14 and one other allele at locus DCA15) from two oleasters, Olivastër Kokërr Gjatë I and Olivastër Cilindrike I; Carolea differed by 2 alleles by from Klon IV-E (one allele at locus DCA16 and the another allele at EMO90) and Ascolana Tenera differed by 2 alleles (one at locus DCA16 and the other at locus GAPU101) from Boç-E. The allelic profile of the Greek cultivar Konservolia deposited in the World Olive Database differed at 2 loci (one allele at locus DCA09 and one allele at locus GAPU71B) with the allelic profile of three genotypes from our sample set, Kokërr Madh Berati-I, Kokërr Madh Elbasani II-I, Sant' Agostino E.

Italian cultivars Frantoio and Lecciono are widespread in Albania, due to their superior agronomical traits and their observed similarity with oleasters suggests that these oleasters probably derived from these two cultivars and are not true oleasters but feral forms. Further investigation would shed more light on this.

Among genotypes with the same name, only one foreign cultivar collected in Albania, Morailo E, showed a perfect match with the allelic profile of the Italian genotype with the same name, Moraiolo, deposited in the World Olive Database. Meanwhile, several cases of subtle allelic differences in 1-2 loci were observed between foreign cultivars collected in Albania and cultivars with the same name deposited in the olive database: Italian cultivars (Carolea –Carolea R, Frantoio-Frantoio-E, Leccino-Lecciono E), Greek cultivars (Mastoidis-Mastoidis E), Spanish cultivars (Picual-Picual R, Manzanilla de Sevilla-Manzanilla R), French cultivars (Picholine-Picholine E). These discrepancies might be due to intra-cultivar variability, which has also been reported for some other well known cultivars obtained from different germplasm collections (Bandelj and Javornik, 2007).

No perfect matches were found between our sample set and the 19 olive cultivars from Slovenian olive collection analysed at 26 microsatellite loci.

A comparison between our sampled genotypes and 306 genotypes of the WOGB of Córdoba published by Trujillo *et al.* (2014) was possible at 13 microsatellite loci (DCA03, DCA05, DCA09, DCA11, DCA15, DCA16, DCA18, GAPU59, GAPU89, GAPU71B, GAPU101, UDO24, and EMO90).

This comparison did not reveal any perfect matches among the SSR profiles of our genotypes and those reported by Trujillo *et al.* (2014). Only one case was found of mismatching at two loci (DCA11 and GAPU71B), between the Albanian cultivar Mixan of the WOGB collection and three of our genotypes: Ulli i Bardhë i Tiranës IV-I, Ulli i Bardhë Pobrati E and Kushan Kokërr Mesëm I-E = Kushan Kokërr Madh III-E = Frëng I-I. The cultivar Mixan from WOGB collection showed homozigosity at locus DCA11, while our genotypes were heterozygous, and at locus GAPU71B there was a difference of only 1bp in the 142 bp allele in Mixan, which was 141 bp in our genotypes, which is probably due to the different allele size rounding.

The identity analysis revealed several cases of synonymy, intra-cultivar variability and homonymy, indicating the high level of varietal disorder in Albanian olive germplasm, which should be re-evaluated and accurately resolved.

The comparative analyses did not find cases of synonymy among the 183 genotyped Albanian cultivars and the 80 genotypes of the World Olive Database, the 19 genotypes of Slovenian collection, or the 306 olive genotypes maintained at the Worldwide Olive Germplasm Bank of Córdoba, thus supporting the hypotheses of an autochthonous origin of Albanian cultivars (Kafazi and Muço, 1984; Belaj *et al.*, 2003a; Ismaili, 2013) and that Albanian olive cultivars represent a unique gene pool.

5.5 PARENTAGE ANALYSIS

Pedigree knowledge about kinship would be of great value for future olive breeding programs, since this knowledge reduces the time and cost of development of new progeny on plants with a long juvenile phase (Bracci *et al*., 2011).

A likehood-based parentage analysis was performed using IDENTITY 1.0 software (Wanger and Sefc, 1999). Several parent-offspring trios were revealed. In all cases, the offspring showed an offspring relation to one of the putative parents. The offspring of the false trios differed by one allele in one or two loci. The similarity between the offspring

and one of the putative parents ranged from 96-98 %. They also clustered close together, thus confirming sibship relationship by UPGMA clustering.

The parentage analysis did not reveal any true parentage relationships. However, it gave useful information about some trios that resulted to be false excluding as well false putative parents.

In two parent-offsprings trios: Boç I=Bahut I x Gjykatës I; Boç I=Gjykatës I x Kalinjot I, the offspring Boç I shared 98 % of similarity with its putative parent Gjykatës I. In three trios; Karen I = Frëng II-I x Mixan I; Karen I=Frëng II-I x Qumështor I; Karen I=Frëng II-I x Ulli i Bardhë i Tiranës II-I, the offspring Karen I shared 98 % of similarity with its putative parent Frëng II-I. In the case of the other two trios: Karen I=Mixan I x Qumështor I; and Karen I=Qumështor I x Ulli I Bardhë I Tiranës II-I, the offspring Karen I shared a similarity of 97 % with Mixan and Ulli i Bardhë i Tiranës II-I. In the case of the trio Llashtrak I-I=Krypsi I Shkodrës I x Ulli I Bardhë i Shkodrës I, the offspring Llashtrak I-I was similar to its putative parent Kryps i Shkodrës I, with 98 % similarity confirming their possible synonymy. In two trios: Kokërr Madh Berati II-R=Kalinjot I x Kokërr Madh Berati I-R; and Kokërr Madh Berati II-R=Kala I x Kokërr Madh Berati I-R, the offspring Kokërr Madh Berati II-R had 94 % similarity to Kokërr Madh Berati I-R. In the final two trios: Klon XI-E=Boç R x Kalinjot Pobrati E; and Klon VIII E=Kalinjot Pobrati E x Klon XI-E, Klon XI-E was 89 % similar to the putative parent Kalinjot Pobrati E and Klon VIII-E was 93 % and 96 % similar to the two parents, the Kalinjot Pobrati E and Klon XI-E, respectively.

Determining the exact pedigrees of olive cultivars is complicated due to the outcrossing nature of this species.

5.6 GENETIC RELATEDNESS

A dendrogram was constructed by using NTSYS v. 2. 2 (Rohlf, 2005) in order to evaluate the genetic relationships among the 183 unique genotypes, using the unweighted pair group (UPGMA) hierarchical clustering method based on the DICE similarity coefficient.

The genetic similarity among the 183 genotypes, based on 26 microsatellite loci, ranged from 0.2 to 1.00, with an average of 0.509. The average level of similarity was higher than the average similarity found in olive cultivars from Slovenian olive collection (0.26) (Bandelj *et al*., 2004) and within the range of mean similarity found in Tunisian olives (0.574) (Rekik *et al.*, 2008).

The genetic relationships among 183 olive genotypes were visualized by a dendrogram (Figure 7). Genotypes clustered into 12 groups, presented in Table 16. Each of the main groups subdivides into several small clusters consisting of related genotypes. In many cases, the high level of genetic similarity between genotypes of the same subcluster (> 85 %) provides confirmation of the high number of synonyms and the intra-cultivar variability revealed in the identity analysis.

The dendrogram showed that native Albanian genotypes were in close relationship with each other. Clusters IV, V and VI contained 99 (95 %) of Albanian genotypes and 5 (5 %) foreign cultivars (4 Italian and 1 USA). The other Albanian genotypes were dispersed within other clusters that also contained foreign cultivars. Clusters II and IX contained mainly foreign cultivars, 15 (75 %) and 13 (62 %), respectively (Table 16 and Figure 7). The tight clustering of Albanian genotypes was also observed even when in the UPGMA analysis were considered also olive cultivars from the Slovenian collection (Annex F). Tight clustering and a clear distinction of Albanian cultivars from foreign ones (Greek, Italian and Turkish) was also observed by Belaj *et al*. (2003a), indicating their autothoctonous origin.

Five genotypes with the same name, Ulli i Hollë i Himarës I/ I-E/ II-E/ III-E/ IV-E collected from the *ex situ* collection and Ulli i Hollë i Himarës I collected *in situ* in the region of Himara in the southern part of the country, one Greek cultivar (Mastoidis) and an oleaster genotype from the central region of the country (Olivastër Kokërr Hollë Elbasani-E) clustered together in cluster XII; all these cultivars are used for oil production. The Greek cultivar Mastoidis has 46 % genetic similarity to the cultivars of this group. This level of similarity between the Greek cultivar and 'Ulli i hollë i Himarës' suggests possible common ancestors. Belaj *et al*. (2003a) also observed the grouping of this native cultivar to Greek olive cultivars.

There is a clear structuring of the variability relative to the geographic origin/main distribution area of Albanian native genotypes. Genotypes from central and northern regions (Adriatic region) tend to be grouped in distinct clusters from those from southern Albania (Ionic region). The Adriatic genotypes grouped mainly in cluster IV (subcluster I) 93.6 % (44 genotypes out of 47), cluster V 90 % (18 genotypes out 20) and cluster VI 83.3 % (10 Adriatic genotypes out of 12). The majority of Ionic genotypes, on the other hand, were in cluster V (subcluster II) 80 % (20 out of 25) and cluster XII 71.4 % (5 out of 7 cultivars). The obtained clustering may also be due to selection pressure towards cultivars that are adapted to the different climatic conditions of these two regions (Belaj *et al*., 2003a).

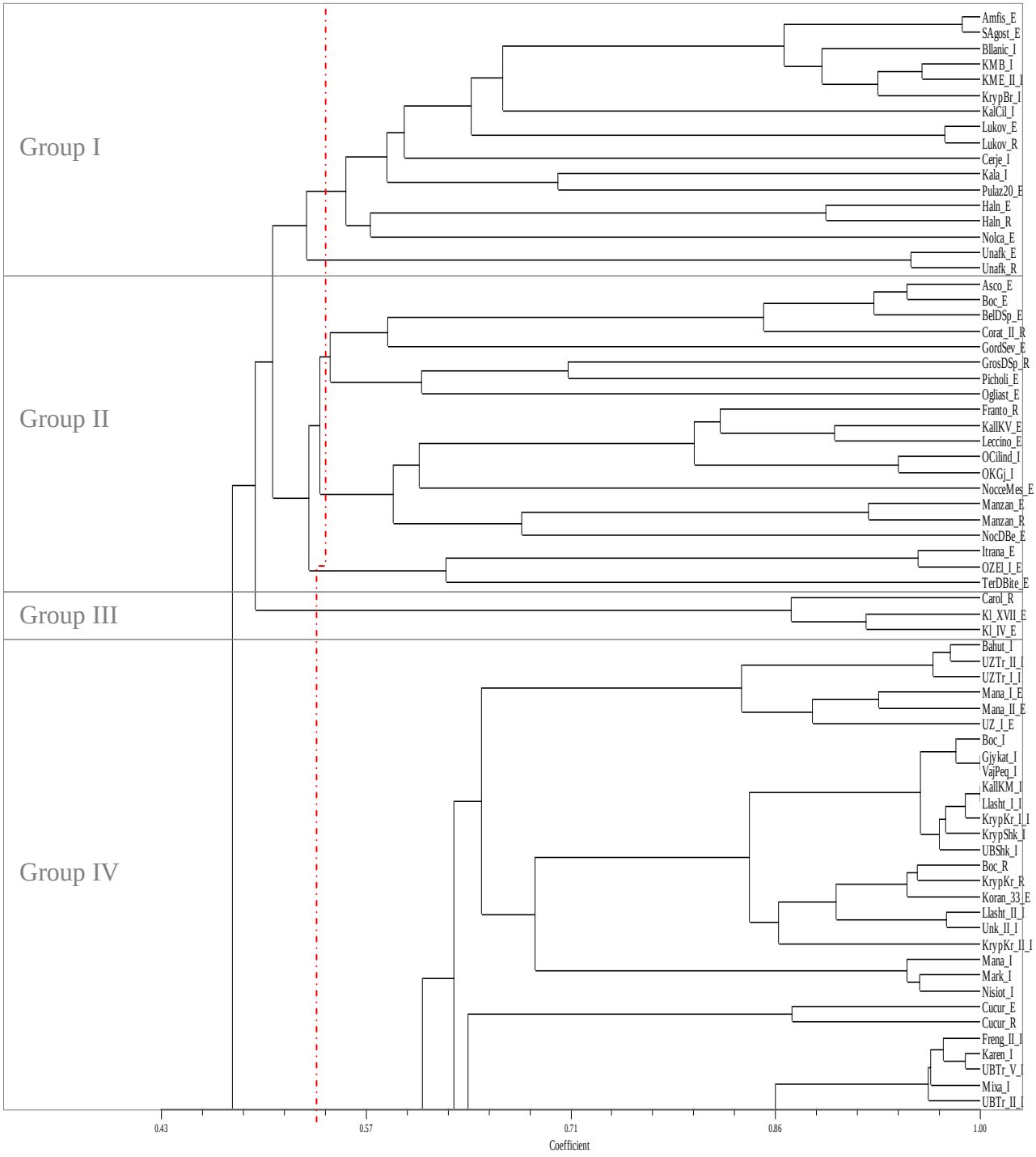
The area of origin can be identified because of the climatic and pedological conditions under which populations experienced a secondary structuration of variability.

Distinguishing among varieties belonging to different gene pools would enable selection of the most adapted cultivars for the new breeding policies for each geographically defined area (Sarri *et al*., 2006). The olive distribution area in Albania is divided into two main regions, based on the correlation of morphological diversity with the climatic conditions and geographical elements of these areas. The olive germplasm is thus distributed into the southern-western area, under the influence of the Ionian Sea (Ionic region) and the central and northen areas under the influence of the Adriatic Sea (Adriatic region) (Ismaili *et al*., 2013)

Weak clustering relative to the fruit end-use was observed in overall sample set, in line with the findings of Abdelhamid *et al*. (2013). The majority of genotypes used as table olives and dual use were mainly mixed into clusters I, II and IV (subcluster II). The majority of genotypes used for oil production were grouped in clusters IV (subcluster I), V, VI and XII. Only native genotypes, especially those from the central region, showed fairly good clustering according to their fruit end-use (Table 16).

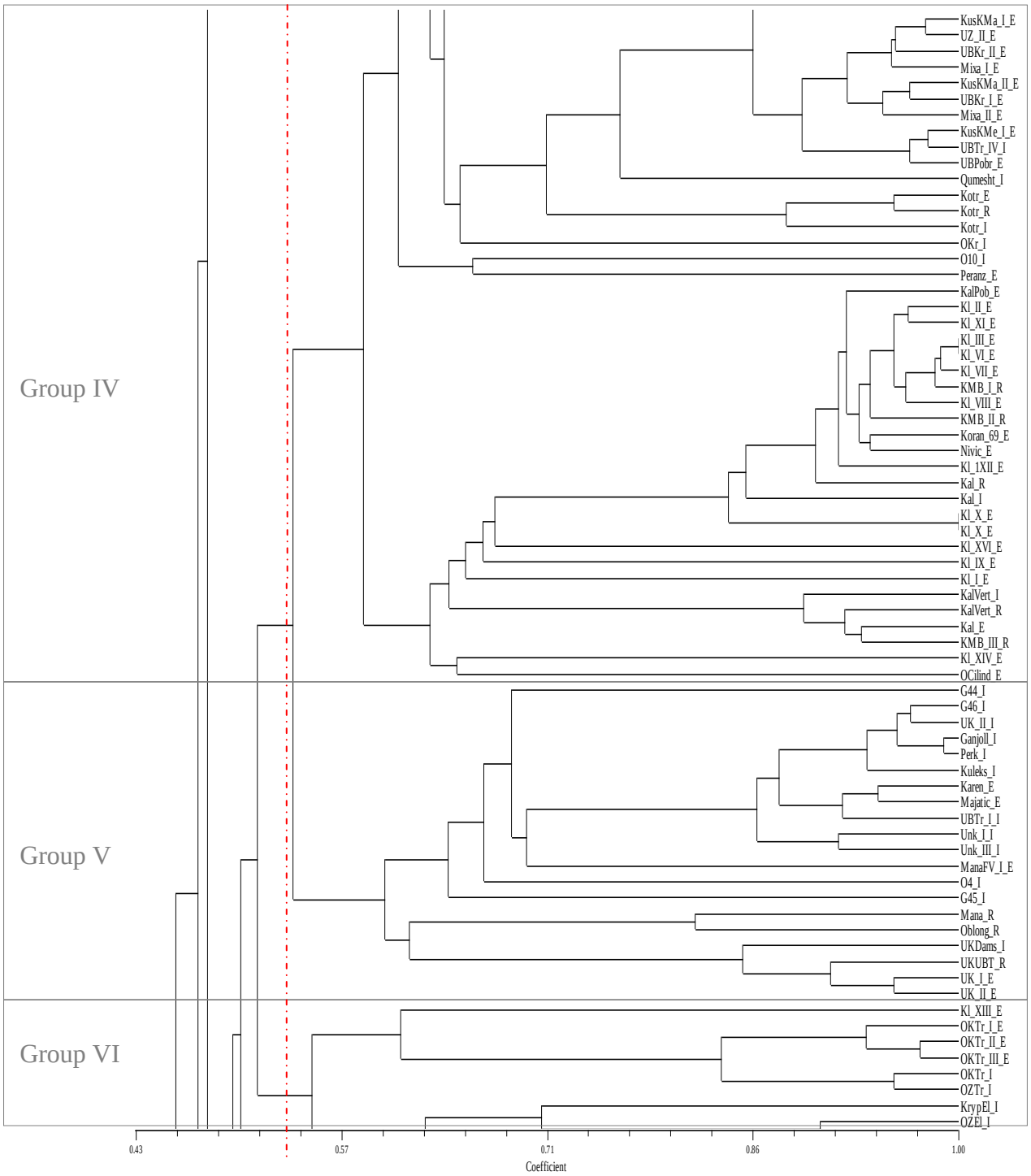
Only five oleasters out of 18 in cluster VI were subclustered distinctly (Olivastër e Kuqe e Tiranës I-E, Olivastër e Kuqe e Tiranës II-E, Olivaster e Kuqe e Tiranës III-E, Olivastër e Kuqe e Tiranës I, Olivastër e Zezë e Tiranës I) (Figure 7), suggesting a genuine originality of these oleaster genotypes. The other oleasters clustered together with Albanian cultivated forms, especially with foreign cultivars, in the different groups Olivastër e Zezë Elbasani I-E, Olivastër Cilindrike I, Olivastër Kokërr Gjatë I and) in cluster II, (Olivastër10 I) in cluster VI, (Olivastër ILB-I E and Olivastër ILB-II E) in cluster IX and (Olivastër Kokërr Hollë Elbasani E) in cluster XII. These results support two hypotheses: 1) oleasters grouped in the same cluster with cultivated forms may be feral olives and are showing high similarity with the cultivated genotypes from which they derive or 2) according to the second hypothesis, this may be a case of clustering of true oleasters together with cultivated forms that derived from them.

No clear distinction of cultivated olives and oleasters was observed, which is in accordance with previous studies (Belaj *et al*., 2010), indicating the presence of feral forms in our sample set. However, the co-existence of true oleasters and feral forms should not be excluded. It can be concluded that analysis of a larger number of oleasters, including oleasters from neighbouring countries, would contribute to a better understanding of origin and distribution in Albanian olives.


Group I
Group II
Group III
Group IV
Amfis_E
SAgost_E
Bllanic_I
KMB_I
KME_II_I
KrypBr_I
KalCil_I
Lukov_E
Lukov_R
Cerje_I
Kala_I
Pulaz20_E
Haln_E
Haln_R
Nolca_E
Unafk_E
Unafk_R
Asco_E
Boc_E
BelDSp_E
Corat_II_R
GordSev_E
GrosDSp_R
Picholi_E
Ogliast_E
Franto_R
KallKV_E
Leccino_E
OCilind_I
OKGj_I
NocceMes_E
Manzan_E
Manzan_R
NocDBe_E
Itrana_E
OZEl_I_E
TerDBite_E
Carol_R
Kl_XVII_E
Kl_IV_E
Bahut_I
UZTr_II_I
UZTr_I_I
Mana_I_E
Mana_II_E
UZ_I_E
Boc_I
Gjykat_I
VajPeq_I
KallKM_I
Llasht_I_I
KrypKr_I_I
KrypShk_I
UBShk_I
Boc_R
KrypKr_R
Koran_33_E
Llasht_II_I
Unk_II_I
KrypKr_II_I
Mana_I
Mark_I
Nisiot_I
Cucur_E
Cucur_R
Freng_II_I
Karen_I
UBTr_V_I
Mixa_I
UBTr_II_I
0.43
0.57
0.71
0.86
1.00
Coefficient

Continued

Continuation of Figure 7: Dendrogram of 183 olive genotypes based on Dice coefficient and the UPGMA clustering method



Continued

Continuation of Figure 7: Dendrogram of 183 olive genotypes based on Dice coefficient and the UPGMA clustering method

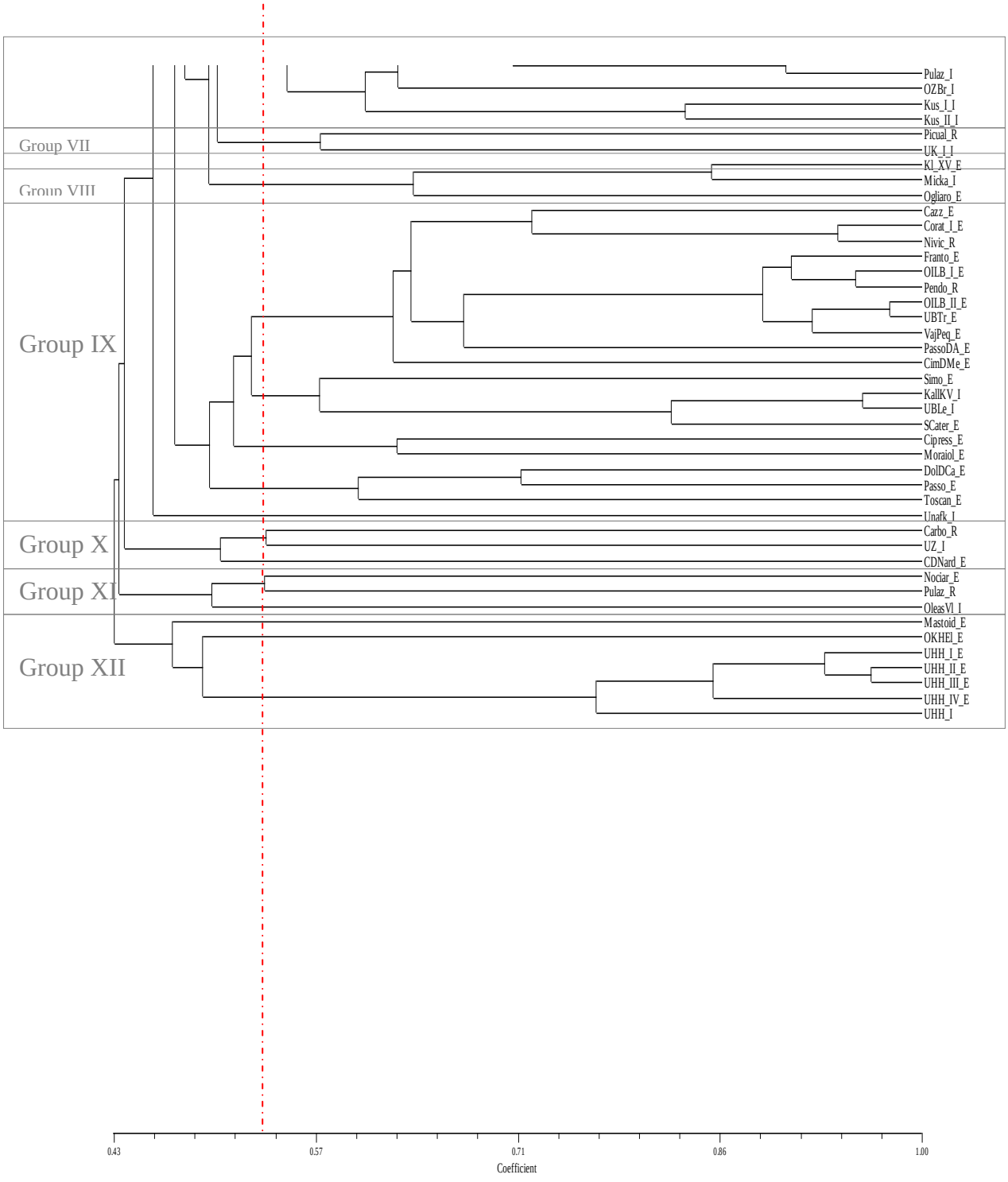


Figure 7: Dendrogram of 183 olive genotypes based on Dice coefficient and the UPGMA clustering method
Slika 7: Dendrogram analize genetske sorodnosti 183-ih genotipov oljk, ki temelji na razdalji koeficineta Dice in UPGMA metodi razvrščanja v skupine

Table 16: Genetic relationships of 183 olive genotypes clustered according to a UPGMA dendrogram
Preglednica 16: Prikaz genetske sorodnosti in razvrščanja v skupine za 183 genotipov oljk glede na UPGMA dendrogram

Cluster	Genotype	Abbreviation	End-use	Geographic region
	Amfisse E	Amfis_E	T	Greece
	Sant Agostino E	Sagost_E	T	Italy
	Bllanic I	Bllanic_I	D	Ionic
	Kokërr Madh Berati I	KMB_I	T	Adriatic*
	Kokërr Madh Elbasani II-I	KME_II_I	T	Adriatic
	Krypsi I Beratit I	KrypBr_I	T	Adriatic
	Kalinjot Cilindrik I	KalCil_I	D	Ionic*
Cluster I	Lukova E	Lukov_E	D/O	Ionic
	Lukova R	Lukov_R	D/O	Ionic
	Cerje I	Cerje_I	T	Ionic
	Kala I	Kala_I	D	Ionic
	Pulazeqin 20 E	Pulaz20_E	O	Ionic
	Halneiqis E	Haln_E	T	Greece
	Halneiqis R	Haln_R	T	Greece
	Nolca E	Nolca_E	T	Italy
	Unafka E	Unafka_E	O	Adriatic
	Unafka R	Unafka_R	O	Adriatic
	Ascolana E	Asco_E	T	Italy
	Boç E	Boç_E	D	Adriatic
	Bella di Spagna E	BelDSp_E	T	Spain
	Coratina II-R	Corat_II-R	O	Italy
	Gordal Sevillana E	GordSev_E	T	Spain
	Grossa di Spagna R	GrosDSp_R	T	Italy
	Picholine E	Picholi_E	D	France
	Ogliastro E	Ogliast_E	O	Italy
	Frantoio R	Franto_R	O	Italy
Cluster II	Kallmet Kokërr Vogël E	KallKV_E	O	Adriatic
	Leccino E	Leccino_E	O	Italy
	Olivastër cilindrike I	Ocilind_I	O	Adriatic
	Olivastër Kokërr Gjatë I	OKGj_I	O	Adriatic
	Noccellara Messinese E	NocceMes_E	D	Italy
	Manzanilla E	Manzan_E	D	Spain
	Manzanilla R	Manzan_R	D	Spain
	Nocellara Dell Bellice E	NocDBe_E	T	Italy
	Itrana E	Itrana_E	D	Italy
	Olivastër e Zezë Ebasani I-E	OZEl_I_E	O	Adriatic

Continued

Continuation of Table 16: Genetic relationships of 183 olive genotypes clustered according to a UPGMA dendrogram

Cluster	Genotype	Abbreviation	End-use	Geographic region
	Termite di Bitteto E	TerDBite_E	D	Italy
	Carolea R	Carol_R	D	Italy
Cluster III	Klon XVII-E	Kl_XVII_E	D	Ionic
	Klon IV-E	Kl_IV_E	D	Ionic
	Bahuta I	Bahut_I	O	Adriatic
	Ulli I Zi I Tiranës II-I	UZTr_II_I	O	Adriatic
	Ulli I Zi I Tiranës I-I	UZTr_I_I	O	Adriatic
	Managjel I-E	Mana_I_E	D	Adriatic
	Managjel II-E	Mana_II_E	D	Adriatic
	Ulli I Zi I-E	UZ_I_E	O	Adriatic
	Boç-I	Boc_I	D	Adriatic
	Gjykates I	Gjykat_I	D	Adriatic
	Vajs Peqini I	VajPeq_I	O	Adriatic
	Kallmet Kokërr Madh I	KallKM_I	D	Adriatic
	Llashtrak I-I	Llasht_I_I	T	Adriatic
	Krypsi I Krujës I-I	KrypKr_I_I	D	Adriatic
	Krypsi I Shkodrës I	KrypShk_I	D	Adriatic
	Ulli I Bardhë I Shkodrës I	UBShk_I	O	Adriatic
	Boç R	Boc_R	D	Adriatic
	Krypsi I Krujës R	KrypKr_R	D	Adriatic
	Koran 33 E	Koran_33_E	O	Adriatic
	Llashtrak II-I	Llasht_II_I	T	Adriatic
	Unknown II-I	Unk_II_I	D	Adriatic
	Krypsi I Krujës II-I	KrypKr_II_I	D	Adriatic
	Managjel I	Mana_I	D	Adriatic
	Marks I	Mark_I	D	Adriatic
	Nisiot I	Nisiot_I	D	Adriatic
	Cucurelia E	Cucur_E	D	Italy
	Cucurelia R	Cucur_R	D	Italy
	Frëng II-I	Freng_II_I	O	Adriatic
	Karen I	Karen_I	O	Adriatic
	Ulli I Bardhë I Tiranës V-I	UBTr_V_I	O	Adriatic
	Mixan I	Mixa_I	O	Adriatic
	Ulli I Bardhë I Tiranës II-I	UBTr_II_I	O	Adriatic
	Kushan Kokërr Madh I-E	KusKMa_I_E	O	Adriatic
	Ulli I Zi II-E	UZ_II_E	O	Adriatic
	Ulli I Bardhë I Krujës II-E	UBKr_II_E	O	Adriatic

Continued

Continuation of Table16: Genetic relationships of 183 olive genotypes clustered according to a UPGMA dendrogram

Cluster	Genotype	Abbreviation	End-use	Geographic region
	Mixan I-E	Mixa_I_E	O	Adriatic
	Kushan Kokërr Madh III-E	KusKMa_III_E	O	Adriatic
Cluster IV	Ulli I Bardhë I Krujës I-E	UBKr_I_E	O	Adriatic
	Mixan II-E	Mixa_II_E	O	Adriatic
	Kushan Kokërr Mesëm I-E	KusKMe_I_E	O	Adriatic
	Ulli I Bardhë I Tiranës IV-I	UBTr_IV_I	O	Adriatic
	Ulli I Bardhë Pobrati E	UBPobr_E	O	Adriatic
	Qumështor I	Qumesht_I	D	Adriatic
	Kotruvs E	Kotr_E	O	Adriatic
	Kotruvs R	Kotr_R	O	Adriatic
	Kotruvs I	Kotr_I	O	Adriatic
	Olivastër e Krujës I	Okr_I	O	Adriatic
	Olivastër 10-I	O10_I	O	Adriatic
	Peranzana E	Peranz_E	D	Italy
	Kalinjot Pobrat E	KalPob_E	D	Ionic
	Klon II-E	Kl_II_E	D	Ionic
	Klon XI-E	Kl_XI_E	D	Ionic
	Klon III-E	Kl_III_E	D	Ionic
	Klon VI-E	Kl_VI_E	D	Ionic
	Klon VII-E	Kl_VII_E	D	Ionic
	Kokërr Madh Berati I-R	KMB_I_R	T	Adriatic
	Klon VIII-E	Kl_VIII_E	D	Ionic
	Kokërr Madh Berati II-R	KMB_II_R	T	Adriatic
	Koran 69-E	Koran_69_E	O	Adriatic
	Nivica-E	Nivic_E	O	Ionic
	Klon XII-E	Kl_XII_E	D	Ionic
	Kalinjot-R	Kal_R	D	Ionic
	Kalinjot I	Kal_I	D	Ionic
	Klon V-E	Kl_V_E	D	Ionic
	Klon X-E	Kl_X_E	D	Ionic
	Klon XVI-E	Kl_XVI_E	D	Ionic
	Klon IX-E	Kl_IX_E	D	Ionic
	Klon I-E	Kl_I_E	D	Ionic
	Kalinjot Vertikal I	KalVert_I	D	Ionic
	Kalinjot Vertikal R	KalVert_R	D	Ionic
	Kalinjot E	Kal_E	D	Ionic
	Kokërr Madh Berati III-R	KMB_III_R	O	Adriatic

Continued

Continuation of Table16: Genetic relationships of 183 olive genotypes clustered according to a UPGMA dendrogram

Cluster	Genotype	Abbreviation	End-use	Geographic region
	Klon XIV-E	Kl_XIV_E	D	Ionic
	Olivastër cilindrike E	Ocilind_E	O	Adriatic
	G44-I	G44_I	O	Adriatic
	G46-I	G46_I	O	Adriatic
	Ulli I Kuq II-I	UK_II_I	O	Adriatic
	Ganjoll I	Ganjoll_I	D	Adriatic
	Perk I	Perk_I	D	Adriatic
	Kuleks I	Kuleks_I	O	Adriatic
	Karen E	Karen_E	O	Adriatic
	Majatica E	Majatic_E	D/T	Italy
	Ulli I Bardhë I Tiranës I-I	UBTr_I_I	O	Adriatic
Cluster V	Unknown I-I	Unk_I_I	D	Adriatic
	Unknown III-I	Unk_III_I	D	Adriatic
	Managjel Formë Veçantë I-E	ManaFV_I_E	D	Adriatic
	Olivaster 4-I	O4_I	O	Adriatic
	G45-I	G45_I	O	Adriatic
	Managjel R	Mana_R	D	Adriatic
	Oblonga R	Oblong_R	O	USA
	Ulli I Kuq I Damsit I	UKDams_I	O	Adriatic
	Ulli I Kuq UBT R	UKUBT_R	O	Adriatic
	Ulli I Kuq II-E	UK_I_E	O	Adriatic
	Ulli I Kuq II-E	UK_II_E	O	Adriatic
	Klon XIII-E	Kl_XIII _E	D	Ionic
	Olivastër e Kuqe e Tiranës I-E	OKTr_I_E	O	Adriatic
	Olivastër e Kuqe e Tiranës II-E	OKTr_II_E	O	Adriatic
	Olivastër e Kuqe e Tiranës III-E	OKTr_III_E	O	Adriatic
	Olivastër e Kuqe e Tiranës I	OKTr_I	O	Adriatic
Cluster VI	Olivastër e Zezë e Tiranës I	OZTr_I	O	Adriatic
	Krypsi i Elbasanit I	KrypEl_I	D	Adriatic
	Olivastër e Zezë Elbasani I	OZEl_I	O	Adriatic
	Pulazeqin I	Pulaz_I	O	Ionic
	Olivastër e Zezë Berati I	OZBr_I	O	Adriatic
	Kushan I-I	Kus_I_I	O	Adriatic
	Kushan II-I	Kus_II_I	O	Adriatic
	Picual R	Picual_R	O	Spain
Cluster VII	Ulli i Kuq I-I	UK_I_I	O	Adriatic

Continued

Continuation of Table16: Genetic relationships of 183 olive genotypes clustered according to a UPGMA dendrogram

Cluster	Genotype	Abbreviation	End-use	Geographic region
	Klon XV-E	Kl_XV_E	D	Ionic
Cluster VIII	Micka I	Micka_I	O	Adriatic
	Ogliarola E	Ogliaro_E	O	Italy
	Cazzinicchio E	Cazz_E	D/O	Italy
	Coratina I-E	Corat_I_E	O	Italy
	Nivica R	Nivic_R	D/O	Ionic
	Frantoio E	Franto_E	O	Italy
	Olivastër ILB I-E	OILB_I_E	O	Adriatic
	Pendolino R	Pendo_R	O	Italy
	Olivastër ILB II-E	OILB_II_E	O	Adriatic
	Ulli I Bardhë I Tiranës E	UBTr_E	O	Adriatic
	Vajs Peqini E	VajPeq_E	O	Adriatic
	Passola di Andria E	PassoDA_E	T	Italy
Cluster IX	Cima di Melfi E	CimDMe_E	O	Italy
	Simona E	Simo_E	O	Italy
	Kallmet Kokërr Vogël I	KallKV_I	O	Adriatic
	Ulli I Bardhë I Lezhës I	UBLe_I	O	Adriatic
	Santa Caterina E	Scater_E	T	Italy
	Cipressino E	Cipress_E	O	Italy
	Moraiolo E	Moraiol_E	O	Italy
	Dolce di Cassano E	DolDCa_E	O	Italy
	Passola E	Passo_E	T	Italy
	Toscanina E	Toscan_E	O	Italy
	Unafka I	Unafk_I	O	Adriatic
	Carboncella R	Carbo_R	O	Italy
Cluster X	Ulli I Zi I	UZ_I	O	Adriatic
	Cellina di Nardo E	CDNard_E	D	Italy
	Nocciara E	Nociar_E	D	Italy
Cluster XI	Pulazeqin R	Pulaz_R	O	Ionic
	Oleastër e Vlorës I	OleasVI_I	O	Ionic
	Mastoidis E	Mastoid_E	O	Greece
	Olivastër Kokërr Hollë Elbasani E	OKHEl_E	O	Adriatic
	Ulli I Hollë I Himarës I-E	UHH_I_E	O	Ionic
Cluster XII	Ulli I Hollë I Himarës II-E	UHH_II_E	O	Ionic
	Ulli I Hollë I Himarës III-E	UHH_III_E	O	Ionic
	Ulli I Hollë I Himarës IV-E	UHH_IV_E	O	Ionic
	Ulli I Hollë I Himarës I	UHH_I	O	Ionic

*Olive samples collected in Albania are presented according to their area of origin/main distribution into Adriatic and Ionic regions

5.7 PRINCIPAL COORDINATE ANALYSIS AND AMOVA

Principal coordinate analysis (PCoA) available in GenAlEx v 6.501 software (Peakall and Smouse 2006; 2012) was performed on genetic distance based matrix on the complete data set of 183 unique olive genotypes, and on specific groups of olive genotypes defined by their origin/main area of distribution, breeding (cultivars *vs* oleasters) and by their product end-use. PCoA plots shows graphically the relationship between individuals/groups of olives and determines whether partititioning into these groups is supported by genetic variation.

Principal coordinate analysis of the 183 olive genotypes generated a total variation of 23.21 %; the first and the second principal coordinates explained 9.22 and 7.98 of genetic variation, respectively (Figure 8).

Only a few individual genotypes were separated from the others, such as, 'Grosso di Spagna'; three accessions of 'Ulli i Hollë i Himarës' (UHH I-E, UHH II-E, and UHH III-E), which were also clustered together in the UPGMA phenogram; while the two unknown olive genotypes from the northern region of Albania (G44 and G45); and two clones of the cultivar 'Kalinjot' ('Klon IX' and 'Klon X'), which were not closely clustered into the phenogram but still present in the same big group; five and four, respectively.

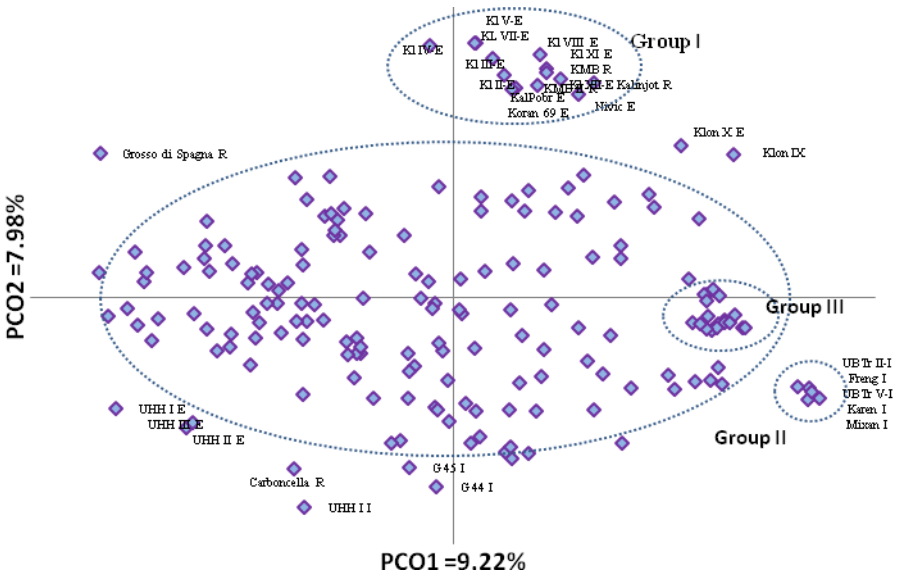


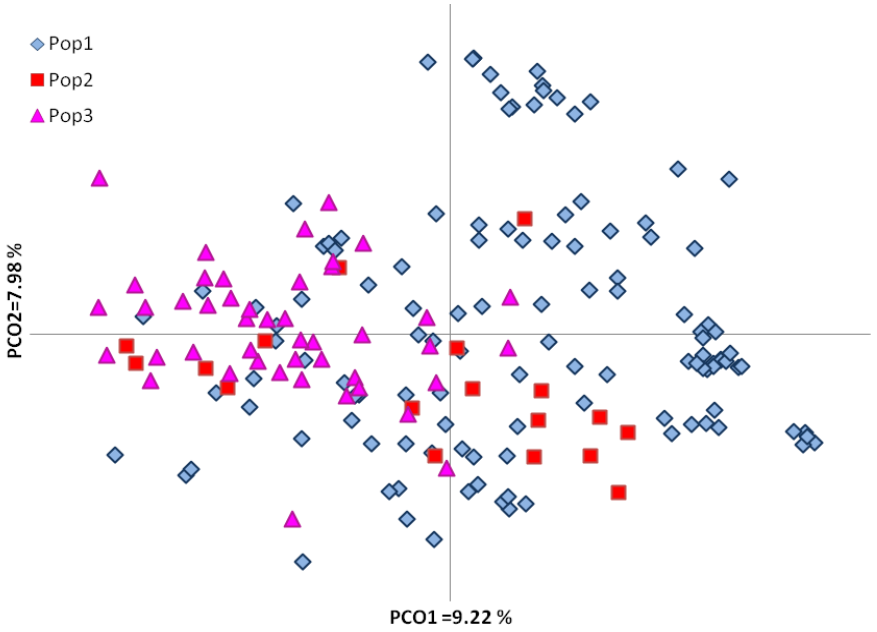
Figure 8: Principal coordinate analysis of 183 olive genotypes based on SSR data. The first two principal coordinates explained 9.22 and 7.98 % of the variance, respectively

Slika 8: Analiza glavnih koordinat (PCoA) narejena na osnovi mikrosatelitnih podatkov za 183 genotipov oljk. Prvi dve glavni koordinati pojasnita 9.22 in 7.98 % raznolikosti.

Three small groups of genotypes of high genetic similarity were distinguished in the PCoA scatter plot (Figure 8). These three groups were also observed in the UPGMA cluster analysis based on Dice's similarity coefficient. The first group is composed mainly of clones of cultivar Kalinjot grown in the South of the country, which are used for oil production. The second group is comprised of olives genotypes from middle of the country used for oil production and in the third group 13 olive genotypes are clustered, they are grown in the middle-north area of Albania, the majority of which are used for oil production.

A high level of overlapping was shown in the PCoA scatter plot of the analysis of the entire data set (Figure 8). The rest of olive genotypes were not clearly separated by PCoA, which is in agreement with the high number of synonyms and near synonyms revealed by identity analysis. This admixture also supports the high number of small clusters among closely related genotypes observed in the UPGMA based phenogram (Figure 7).

PCoA analysis was also performed for different groups defined according to their geographical origin, (native vs foreign genotypes and Ionic region *vs* Adriatic region genotypes), and their breeding (cultivars *vs* oleaster; ancient cultivars *vs* oleasters, foreign cultivars *vs* native cultivars and oleasters) (Figures 9 and 10).



*Pop 1-Native cultivars, Pop 2-Oleasters, Pop 3- Foreign imported cultivars

Figure 9: Principal coordinate analysis of olive genotypes based on their origin and breeding. The first two principal coordinates explained 9.22 and 7.98 % of the variance, respectively
Slika 9: Analiza glavnih koordinat (PCoA), ki temelji na izvoru posameznih genotipov. Prvi dve glavni koordinati pojasnita 9.22 in 7.98 % raznolikosti

In order to investigate the genetic partitioning and relationship among the different olive groups, (1) native cultivars (2) oleasters and (3) foreign cultivars, principal coordinate analysis of these three groups was performed (Figure 10).

The PCoA explained a total variation of 23.21 % (Figure 9). The first principal coordinate accounted for 9.22 % of the total variation and allowed the discrimination of the majority of foreign olive cultivars from native olive cultivars and oleasters. The second coordinate, which accounted for 7.98 % of the total variation, differentiated the majority of the oleasters. There was no clear separation of oleasters from native cultivated cultivars or foreign ones. The same was observed in the phenogram (Figure 7), in which some oleasters were closely clustered with cultivated olive cultivars.

PCoA in pairwaised comparison of olive groups (Figure 10) revealed a clear differentiation between foreign cultivars and the group of oleasters (Figure 10-B), the differentiation was not as clear between the groups of foreign cultivars and native cultivated olives (Figure 10-A). Oleaster genotypes were admixed much more with the ancient than with the current olive cultivars (Figure 10-C and Figure10-D), which might be due to the age of ancient cultivars, which are representative of ancient domestication events (Díez *et al.*, 2011) or the admixed oleaster genotypes might have been derived from ancient olive cultivars.

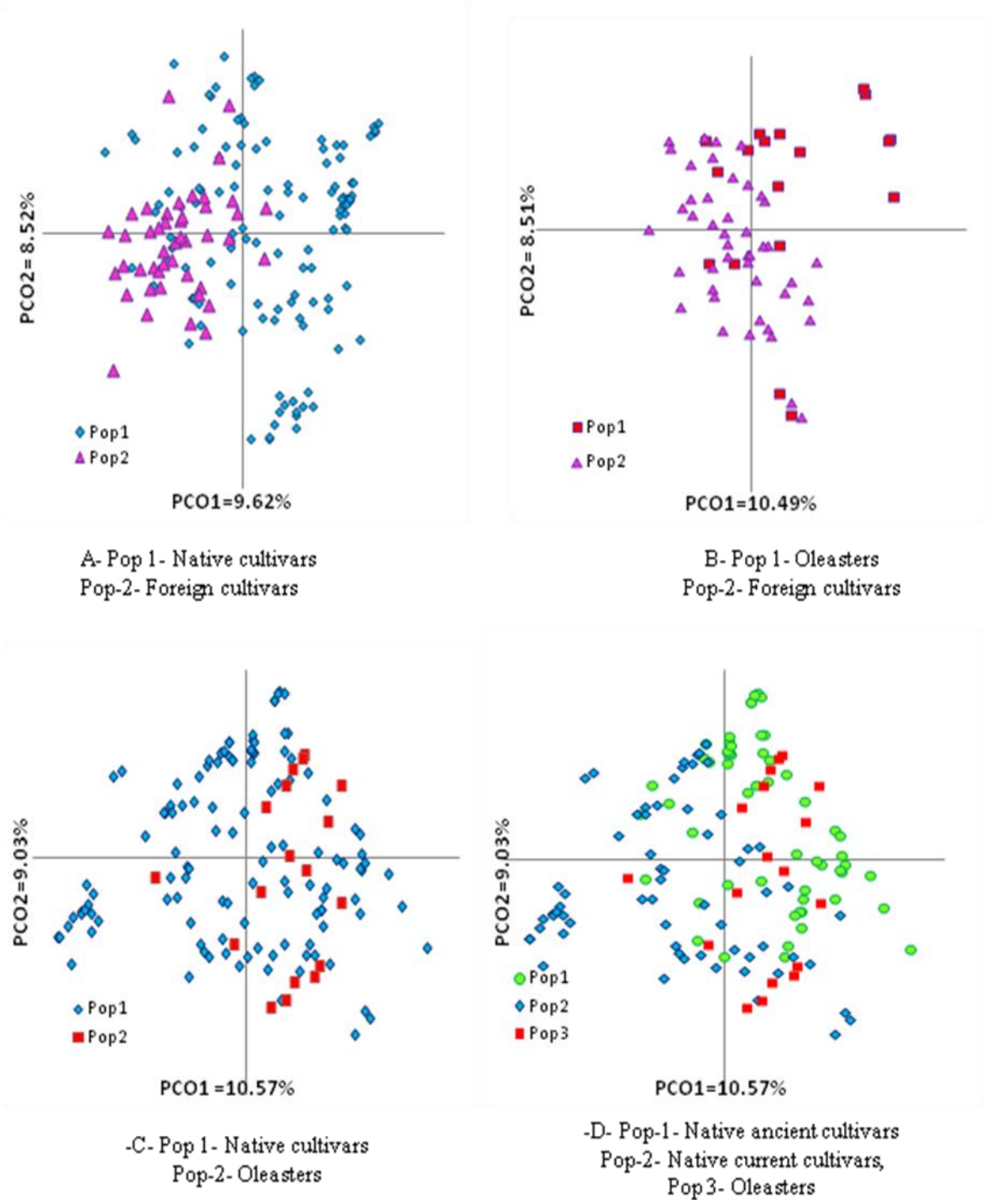


Figure 10: Principal coordinate analysis of olive genotypes based on their origin (-A & -B) and breeding (-C &-D)
Slika 10: Analiza glavnih koordinat (PCoA), ki temelji na izvoru (-A & -B) in žlahtnjenju (-C &-D)

Analysis of molecular variance (AMOVA) was performed to evaluate the partitioning of molecular variance among native cultivars (current and ancient cultivars), oleasters and foreign olive cultivars.

AMOVA showed that 93 % of variation is due to genetic variation within groups, with only 7 % of genetic variance being observed among groups, indicating the high heterozygous nature and mixed genetic structure of the olive genotypes (do Val *et al*., 2012). The calculated PhiPT (analogue of Fst index) (0.074) was significant $P < 0.001$, indicating low genetic differentiation among groups. The P values were calculated for a random permutation test of 9999 permutations (Table 17).

Table17: Analysis of variance of oleasters, native and foreign cultivars
Preglednica 17: Analiza variance vzorcev divije oljke ter domačih in tuiih kultivariev

Source	**Df**	**SS**	**MS**	**Est. Var.**	**%**
Among Pops	2	157	78	1	7
Within Pops	180	3026	17	17	93
Total	182	3183		18	100
PhiPT	**0,074**		**P=0.0001**		

PhiPT = AP/(WP+AP)=AP/Tot; AP = Est.var. among groups;WP = Est. var. within groups

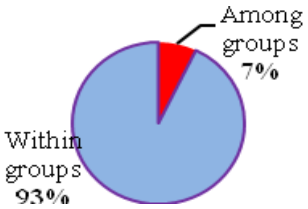


Table18: Pairwaise population matrix of Nei's genetic distance between native olive cultivars, foreign cultivars and oleasters
Preglednica 18: Populacijska matrika genetskih razdalj po Nei-u med pari obravnavanih vzorcev (med domačimi in tujimi kultivarji ter divjimi oljkami)

	*Native cultivars	Oleasters	Foreign cultivars
Native cultivars	0.000		
Oleasters	0.106	0.000	
Foreign cultivars	0.092	0.127	0.000

*Native cultivars (current and acient cultivars)

Based on a pairwise population matrix of Nei's genetic distances (Table 18), the highest genetic distance was observed between foreign cultivars and oleasters (0.127) and the lowest between native and foreign cultivars (0.092). Oleasters were shown to be closer to native cultivars (GD=0.106) than foreign cultivars (GD=0.127), suggesting that domestication events may have involved native oleasters.

Table19: Pairwaise population matrix of Nei's genetic distance between native ancient, current cultivated olive cultivars , foreign cultivars and oleasters
Preglednica 19: Populacijska matrika genetskih razdalj po Nei-u med pari obravnavanih vzorcev (med domačimi/starimi sortami, novejšimi kultivarji, tujimi kultivarji ter divjimi oljkami)

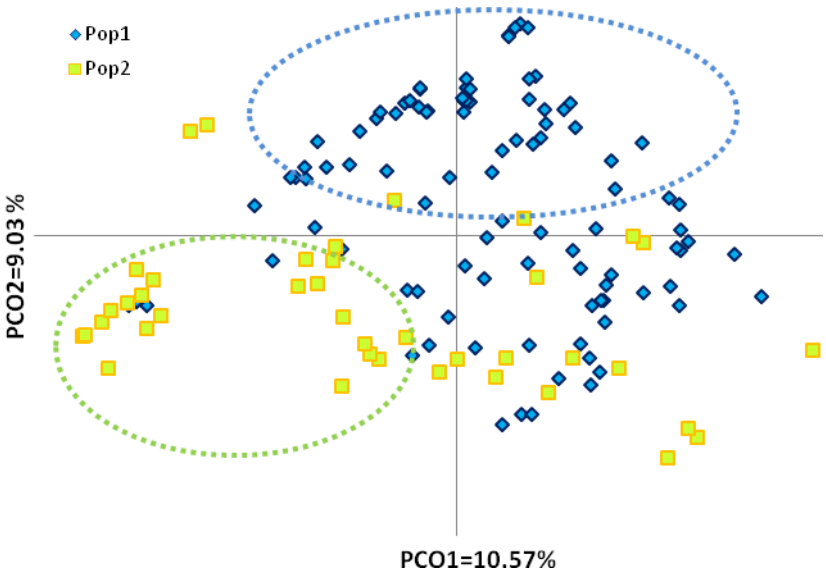
	Ancient	*Current	Oleaster	Foreign
Ancient	0.000			
Current	0.116	0.000		
Oleaster	0.135	0.133	0.000	
Foreign	0.163	0.086	0.127	0.000

*Current cultivated genotypes

It has been suggested that ancient cultivars are of great value for understanding domestication events, due to their great age. Nei's genetic distance was calculated among the group of native ancient cultivars, current cultivated olive cultivars, oleasters and foreign cultivars (Table 19). The highest genetic distance was observed between ancient cultivars and cultivars of foreign origin (0.163). Ancient cultivars were closer to current cultivated olives (0.116) than to oleasters (0.135). The oleaster genotypes were at approximately the same distance from ancient (0.135), current (0.133) and foreign cultivars (0.127), which suggest the common origin of cultivated olives and oleasters. The low genetic distance of native current cultivated olives with foreign cultivars compared with ancient ones suggests that some of the native current olive cultivars may be derived from intercrossing of native and introduced foreign cultivars (Besnard *et al*., 2001a).

Another analysis of principal coordinates based on assignment of two geographical groups according to their main area of distribution/origin (Ionic *vs* Adriatic region) explained a total of 26.30 % of genetic variation. The first two principal coordinates accounted for 10.57 and 9.03 of variation, respectively (Figure 11). The slight overlapping of olive genotypes did not allow them to be divided into two clear groups, although he differentiation of genotypes of Ionic and Adriatic regions was obvious because the majority of them were separated by the second principle coordinate, confirming previous results that were based on morphological characters (Ismaili *et al.,* 2013).

In addition, an analysis of molecular variance (AMOVA) was performed to estimate the level of differentiation between these two groups (Ionic and Adriatic). The analysis revealed higher values of variation between individuals within each group (92 %) than between these two groups (8 %). The calculated PhiPT (0.081) was weakly significant ($P < 0.05$). Nei's genetic distance between the group of genotypes of the Ionic region and that of Adriatic region was 9.4 %.



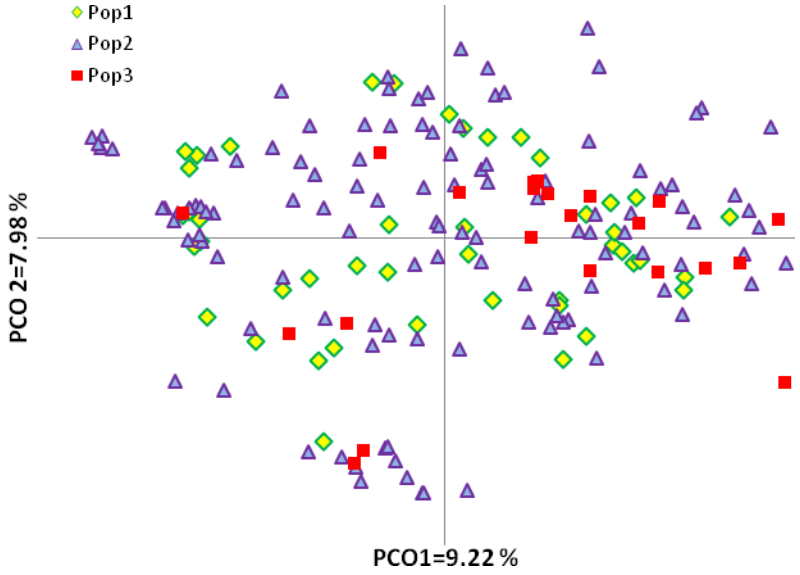
*Pop.1- Genotypes distributed at Adriatic region, Pop 2- Genotypes distributed at Ionic region

Figure 11: PCoA plot of principal coordinate analysis of 183 native olive genotypes, defined according to their main geographical area of distribution/ origin (Adriatic vs. Ionic region). The first two principal coordinates explained 10.57 and 9.03 % of the variance, respectively
Slika 11: Analiza glavnih koordinat (PCoA) narejena na osnovi geografske razvrstitve/izvora (Jadranska vs. Jonska regija) 183-ih genotipov oljk. Prvi dve glavni koordinati pojasnita 10.57 in 9.03 % raznolikosti.

Olives are used mainly for oil production or as table olives. Some olives are used for both purposes because they give a good percentage of oil and have large fruit. The links among these three groups of olive genotypes based on their product end use was evaluated by PCoA (Figure 12).

The olive genotypes were not separated into groups according to their end use in the PCoA diagram. Principal coordinates explained 23.21 % of genetic variation; the first and second coordinates revealed 9.22 % and 7.88 % of variation, respectively.

AMOVA analysis revealed that most genetic diversity was attributable to variability within olive groups (96 %) rather than between groups (4 %) of different product end-use. The calculated PhiPT=0.045 for all olive groups used for oil, table or dual use was significant ($P < 0.05$). The admixture observed in the principal coordinate analysis suggests that the abundant diversity existing in olives cannot be differentiated based on the end-use of the product.



*Pop.1- olive cultivars used for double purposes, Pop 2- cultivars used for oil production, Pop 3- cultivars used as table or canned olives production

Figure 12: PCoA plot of principal coordinate analysis of 183 olive genotypes defined according to their fruit end-use. The first two principal coordinates explained 9.22 and 7.98 % of the variance, respectively
Slika 12: Analiza glavnih koordinat (PCoA) narejena na osnovi končne uporabe 183-ih genotipov oljk. Prvi dve glavni koordinati pojasnita 9.22 in 7.98 % raznolikosti.

PCoA indicated that, despite slight overlapping, the olive genotypes showed grouping according to their origin and their area of distribution. The differentiation of genotypes according to their fruit end-use obtained in UPGMA analysis was not supported in the PCoA analysis.

Analysis of molecular variance confirmed a wide range of variability within groups and significant differences among groups. Our results are in line with previous studies, in which most variation was maintained within olive populations (Belaj *et al*., 2007, 2010, 2011), suggesting that, especially for wild olive, it is necessary to collect a higher number of samples within sampling sites in future studies (Belaj *et al.*, 2011). The obtained high level of variance within cultivars may be due to misllabelling and the presence of homonyms in olive germplasm (Omrani-Sabbaghi *et al*., 2007).

The exchange of olive plant material that may have occurred between Italy and Albania is also supported by the high similarity observed between some native olive cultivars or oleasters and Italian cultivars grown in Albania, as well with the cultivars Frantoio, Leccino, Carolea, Ascolana Tenera and those from the olive database. The Italian cultivars showed similarity with some ancient genotypes of the central and northern parts of

Albania, suggesting that the trade developed in ancient times between these two countries through the port city of Durrësi in Albania, or the well known trading activities of northen Albania with Venice indicate how the material exchange took place. Further studies employing a larger number of oleasters, including those of neighbouring countries will elucidate the origin of olive origin in Albania as well as the direction of olive exchange among countries.

5.8 ANALYSIS OF GENETIC STRUCTURE

Using model–based Bayesian clustering, the genetic structure of the sampled individuals was examined under K=2 clusters, which, as estimated by Evanno's ΔKs statistics (Evanno *et al.,* 2005), resulted in the best fit to the model for olive genetic structuring.

The proportion of its genome derived from different clusters was estimated. Genotypes were assigned to a cluster when 85% or more of their inferred genomes belonged to the cluster, with genotypes having a lower percentage being considered to be admixed (Štajner *et al.,* 2014). Using a threshold of >85% for the designation of group representatives, 81 genotypes were assigned to group A (red bars) and 53 genotypes to group B (green bars) and the remaining 49 genotypes (26.7%) were assigned as admixed genotypes.

Analysing the set of genotypes that were assigned to different gene pools, hierarchical levels of structure clusters could hardly be recognized. Both groups, A and B, were of mixed usage origin (oil, table or dual use) and no well- assigned population designations could therefore be recognized within genetic structure analysis probably due to the presence of oleaster genotypes in the sample set. No hierarchical differentiation based on olive end-use products was revealed by PCoA, either.

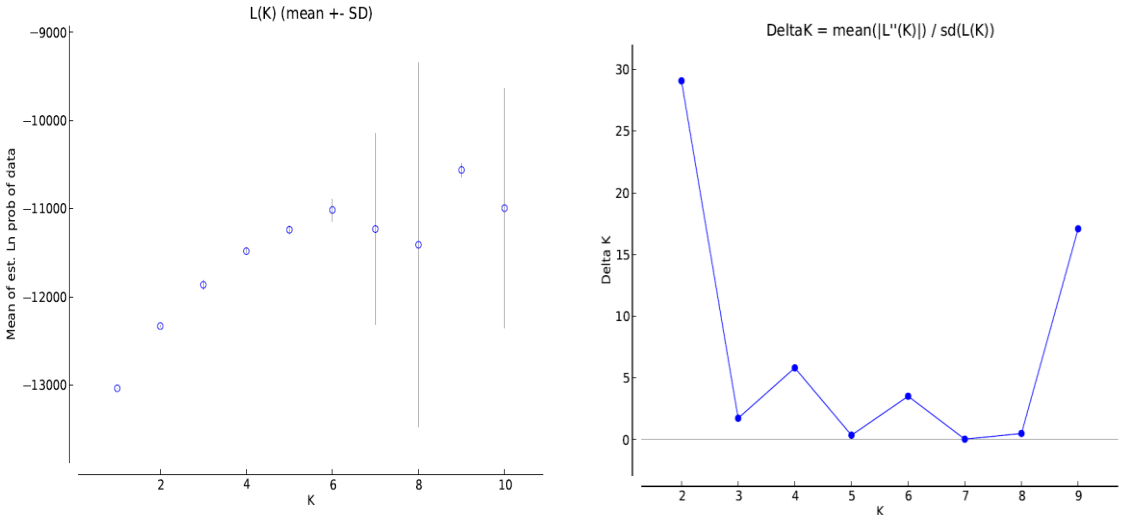
Weak differentiation was noticed in relation to the fact that group A (red color) contains 81 olive genotypes, consisting of oleasters (9), the majority of foreign imported olive cultivars (39) and native cultivars (33). While group B (green color) is composed of 53 native cultivars and lacks the oleaster genotypes and foreign imported cultivars. This weak differentiation was also observed in PCoA and AMOVA analysis, where only 7 % of variation was observed among native cultivars, oleasters and foreigner cultivars, indicating a mixed genetic structure of the olive genotypes.

The overall proportions of the membership of the sample to each of two clusters were 58.9 % and 41.1 % for cluster A and B, respectively. Average distances (expected heterozygosity) among individuals within a cluster were 0.708 and 0.559 for clusters A and B, respectively.

The low differentiation structure (K=2) detected within our analysed set of samples may be due to complex relationships among the olive cultivars or due to the inclusion of oleasters in the cluster.

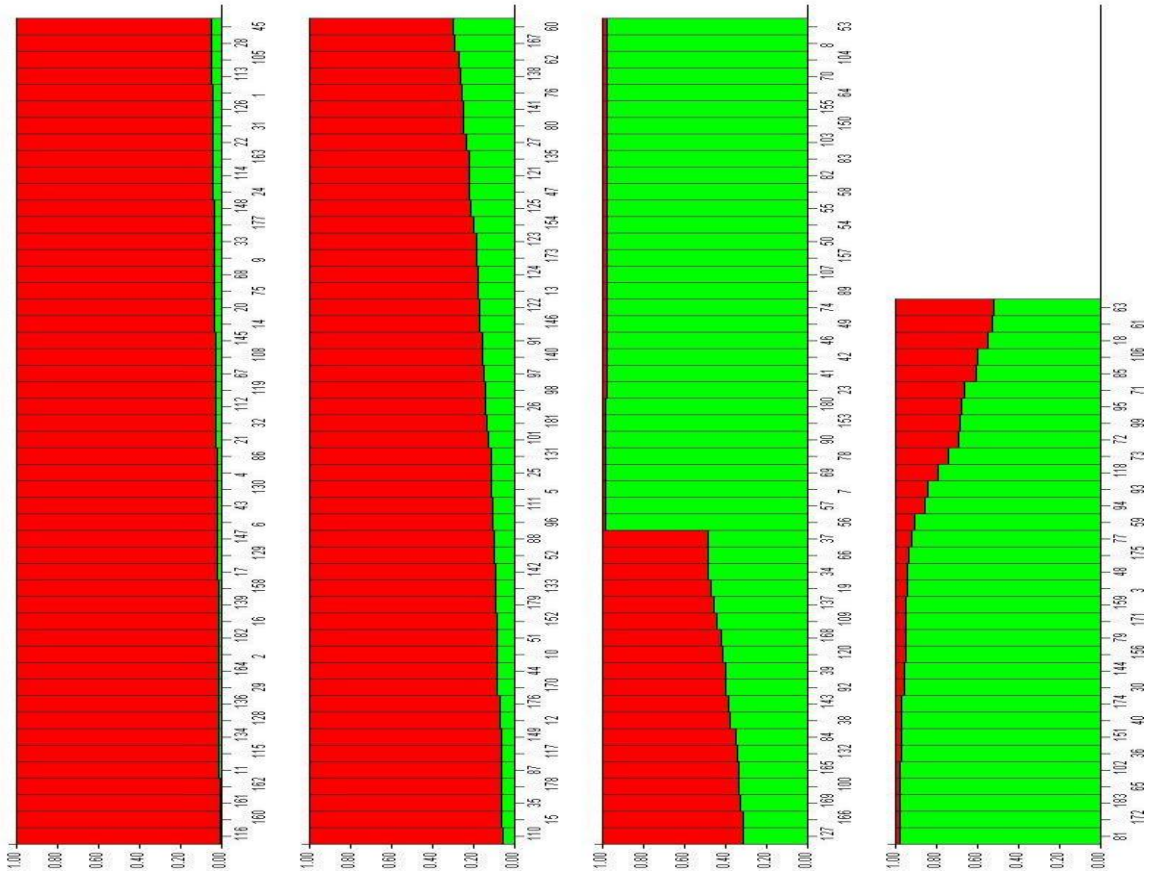
A weak differentiation of cultivars (51) and wild olives (107) was also reported by Belaj *et al.* (2010), where by wild olives and cultivars collected in three main Spanish olive growing regions were divided into four main gene pools. Only in two cases did the majority of wild olives and cultivars coming from the same region, Andalusia and Catalonia, cluster in separate gene pools. The wild olives of Valencia were assigned to another gene pool, while the remaining Andalusian and Valencian cultivated olives were clustered into the fourth gene pool. In contrast, Erre *et al*. (2010) reported the assignment of wild (21) and cultivated olives (57) of Sardinia into two different clusters.

Our population was not structurally based on the end use of the fruit (oil or table), in contrast to Do Val *et al*. (2012), who reported genetic differentiation to a certain extent according to the end-use of 60 Brasilian olive cultivars, indicating a relationship between their genetic makeup and agronomical traits such as size of fruit and percentage of the oil. This level of admixture is possibly also due to make up of the sample set, in which are included currently cultivated olive (native and foreign) ancient cultivars and oleaster forms.



A B

Figure 13: A- Mean Log likehood (Ln(K)+-s.d) averaged over 10 iterations for model based structure analysis (STRUCTURE); -B-Evanno test for delta K
Slika 13: A- povprečje Ln verjetnosti (Ln(K)+-s.d) za 10 ponovitev modelno-strukturne analize (STRUCTURE); -B- Test Evanno za delta K



116-Oleastër e Vlorës I; 160-Ulli I Hollë I Himarës I E; 161-Ulli I Hollë I Himarës II E; 162-Ulli I Hollë I Himarës III E; 11-Cazzinicchio E; 115-Ogliastro E; 134-Passola di Andria E; 128-Olivastër ILB I E; 136-Pendolino R; 29-Grossa di Spagna R; 164-Ulli I Hollë I Himarës I; 2-Ascolana E; 182- Vajs Peqini E; 16-Coratina I E; 139- Picholine E; 158- Ulli I Bardhë I Tiranës E; 17- Coratina II R; 129- Olivastër ILB II E; 147- Simona E; 6- Boç E; 43- Kallmet Kokërr Vogël E; 130- Olivastër Kokërr Gjatë I; 4- Bella di Spagna E; 86-Leccino E; 21- Frantoio E; 32- Halneiqis R; 112- Nolca E; 119- Olivastër Cilindrike I; 67- Kokërr Madh Berati I; 108- Nivica R; 145- Sant Agostino E; 14- Cima di Melfi E; 20- Dolce di Cassano E; 75- Krypsi I Beratit I; 68- Kokërr Madh Elbasani II I; 9-Carboncella R; 33- Itrana E; 177-Unafka I; 148- Termite di Bitteto E; 24- G44 I; 114- Ogliarola E; 163- Ulli I Hollë I Himarës IV E; 22-Frantoio R; 31- Halneiqis E; 126- Olivastër e Zezë Elbasani I E; 1- Amfisse E; 113- Oblonga R; 105- Moraiolo E; 28- Gordal Sevillana E; 45- Karen E; 110- Nociar E; 15- Cipressino E; 35- Kalinjot Cilindrik I; 178- Unafka R; 87- Lukova E; 117-Olivastër 4 I; 149-Toscanina E; 12- Cellina di Nardo E; 176- Unafka E; 170- Ulli I Kuq II-I; 44- Kallmet Kokërr Vogël I; 10- Carolea R; 51-Klon XVII E; 152- Ulli I Bardhë I Lezhës I; 179- Unknown I-I; 133- Olivastër e Zezë Berati I; 142- Pulazeqin R; 52- Klon IV E; 88- Lukova R; 96-Managjel R; 111- Nocellara Dell Bellice E; 5- Bllanic I; 25- G45 I; 131- Olivastër Kokërr Hollë Elbasani E; 101- Micka I; 181-Unknown III I; 26- G46 I; 98- Manzanilla R; 97- Manzanilla E; 140- Picual R; 91- Majatica E; 146- Santa Caterina E; 122- Olivastër e Kuqe e Tiranës II E; 13- Cerje I; 124- Olivastër e Kuqe e Tiranës I; 173- Ulli I Zi I; 123- Olivastër e Kuqe e Tiranës III E; 154- Ulli I Bardhë I Tiranës I-I; 125- Olivastër e Zezë e Tiranës I; 47- Klon I E; 121- Olivastër e Kuqe e Tiranës I E; 135- Passola E; 27- Ganjoll I; 80- Kuleks I; 141- Pulazeqin I; 76- Krypsi I Elbasanit I; 138- Perk I; 62- Klon XV E; 167- Ulli I Kuq I E; 60- Klon XIII E; 127-Olivastër e Zezë Elbasani I; 166- Ulli I Kuq UBT R; 169- Ulli I Kuq II E; 100- Mastoidis E; 165- Ulli I Kuq I Damsit I; 132-Olivastër10 I; 84- Kushan I-I; 38- Kalinjot Vertikal R; 143- Pulazeqin20 E; 92- Managjel Formë e Veçantë I E; 39- Kalinjot E; 120-Olivastër e Krujës I; 168- Ulli I Kuq I-I; 109- Noccellara Messinese E; 137- Peranzana E; 19- Cucurelia R; 34- Kala I; 66- Kokërr Madh Berati III R; 37- Kalinjot Vertikal I; 56-Klon IX E; 57-Klon X E; 7-Boç I; 69-Koran33-E; 78-Krypsi I Krujës R; 90-Llashtrak II-I; 153-Ulli I Bardhë I Shkodrës I; 180-Unknown II-I; 23-Frëng II-I; 41-Kalinjot R; 42-Kallmet Kokërr Madh I;46-Karen I; 49-Klon III-E; 74-Krypsi I Krujës I-I; 89-Llashtrak I-I; 107-Nivica E; 157-UBTr V-I; 50-Klon V E; 54-Klon VII-E; 55-Klon VIII E; 58-Klon XI E; 82-Kushan Kokërr Madh II-E; 83-Kushan Kokërr Mesëm I-E; 103-Mixan II-E; 150-Ulli I Bardhë I Krujës I-E; 155-Ulli I Bardhë I Tiranës II-I; 64-Kokërr Madh Berati I-R; 70-Koran69-E; 104-Mixan I; 8-Boç R; 53-Klon VI E; 81-Kushan Kokërr Madh I E; 172-Ulli I Zi II E; 183-Vajs Peqini I; 65-Kokërr Madh Berati II-R; 102-Mixan I E; 36-Kalinjot Pobrat E; 151-Ulli I Bardhë I Krujës II E; 40-Kalinjot I; 174-Ulli I Zi I Tiranës I-I; 30-Gjykatës I; 144-Qumështor I; 156-Ulli I Bardhë I Tiranës IV I; 79-Krypsi I Shkodrës I; 171-Ulli I Zi I E; 159-Ulli I Bardhë Pobrati E; 3-Bahuta I; 48-Klon II E; 175-Ulli I Zi I Tiranës II-I; 77-Krypsi I Krujës II-I; 59-Klon XII E; 94-Managjel II E; 93-Managjel I E; 118-Olivastër Cilindrike E; 73-Kotruvsi R; 72-Kotruvsi I; 99-Marks I; 95-Managjel I; 71-Kotruvsi E; 85-Kushan II-I; 106-Nisiot I; 18-Cucurelia E; 61-Klon XIV E; 63-Klon XVI E.

Figure 14: Inference of population structure based on microsatellite data and Bayesian simulation
Slika 14: Strukturna analiza populacij, ki temelji na simuliranju (metoda Bayes) mikrosatelitnih podatkov

5.9 CORE COLLECTION

The conservation of cultivated plants in *ex situ* collections is essential for the optimal management and use of their genetic resources (Haouane *et al*., 2011). The construction of a core collection, which contains the minimum number of genotypes that represent the entire germplasm diversity, as first proposed by Brown (1989) for the purpose of facilitating the exploitation of germplasm, will provide better management of the huge diversity found in olive germplasm.

The dataset of microsatellite profiles of 183 unique olive cultivars for 26 microsatellite loci was used to construct an olive core collection, using the M strategy (Gouesnard *et al*., 2001; Ecribano *et al*., 2008), implemented in CoreFinder software. A core collection was herein assembled for each of the olive germplasm collections maintained *ex situ*, *in situ* and for the overall Albanian olive germplasm, aiming to represent the entire genetic diversity identified in this study.

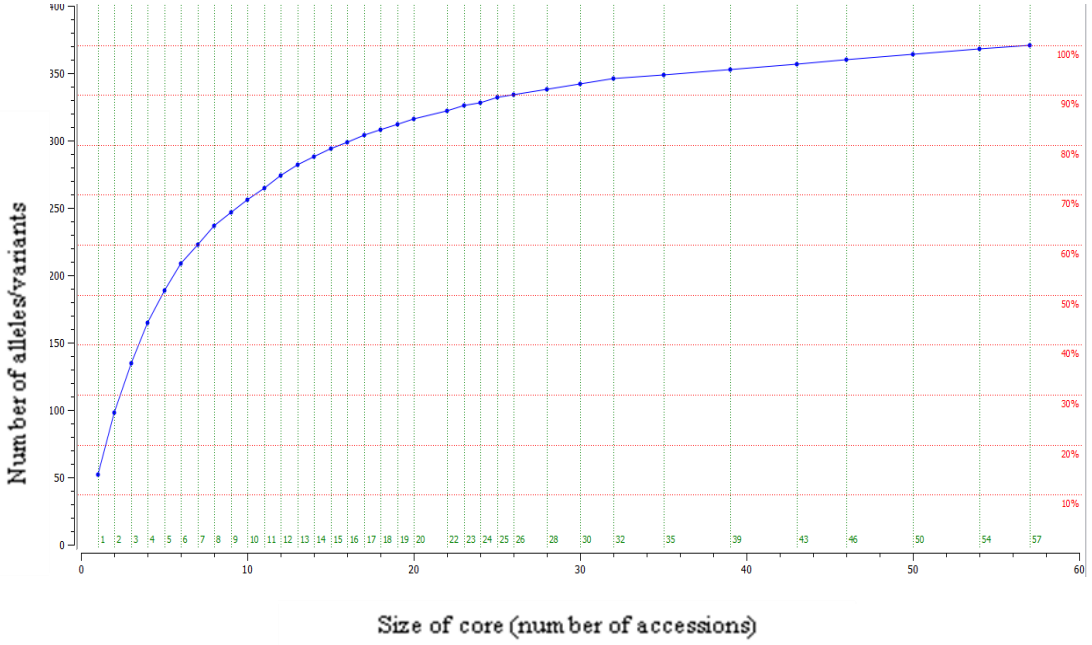


Figure 15: Genetic diversity as a function of the number of accessions included in the core collection
Slika 15: Genetska variabilnost kot funkcija števila akcesij vključenih v centralno zbirko

CoreFinder analysis showed that the 203 alleles identified by characterization of 183 unique genotypes with twenty six microsatellite loci could be represented by a core collection of 57 accessions (31 %) of all genotypes:‘Vajs Peqini I’, Boç E, Unafka I, Klon XVI E, Kala I, Oleastër e Vlorës I, Olivastër Kokërr Hollë Elbasani -E, Gordall Sevillana E, Olivastër 4 I, Pulazeqin R, Bllanic I, Pulazeqin I, Mastoidis E, Nolca E, Cucucrelia R,

Kallmet Kokërr Vogël E, Amfisse E,Cazzinicchio E,Olivastër e Zezë Berati I, Nocellara Dell Belice E, Olivastër e Zezë e Tiranës I, G44 I, KlonXIII E, KlonX E, Cellina di Nardo″ E, Karen E, Majatica E, Ogliarola E, Oblonga R, Ulli I Kuq II-I, Carboncella-R, Managjel Formë Veçantë I-E, Micka I, Ulli I Hollë Himare I, Lukova II R, Marks I, Nisiot I, Passola di Andria E, Simona E, Olivastër 10 I, Itrana E, Kushan I-I, Unafka E, Kryps Elbasani I, Cipressino E, Dolce di Cassano E, Toscanina E, KlonXII E, Ulli i Bardhë Pobrati E, Sant Agostino E, Cima di Melfi E, Noccellara Messinese E, Manzanilla E, Olivastër Kruje I

Several studies have arranged core collections of large and important collections. Haouane *et al.* (2011) developed a core collection of 67 (13 %) olive accessions from a total of 505 accessions from 14 Mediterranean countries planted at OWGB in Marrakech (Morocco). Only 15 % (59) of cultivars from the collection of WOGB Cordoba Spain, coming from 21 different Mediterranean countries, were necessary to capture 236 alleles identified I WOGB collection (Díez *et al.*, 2012). Muzzalupo *et al.* (2014) proposed an Italian olive core collection capturing 81 detected alleles of 489 olive cultivars of Italian germplasm CRA-OLI. Only 5 % of olive accessions were sufficient to construct the core collection. According to Belaj *et al*. (2012) a core collection containing 19-10 % of the total collection size is considered optimal to retain the bulk of genetic diversity found in the IFAPA, a germplasm collection of 361 olive accessions.

In our study, high percentages of accessions (31 %) were necessary to capture all the allelic richness found in olive germplasm in Albania, suggesting the presence of high levels of diversity in our sample genotypes.

The entire allelic richness of the actual *ex situ* core collection (117 accessions) may be represented by a minimum number of 45 accessions (38 %), while the diversity of the *in situ* collection (66 accessions) could be represented by 32 accessions (48 %). The olive genotypes for each partial collection, *ex situ* and *in situ* are given in Annex E.

However, we propose that for better management of the two olive collections in Albania and the preservation of the germplasm in a thorough and efficient way, the collections should be unified.

6 CONCLUSION

The Albanian olive germplasm material represents a large and very important genetic patrimony, which includes native cultivated olives, ancient and foreign cultivars and an abundant number of wild genotypes.

In order to include the whole Albanian olive germplasm, plant material of a total of 194 olive samples was collected from both *in situ* and *ex situ* germplasm collections. The samples were analysed by 26 microsatelite loci, of which 14 were previously published genomic-derived microsatellite markers and 12 were newly developed EST-derived microsatellites markers.

Identity analysis revealed 183 unique genotypes and their allelic profiles were used to calculate several indices of genetic diversity. Based on two distinct discriminative indices, PIC value (0.873) and probability of identity (0.052), DCA9 locus was shown to be the most informative of the 26 microsatellite loci that were analysed. The least informative locus, showing the lowest PIC value (0.250) and the highest PI (0.628) value, was locus SiBi 03. Our results confirmed that the analysed microsatellite markers area useful tool for exploring the genetic diversity of olive germplasm. Their high information content classified 19 microsatellite markers as highly informative (PIC >0.5) and twelve microsatellites as potential markers for genetic mapping (PIC >0.7), thus allowing confirmation of the robustness of the overall results achieved.

The new set of twelve EST-derived markers developed in our study, especially those of a tri-nucleotide repeat motif, has shown a high degree of polymorphism and could efficiently be applied in analyses of genetic diversity in olive: ancient, wild and collection material. They constitute a novel source of markers, which enable the assessment of variation in expressed genes, since they are related to coding regions of the genome. The development and comparative evaluation of the polymorphism level displayed by the 12 new EST-derived markers in relation to their origin of development and the length of the repeat motif in olive, as well as investigation of their efficiency in olive genotyping, was reported for the first time in our study.

Microsatellite markers were shown to be very useful tools, not only for the study of inter-cultivar diversity of olive but also for addressing the issues of intra-cultivar variation, synonymy and homonymy in olive cultivars. Six cases of perfectly matched synonyms and nine of near synonyms were revealed in our analysis. Only one case of synonymy out of three cases of synonyms previously reported was supported by our analysis. The other two pairs of synonyms cases previously reported showed mistmatches at 20-22 loci between them for the putative synonym pair Ulli i hollë Himare-Nisiot and between Qumështor and clones of Ulli i Bardhë Tirane, the allelic profiles mismatched at 10-24 loci. Sixteen intra-

cultivar variability groups and 50 groups of homonymy were determined, confirming the presence of high varietal disorder in the olive germplasm, which should be accurately resolved. Albanian olive germplasm present a unique gene pool. The comparative analysis of 183 uniqe genotypes with 80 allelic profiles from the Olive Database, 19 olive cultivars from Slovenian collection and 306 allelic profiles from Worldwide Olive Germplasm of Córdoba did not show any perfect match among allelic profiles.

Distance based clustering of genotypes, PCoA and analysis of molecular variance showed a wide range of variability within groups. The geographical grouping according to olive origin and area of distribution was confirmed, although not in a clearcut way. There was weak differentiation of olive genotypes based on their fruit end-use. The native olive genotypes showed tight clustering, suggesting their authoctonous origin. However, the analyses suggested that exchange of plant material occurred with neighbouring countries.

The clear separation of native cultivars from foreign cultivars and the presence of oleasters suggest that the most native olive cultivars originated from native oleasters. The overlapping of foreign cultivars with native olive cultivars from the central region in the PCoA analysis and the closer distance of present cultivars from foreign ones compared to their distance from ancient cultivars as representative of the domestication event, suggest that some of the present cultivars derived from the crossing of native and introduced foreign cultivars.

The minimum number of 57 samples (31 % of all olive germplasm) representing the entire genetic richness was assembled into a core collection, confirming high genetic diversity of the germplasm. The core collection includes also oleasters genotypes, ancient and foreign olive cultivars. Establishment of the core collection is necessary for the development of an improved breeding strategy and for safeguarding the wealth of genetic diversity.

The present study is the first report on the molecular characterization of Albanian germplasm and an assessment of genetic relationships of oleasters, ancient and present cultivars and imported foreign cultivars. The molecular data provide essential information to build the first molecular database of Albanian olive germplasm, which will also be deposited in the Worldwide Olea Database. The reference cultivars/genotypes and specific alleles developed in this study, as well as the identification of the minimum number of microsatellite markers required to distinguish all genotypes, will help in rapid testing and cultivar identification in the future.

7 SUMMARY (POVZETEK)

7.1 SUMMARY

Olive (*Olea europaea* L.) is one of the most important fruit crops, due to its worldwide use for oil production and table olives, as well as its application in the cosmetic sector. It is one of the oldest cultivated fruit trees, with a very long history in the Mediterranean Basin. It is an emblematic tree crop, even in Albania, whose olive germplasm has a very large genetic patrimony, represented by native cultivated olives, ancient and foreign cultivars and an abundant number of wild genotypes.

Historical evidence and archeological findings suggest that olive cultivation in Albania dates back to the 6th century BC. Through the long history of cultivation and mainly vegetative propagation, genetically different cultivars have been given the same name (homonymy) and *vice versa* some cultivars genetically identical or closely related have been denominated differently (synonymy). Generally, the generic denomination based on morphological characters, especially of the fruit is the main cause of high varietal disorder in plantations and olive collections. The existing olive germplasm thus needs to be properly identified and evaluated to ensure proper genetic resource management and preservation. The Albanian government is currently promoting an increase in olive trees planted in the country, stressing the need for accurate cultivar identification. Albanian olive germplasm have been traditionally characterized and identified by morphological methods, which are known to be influenced by environmental factors and the development stage of the plant. Only a limited number of cultivars have been analysed by 16 RAPD markers or by only two microsatellite loci.

Of the enormous number of molecular markers available, simple sequence repeats (SSR) are considered to be the most powerful markers for olive cultivar identification, genetic relationship assessment and parentage analysis, because of their advantages as co-dominant and highly polymorphic markers and producing data that can be easily transferred among laboratories. A set of 11 microsatellites have been chosen as highly polymorphic and informative, and they comprise a consensus list to be used in olive characterization in order to provide comparable data and to construct a worldwide olive molecular database.

The aim of the present study was the identification and evaluation of the genetic diversity of olive genotypes collected in Albania. Molecular characterization will provide accurate identification of olive germplasm, as the key to biodiversity preservation. Possible cases of synonymy and homonymy, and intra-cultivar variability, as well as planting misnaming or mislabeling and parentage kinship will be revealed. Reference cultivars were introduced to

enable data comparison with the worldwide olive genotypes deposited in the olive database. This comparison will highlight unknown cases of synonyms and homonyms.

The genetic relationships among olive genotypes and the variability partitioning among different geographical and breeding groups of olive will be evaluated and visualized by dendrogram by applying genetic distance analysis and model based clustering methods. The genetic relatedness analysis between native and foreign olive genotypes will give an insight into the origin of Albanian olive cultivars.

Another aim of this study was the development of a new set of expressed sequence tags based microsatellites (EST-SSR) markers and assessment of their efficiency in genotyping and estimation of genetic diversity of olive germplasm, and in comparison with genomic derived microsatellites. The validation of microsatellite markers in genotyping of olive will enable the identification of a set of the minimum number of markers required to distinguish all olive genotypes and the proper identification of olive germplasm, which can be used in further studies or in rapid germplasm evaluation.

The molecular data will provide also for the first time an inventory and a reference database of autochthonous olive genetic resources in Albania and enabled the arrangement of a core collection that accurately represents the entire genetic diversity in a minimum number of entities, providing proper management of olive germplasm.

A total of 194 accessions of *Olea europaea* L., of which 149 accessions were autochthonous Albanian olive genotypes and 45 of foreign origin, collected in Albania were genotyped by 26 microsatellite loci of which 14 were di-nucleotide microsatellite motifs originating from genomic libraries and 12 were new primers developed from EST containing tri and tetra- nucleotide repeat motifs. A group of 74 olive trees was represented by ancient olive genotypes, with an estimated age up to 3000 years old, maintained *in situ,* while 19 olive samples represented the wild form of *Olea europaea var. sylvestris,* known also as ‘oleaster’.

Identity analysis revealed 183 unique genotypes from the 194 olive accessions analyzed. The 183 unique microsatellite allelic profiles were further used to calculate several indices of genetic diversity. The degree of polymorphism revealed by means of 26 microsatellite loci was evaluated based on parameters of variability, such as: the number of alleles per loci (N_o); observed heterozigosity (H_o), expected heterozygosity (H_e), polymorphic information content(PIC), estimated frequency of null alleles (F_{null}); Hardy-Weinberg equilibrium (HW) estimated for each locus using Cervus 3.0 (Kalinowski *et al*., 2007). The probability of identity (PI) was calculated by statistical software Identity 1.0 (Wagner and Sefc, 1999). Ne- number of effective alleles and the frequency of alleles (Freq ≤ 5) were

calculated for each olive group defined by geographical origin and breeding by means of GenAlEx v 6.501 (Peakall and Smouse, 2006, 2012).

The effectiveness of the set of microsatellite markers in discriminating olive genotypes was evaluated, based on the observed diversity indices in relation to their origin (genome-derived *vs* EST-derived microsatellites) and the length of their repeat motif (di-, tri- and tetra-nucleotidic microsatellites). The minimum number of markers required to distinguish all the observed genotypes was identified by the AMaCAID script (Caroli *et al.,* 2011) designed in R program

A total of 203 alleles were amplified across 183 unique genotypes by the 26 microsatellite markers,with an average number of alleles per locus was 7.8. There were 92 rare alleles detected (freq <5 %). In total, 31 specific alleles were found in 24 genotypes, 29 alleles being specific to 22 Albanian olive genotypes. These alleles should be further evaluated and analysed, since they may be possible genotype specific markers, hence useful for further characterization of these genotypes. The occurrence of three alleles amplified by one genotype was observed in three microsatellite loci, DCA09, DCA11 and GAPU71 B.

The mean values of expected and observed heterozygosity were 0.678 and 0.744, respectively, while the mean polymorphic information content for the 26 microsatellite loci was 0.630. Based on two distinct discriminative indices, higher PIC value (0.873) and lower probability of identity (0.052), among the 26 microsatellite loci that were used, locus DCA9 was shown to be the most informative, while the least informative locus, showing a lower PIC value (0.250) and higher PI (0.628) value was SiBi 03. High information content classified 19 microsatellite markers as highly informative (PIC> 0.5) and 12 microsatellites as potential markers for genetic mapping (PIC> 0.7), thus allowing confirmation of the robustness of the overall results achieved. The estimated total probability of identical genotypes for the 26 microsatellite loci was 5.19×10^{-18}, showing the high discriminatory power of the microsatellite set used and ensuring accurate discrimination of our olive genotypes.

A set of 11 out of 26 microsatellite loci discriminated all the 183 olive genotypes in this study. This highly discriminatory set of 11 markers (DCA9, DCA18, GAPU59, GAPU71B, UDO24, EMO90, SNB19, SNB20, SiBi04, SiBi07 and SiBi19) consisted of six previously developed di-nucleotide and five tri-nucleotide and tetra-nucleotide new markers developed from ESTs in this study. With a combination of only three primers DCA18, GAPU59 and EMO90 it was able to discriminate 83 % (152) of unique genotypes. These three microsatellite markers may be very useful for rapid screening of olive germplasm in the collection.

The new set of twelve EST-derived markers was developed in our study. The EST-SSR produced a better quality of allelic patterns than genomic derived SSR. Tri- and tetra-nucleotide SSRs were not as prone to stutters as di-nucleotide genomic-SSR, due to their long core repeat of tri- and tetra-nucleotide, enabling better separation of alleles, although the application of a long final extension time also greatly reduced the stutters at di-nucleotide loci.

The average polymorphism of the analysed loci was 71 %. 64.9 % and 42.5% for di-, tri- and tetra-nucleotide microsatellites, respectively. The lowest level of polymorphism was obtained by tetra-nucleotide microsatellites, while the other two microsatellite categories showed roughly the same level of polymorphism. The majority of markers showed PIC value >0.5, indicating their ability to assess molecular diversity and genetic relatedness.

Microsatellite markers have been shown to be very useful tools not only for study of the inter-cultivar diversity of olive but also for addressing the issues of intra-cultivar variation, synonymy and homonymy in olive cultivars. Six cases of perfectly matched synonyms and nine of near synonyms were revealed in our analysis. Only one case of synonymy from three cases of synonyms previously reported was supported by our analysis. Sixteen intra-cultivar variability groups and 50 groups of homonymy were determined, revealing the presence of high varietal disorder of olive germplasm, which should be accurately resolved.

Olive cultivars from Albania present a unique genpool, the comparative analysis of 183 unique genotypes with 80 allelic profiles deposited olive database, 19 olive cultivars from Slovenian olive collection and 306 allelic profiles of Worldwide Olive Germplasm of Córdoba did not show any perfect match among the allelic profiles.

The parentage analysis did not reveal any true parentage relationships. However, it gave useful information about some trios that resulted to be false exluding as well false putative parents. Determining the exact pedigrees of olive cultivars is complicated due to the outcrossing nature of this species.

The genetic relationship among olive cultivars were evaluated using a genetic distance-based method. In the UPGMA constructed dendrogram, olive cultivars were clustered into 12 groups. The average genetic similarity among the 183 genotypes based on 26 microsatellite loci was 0.509. The average similarity among oleaster genotypes was lower (0.514) than among native cultivated olives (current and ancient cultivars) (0.554), indicating a high level of diversity within oleaster genotypes. Imported foreign cultivars showed the lowest average similarity (0.474), which may be due to the different origin of these cultivars. The native olive cultivars showed tight clustering, suggesting their

autochthonous origin. However, the analyses suggested that exchange of plant material may have occurred with neighbouring countries

The model-based method, PCoA and analysis of molecular variance showed a wide range of variability within olive groups. The geographical grouping according to olive origin and area of distribution was confirmed, though not in a clearcut way. There was weak differentiation of olive genotypes based on fruit end-use. This level of admixture is possibly also due to make up of the the sample set, in which are included currently cultivated olive (native and foreign) ancient cultivars and oleaster forms.

A minimum of 57 samples (31% of all olive germplasm), representing the entire genetic richness, was assembled into a core collection, confirming the high genetic diversity of germplasm. The core genotypes, including also oleasters, ancient and foreign olive cultivars, could be used for development of breeding strategy and for safeguarding the richness of the genetic diversity.

The present study is the first report on molecular characterization of Albanian germplasm and on the assessment of the genetic relationships of oleasters, ancient and present cultivars, and imported foreign cultivars. The molecular data provided essential information for constructing the first molecular database of olive germplasm which will also be deposited in the Worldwide Olea Database. The EST-based microsatellite markers developed in this study, as well as the identification of a minimum number of microsatellite markers required to distinguish all genotypes, will help in rapid testing and cultivar identification in the future.

The development and comparative evaluation of the level of polymorphism displayed from the 12 new EST-derived markers in relation to their origin of development and the length of the repeat motif in olive, as well as investigation of their efficiency in olive genotyping was reported for the first time in our study. They constitute a novel source of markers, which enable assessment of variation in expressed genes since they are related to coding regions of the genome.

7.2 POVZETEK

Oljka (*Olea europaea* L.) je zelo pomembna sadna vrsta, v svetovnem merilu, zaradi pridelave oljčnega olja in namiznih oljk ter uporabe v kozmetične namene. Oljka je ena izmed najstarejših kultiviranih sadnih vrst in se že zelo dolgo goji v mediteranskem prostoru. Tudi v Albaniji, kjer rastejo predstavniki avtohtonih gojenih oljk, starih in tujih sort in veliko število divjih genotipov ima dolgo zgodovino gojenja.

Zgodovinski dokazi in arheološke najdbe kažejo na to, da se je oljka v Albaniji začela gojiti v šestem stoletju pr. n. št. Skozi dolgo obdobje gojenja in predvsem vegetativnega razmnoževanja so bile sorte marsikdaj napačno poimenovane oz. identificirane in tako so na primer genetsko različne sorte dobile enaka imena (homonimi) in obratno, genetsko enake sorte so bile različno poimenovane (sinonimi). Poimenovanje osnovano na morfoloških lastnostih, predvsem na osnovi morfologije plodov je glavni razlog za veliko neskladnost in napačna poimenovanja v oljčnih in kolekcijskih nasadih. Zato je zelo pomembna natančna in pravilna identifikacija genskega materiala oljk, da lahko tako zagotovimo pravilno upravljanje in ohranjanje genskih virov. V Albaniji tako poteka akcija zasajanja oljk podprta s strani vlade in s tem tudi potreba po natančni identifikaciji, saj so bili v preteklosti kultivarji identificirani predvsem na osnovi morfoloških lastnosti, za katere pa vemo, da so pod vplivom okolja in razvojnega stadija rastline.

Izmed številnih markerjev, ki se uporabljajo v molekularnih študijah, so mikrosatelitni markerji (SSR – simple sequence repeats) najbolj aktualni za namen identifikacije, analize sorodnosti in starševstva tudi pri oljkah, predvsem zaradi svoje kodominantne narave ter visoke stopnje polimorfizma in ker so lahko primerljivi med laboratoriji. Za molekularno karakterizacijo oljk smo v pričujoči nalogi izbrali set 11-ih mikrosatelitov, ki so najbolj polimorfni in informativni in uporabni v širšem pomenu ter primerni tudi za izdelavo molekularne baze podatkov za oljko.

Cilj pričujoče naloge je identifikacija in evalvacija genetske raznolikosti kultivarjev oljk, ki rastejo v Albaniji. Z molekularno karakterizacijo lahko natančno identificiramo sorte oljk in tako pripomoremo k ohranjanju genskih virov in raznolikosti. Z molekularnimi analizami smo določili primere sinonimov, homonimov, znotraj-sortne variabilnosti, napak pri poimenovanju in relacije starši/potomci. V analizo smo vključili tudi referenčne kultivarje, in s tem omogočili analizo primerjave s kultivarji oljke, ki so dostopni v podatkovni bazi na spletu. S to primerjavo smo odkrili nove primere sinonimov in homonimov. Z genetsko analiza sorodnosti in primerjavo med avtohtonimi albanskimi in tujimi genotipi smo pridobili informacije o izvoru albanskih kultivarjev oljke. Genetska sorodnost med posameznimi sortami oljk in delež variabilnosti med različnimi geografskimi skupinami in skupinami oljk glede na končno uporabo so bili ocenjeni in vizualizirani s pomočjo dendrograma na osnovi analize genetske razdalje in na osnovi modelno-strukturne analize.

Nadalje smo v nalogi razvili set novih mikrosatelitnih markerjev, ki temeljijo na izraženih zaporedjih (EST-SSR) in ovrednotili njihovo uporabnost in učinkovitost v analizah genotipizacije in genetske raznolikosti genskega materiala oljk ter naredili primerjalno analizo z mikrosateliti, ki temeljijo na genomskih zaporedjih. Določili smo tudi minimalno število markerjev, ki so potrebni za ločevanje vseh genotipov vključenih v analizo in so tako uporabni za genetske študije in za hitro identifikacijo genskega materiala oljk. Z

dobljenimi molekularnimi podatki smo prvič zasnovali referenčno bazo podatkov avtohtonih oljk iz Albanije kar bo pripomoglo k organizaciji kolekcijskega nasada, ki naj bi predstavljal celotno genetsko diverziteto ob minimalnem številu posameznih dreves.

Z mikrosatelitnimi lokusi smo genotipizirali 194 akcesij *Olea europaea* L., od katerih je bilo 149 avtohtonih albanskih genotipov oljk, 45 akcesij pa je bilo tujega izvora. Genotipizacija je bila izvedena na osnovi 26 mikrosatelitnih lokusov, od tega jih je bilo 14 z di-nukletidnim motivom iz genomske knjižnice, 12 mikrosatelitov pa je bilo na novo izoliranih na osnovi izraženih zaporedij (EST-SSR in so vključevali tri- in tetra-nukleotidne ponovitve. Skupina starih oljk je vsebovala 74 kultivarjev z ocenjeno starostjo kultivarjev do 3000 let, vzdrževanih v *in situ* pogojih, 19 oljk pa je predstavljalo divjo vrsto *Olea europaea* var. *sylvestris*, sicer znano kot »oleaster«.

Z analizo mikrosatelitnih lokusov smo identificirali 183 edinstvenih genotipov izmed 194 analiziranih akcesij oljk. 183 različnih mikrosatelitnih profilov smo uporabili za izračun indeksov genetske raznolikosti. Stopnja polimorfizma na osnovi 26-ih mikrosatelitnih lokusov je bila z uporabo računalniškega programa Cervus 3.0 (Kalinovski in sod., 2007) ovrednotena z naslednjimi parametri raznolikosti: število alelov na lokus (No); opazovana heterozigotnost (Ho); pričakovana heterozigotnost (He); informacijska vrednost polimorfizma (PIC); ocenjena pogostost ničtih alelov (Fnull); ocena Hardy-Weinbergovega ravnovesja (HW). Verjetnost identitete (PI) smo izračunali z računalniškim programom Identity 1.0 (Wagner in Sefc, 1999). Število uporabnih alelov in pogostost alelov (pogostost $\leq$ 5) smo izračunali za vsako posamezno skupino oljk, definirano glede na geografski izvor, z računalniškim dodatkom GenAlEx 6.501 (Peakall in Smouse, 2006, 2012). Prav tako smo ocenili učinkovitost seta mikrosatelitnih markerjev za razlikovanje genotipov oljk glede na določene indekse raznolikosti in izvor markerja (pridobljen iz genomskega zaporedja vs. pridobljen iz zaporedij EST) ter glede na dolžino ponovljivega motiva (di-, tri- in tetra-nukleotidni mikrosateliti). Z dodatkom AMaCAID, izdelanim v programu R (Caroli in sod., 2011), smo določili minimalno število potrebnih mikrosatelitnih markerjev za razlikovanje genotipov oljk.

Na osnovi verjetnostne statistike smo izvedli analizo starševstva s programom IDENTITY 1.0 (Wagner in Sefc, 1999). Dendrogram, ki je prikazoval genetske odnose med 183-imi analiziranimi genotipi, smo izdelali in izrisali na osnovi DICE-ovega koeficienta podobnosti in metode neponderirane aritmetične sredine (UPGMA), s programom NTSYS 2.2 (Rohlf, 2005). Analizo glavnih komponent (PCoA) smo izvedli na osnovi genetske matrike oddaljenosti parov izračunani za kodominantne podatke in na osnovi matrike izračunane za različnih skupine oz. populacije oljk, ki so bile določene na osnovi podatkov žlahtnjenja, geografske razporeditve ali glede na uporabo oljk (olje, plod). Za analizo molekulske variance smo uporabili programski paket GenAlEx 6.5 (Peakall in Smouse 2006; 2012). To analizo smo izvedli z namenom, da bi določili relativno razporeditev

skupne genetske variacije med skupinami in znotraj skupin določenih oljk. Za določitev genetske strukture na osnovi mikrosatelitnih podatkov smo uporabili metodo na osnovi modela, ki je implementirana v programskem paketu STRUCTURE 2.3.4 (Pritchard s sod., 2000). Osnovna zbirka genotipov (Core collection) je bila v proučevanem vzorcu določena in ovrednotena na osnovi M strategije (Gouesnard et al., 2001; Escribano et al., 2008), z uporabo programa CoreFinder.

Skupno smo v analizi 26 mikrosatelitnih markerjev pri 183-ih genotipih pomnožili 203 alele. Mikrosatelitni aleli so bili dolgi od 136 bp (na lokusu GAPU71B) do 284 bp (na lokusu DCA15). Na posamezen lokus smo pomnožili od 2 alela (lokusa SiBi 03 and SiBi 11) do 19 alelov (DCA09), v povprečju pa 7,8 alela/lokus.

Določili smo frekvence pojavljanja posameznih alelov v analiziranem vzorcu, ki so se gibale od 0,003 to 0,821, s povprečjem 0,22. Odkrili smo 92 redkih alelov s frekvenco <5 %, kar predstavlja približno 45% vseh namnoženih alelov. Najvišje število redkih alelov smo odkrili na lokusu DCA09 in sicer kar 12, na lokusih DCA16 in DCA18 pa smo odkrili po 9 redkih alelov. Rezultati kažejo na to, da so omenjeni markerji učinkoviti pri odkrivanju redkih alelov, in bi bili zelo uporabni v primeru da so ti aleli povezani z agronomsko pomembnimi lastnostmi.

Skupno smo z analizo mikrosatelitnih lokusov našli 31 specifičnih alelov pri 24 genotipih in od tega je bilo 29 alelov specifičnih za 22 albanskih genotipov oljke. V prihodnosti je potrebno te specifične alele še dodatno ovrednotiti, saj so to lahko potencialni genotipsko specifični markerji uporabni za nadaljnje analize in karakterizacijo z njimi povezanih genotipov. Prav tako bi bili uporabni tudi za analizo sledljivosti oz. prisotnosti posameznih genotipov v oljčnem olju. V analizi genotipizacije smo dobili tudi tri-alelne profile in sicer so se po trije aleli pomnožili pri treh mikrosatelitnih lokusih, pri DCA09, DCA11 in GAPU71 B.

Povprečno število alelov na lokus je bilo 7,8, povprečni vrednosti pričakovane ter opazovane heterozigotnosti pa sta bili 0,678 ter 0,744. Povprečna vsebnost polimorfne informacije (PIC) za 26 mikrosatelitnih lokusov je bila 0,630. Izmed 26 mikrosatelitnih lokusov je lokus DCA9 kazal največjo informativnost, ki se kaže v najvišji PIC vrednosti (0,873) ter najnižji PI vrednosti (0.052), najnižjo informacijsko vrednost pa ima lokus SiBi03, ki ima nižjo PIC vrednost (0,250) ter višjo PI vrednost (0,628). Glede na visoko informacijsko vrednost smo 19 mikrosatelitnih markerjev klasificirali kot visoko informativne (PIC > 0.5), 12 mikrosatelitnih lokusov pa smo določili kot potencialne markerje za genetsko mapiranje (PIC > 7). S tem smo potrdili robustnost in uporabnost tovrstne analize in pridobljenih podatkov. Verjetnost identičnih genotipov (PI), izračunana za vsak lokus, je bila od 0,052 (DCA09) do 0,628 (SiBi03) Ocenjena skupna verjetnost identičnih genotipov za 26 mikrosatelitnih lokusov je bila 5,19 x 10-18, kar kaže na visoko

stopnjo razlikovanja pri uporabi tega seta mikrosatelitov ter zagotavlja natančno razlikovanje naših genotipov oljke.

Izmed 26 mikrosatelitnih lokusov smo s setom 11 lokusov razlikoval med vsemi 183 genotipi oljk, vključenih v analizo. Nabor 11 markerjev (DCA9, DCA18, GAPU59, GAPU71B, UDO24, EMO9, SNB19, SNB20, SiBi07 in SiBi19) predstavlja šest predhodno razvitih markerjev, od teh je bil eden z di-nukleotidnimi ter pet s tri-nukleotidnimi ponovitvami, ter na novo razvite tetra-nukleotidne markerje ki smo jih razvili iz zaporedij EST v tej študiji. S kombinacijo treh začetnih oligonukleotidov za markerje DCA18, GAPU59 in EMO90 smo lahko razlikovali 83 % (152) vseh edinstvenih genotipov. Kombinacija teh treh markerjev bi lahko bila uporabna za hitro testiranje genotipov oljk v zbirki.

V študiji smo razvili dvanajst markerjev iz zaporedij EST (EST-SSR), ki so pomnožili alelne vzorce višje kvalitete, kot mikrosateliti iz genomskih zaporedij. Mikrosateliti iz tri- in tetra-nukleotidnih ponovitev so imeli manj sekundarnih namnožitev (angl.-stutter bands), ki nastanejo zaradi zdrsa Taq polimeraze v procesu prepisovanja, v primerjavi z di-nukelotidnimi SSR, saj imajo daljši ponavljajoči motiv, ki omogoča boljše ločevanje alelov. Po drugi strani pa je bilo sekundarnih namnožitev bistveno manj tudi po uvedbi daljšega časa končnega pomnoževanja verige v reakciji PCR.

Povprečni polimorfizem analiziranih lokusov je znašal 71 % za di-, 64,9 % za tri- in 42,5 % za tetra-nukleotidne mikrosatelite. Najnižjo vrednost polimorfizma smo določili pri tetra-nukleotidnih mikrosatelitih, medtem, ko je pri ostalih dveh kategorijah nivo polimorfizma primerljiv. Večina markerjev ima PIC vrednosti > 0,5, kar kaže na to da so uporabni za genetske analize.

Obe skupini mikrosatelitnih markerjev, tako skupina razvita iz genomskih knjižnic kot tista razvita iz izraženih zaporedij (EST), sta bili primerni za analizo naše skupine vzorcev oljk. Genomski mikrosateliti so kazali višji nivo polimorfizma kot EST pridobljeni mikrosateliti. Vseeno so bili slednji še vedno dovolj informativni za uporabo, celo so imeli nekoliko višjo heterozigotnost v primerjavi z genomskimi mikrosateliti. Mikrosatelitni lokusi z di-nukleotidnimi in tri-nukleotidnimi motivi ponovitev so bili bolj polimorfni kot mikrosateliti s tetra-nukleotidnimi ponovitvami. Visoka stopnja polimorfizma, ki smo jo potrdili v pričujoči študiji kaže na to, da so EST mikrosateliti, še posebno lokusi s tri-nukleotidnimi ponovitvami, primerni za analize genetske raznolikosti oljk.

Mikrosateliti pa so poleg velike uporabne vrednosti v analizah raznolikosti med kultivarji oljk uporabni tudi za analize znotraj-sortne variabilnosti, ter za odkrivanje sinonimov in homonimov. Tako smo v študiji določili šest primerov popolnih sinonimov in devet primerov bližnjih sinonimov ter potrdili le en primer sinonima od treh predhodno objavljenih. Določili smo 16 skupin znotraj-sortne variabilnosti in 50 skupin homonimov,

kar kaže na to, da so v oljčnih nasadih drevesa napačno identificirana oz. napačno poimenovana.

Edinstvene genotipe oljk smo nadalje primerjali z mikrosatelitnimi alelnimi profili iz različnih virov; na 10-ih lokusih so bile narejene primerjave z 80 genotipi iz 11-ih držav (World Olive Database; www. oleadb.it), na 26-ih lokusih z 19 kultivarji iz slovenske kolekcije (DNA izolirana na Centru za rastlinsko Biotehnologijo in Žlahtnjenje, Biotehniška fakulteta, Univerza v Ljubljani) in na 13 lokusih s 306 genotipi iz 17 držav (Worldwide Olive Germplasm Bank of Córdoba; Trujillo et al. (2014).

Primerjalna analiza 183-ih genotipov iz Albanije z 80 alelnimi profili oljk shranjenih v bazi podatkov (www.oleadb.it), 19-imi kultivarji iz kolekcijskega nasada iz Slovenije in 306-imi alelnimi profili svetovne kolekcije oljk iz Cordobe kaže na raznolikost med obravnavanimi skupinami, in na osnovi tega rezultata lahko sklepamo, da oljke iz Albanije predstavljajo svojstven genetski fond.

Številne možne kombinacije starši-potomci so bile odkrite v analizi starševstva. Ddobljene kombinacije v večini primerov razkrivajo tesno relacijo med potomcem in enim od staršev. V večini primerov odkritih kombinacij je bila podobnost med potomcem in enim od staršev 96-98% kar kaže na tesno sorodnost. Analiza tudi kaže na to, da je analiza rodovnika zelo zapletena pri vrsti kot je oljka, predvsem zato, ker je tujeprašna rastlina.

Z analizo na osnovi genetske razdalje in metode UPGMA smo dobili 12 ločenih skupin, v katere so se razvrstili vzorci s povprečjem podobnosti 50,9 %. Vsaka od glavnih skupin se je razdelila na več manjših podskupin, kjer so bili združeni sorodni genotipi. Pri številnih primerih genotipov teh podskupin smo potrdili visoko stopnjo genetske podobnosti (>85 %), kar potrjuje veliko število sinonimov in odkrita variabilnost znotraj kultivarjev, ki jo je potrdila analiza identičnosti. Povprečje podobnosti med divjimi genotipi je bilo nižje (0,514) v primerjavi z avtohtonimi kultiviranimi genotipi (novimi in starimi) (0,554) kar kaže na to, da je nivo raznolikosti večji pri divjih oljkah. Tuji kultivarji, ki rastejo v Albaniji pa kažejo najnižjo stopnjo podobnosti (0,474), kar je verjetno posledica različnega izvora. V analizi nismo mogli določiti popolnega razlikovanja kultiviranih oljk in divjih oljk (olesatrov), kar kaže na prisotnost podivjanih oblik kultiviranih oljk v našem vzorcu. Vsekakor ne moremo izključiti tudi dejanske prisotnosti pravih divjih oljk (oleastrov) in podivjanih oblik. Zaključimo lahko, da bi nam analiza večjega števila oleastrov in vključitev v analizo oleastrov iz sosednjih držav pomagala bolje razumeti poreklo in razširjenost albanskih oljk. Avtohtoni kultivarji oljk so bili razvrščeni skupaj, kar bi lahko kazalo na njihov skupni izvor. Analize pa kažejo na to, da je prišlo do izmenjave sadilnega materiala med sosednjimi deželami.

Rezultati modelno strukturne analize, PCoA in analize variance kažejo na veliko variabilnost znotraj posameznih skupin kultivarjev oljk. Z analizo je bila delno potrjena tudi geografska razdelitev glede na izvor oljk in njihovo področje rasti. Medtem ko smo dobili zelo šibko razvrstitev glede na uporabo oljk (za olječno olje ali namizne oljke). Nejasna razvrstitev je lahko tudi posledica sestave analiziranega vzorca, kjer so vbili vključene tako trenutno gojene oljke (domače in tuje), stari kultivarji in divje oljke (olesatri).

Najmanjše število vzorcev, ki predstavljajo celotno genetsko variabilnost je 57 (31 % celotnega analiziranega vzorca), kar predstavlja veliko genetsko raznolikost analiziranega vzorca. Osnovni vzorec 57-ih genotipov, ki ga predstavljajo tako divje oljke kot tudi stari in tuji kultivarji je uporaben v načrtovanju strategije žlahtnjenja in ima velik pomen za ohranjanje genetske raznolikosti.

Vso genetsko raznolikost aktualne *ex situ* kolekcije (117 akcesij) predstavlja minimalno število 45-ih akcesij (38%). Medtem ko je raznolikost *in situ* kolekcije (66 akcesij) zajeta v 32-ih akcesijah (48%). Na osnovi dobljenih rezultatov predlagamo, da se obe kolekciji v Albaniji združita in tako bo upravljanje in ohranjanje genskih virov veliko bolj učinkovito.

Pričujoče delo predstavlja prvo študijo molekulske karakterizacije genetskega fonda albanskih oljk in podaja oceno genetskih odnosov oljk iz skupine divjih oljk, domačih sort in vnesenih tujih sort. Molekulski podatki predstavljajo pomembno informacijo, ki bo služila za izdelavo prve molekulske podatkovne baze albanskih oljk, ki jo nameravamo tudi vključiti v svetovno podatkovno zbirko. Novo razviti EST mikrosateliti v tej študiji in tudi izbor minimalnega števila mikrosatelitnih markerjev potrebnih za razlikovanje vseh genotipov nam bo omogočalo hitro testiranje in identifikacijo kultivarjev v prihodnosti.

Razvoj in primerjalna ocena stopnje polimorfizmov pri 12-ih novo razvitih EST mikrosatelitnih markerjih glede na njihov izvor in dolžino ponovitve motiva pri oljki, kot tudi ocena njihove uporabnosti za genotipizacijo oljke sta bili po našem prepričanju prvič predstavljeni. EST mikrosateliti predstavljajo nov tip markerjev, ki omogočajo oceno variabilnosti izraženih genov, saj so povezani s kodirajočimi regijami genoma.

8 REFERENCES

Abdelhamid S., Grati-kamoun N., Marra F., Caruso T. 2013. Genetic similarity among cultivates olive estimated through SSR markers. Scientia Agricola, 70, 1: 33-38

Alba V., Montemurro C., Sabetta W., Pasqualone A., Blanco A. 2009. SSR-based identification key of Olea europaea L. diffused in Southern-Italy. Scientia Horticulturae, 123: 11-16

Amane M., Ouzzani N., Lumaret R., Debain C. 2000. Chloroplast-DNA variation in the wild and cultivated olives (Olea europaea) of Marocco. Euphytica, 116: 59-64

Angiolillo A., Mencuccini M., Baldoni L. 1999. Olive genetic diversity assessed using amplified fragment length polymorphism. Theoretical and Applied Genetics, 98: 411-421

Arif I.A., Khan H.A, Shobrak M., Al Homaidan A.A., Al Sadoon M., Al Farban A. H., Bahkali A. H. 2010. Interpretation of electrophoretograms of seven microsatellite loci to determine the genetic diversity of Arabian Oryx. Genetics and Molecular Research, 9, 1: 259-265

Atienza A. G., De la Rosa R., Domínguez-García M. C., Martín A., Kilian A., Belaj A. 2013. Use of DarT Markers as a means of better management of the diversity of olive cultivars. Food Research International, 54: 2045-2053

Baali-Cherif D., and Besnard G. 2005. High genetic diversity and clonal growth in relict populations of Olea europaea subsp laperrinei (Oleaceae) from Hoggar, Algeria. Annals of Botany, 96, 5; 823-8330

Baldoni and Belaj 2009. Olive. Oil crops. In: Handbook of plant breeding 4th ed. Vollman J, Rajcan I. (eds). Springer Science: 397-421

Baldoni L., Cultrera N.G., Mariotti R., Ricciolini C., Arcioni S., Vendramin G.G., Buonamici A., Poceddu A., Sarri V., Ojeda M.A., Trujillo I., Rallo L., Belaj A., Perri E., Salimonti A., Muzzalupo I., Casagrande A., Lain O., Messina R., Testolin R. 2009. A consensus list of microsatellite markers for olive genotyping. Molecular Breeding, 24: 213-231

Baldoni L., Tosti N., Ricciolini C., Belaj A., Arcioni S., Pannelli G., Germana M.A., Mulas M., Porceddu A. 2006. Genetic structure of wild olive and cultivated olives in the central Mediterranean basin. Annals of Botany, 98: 935-942

Bandelj D., Jakše J., Javornik B. 2002. DNA fingerprinting of olive varieties by microsatellite markers. Food technology Biotechnology, 40, 3: 185-190

Bandelj D., Jakše J., Javornik B. 2004. Assessment of genetic variability of olive varieties by microsatellite and AFLP markers. Euphytica, 136: 93-102

Bandelj D., Javornik B. 2007. Microsatellites as a powerful tool for identification of olive (Olea europaea L.) planting material in nurseries. Annales, 17: 133-138

Banilas G., Minas J., Gregoriou C., Demoliou C., Kourti A., Hatzopoulos P. 2003. Genetic diversity among accessions of an ancient olive variety of Cyprus. Genome, 46: 370-376

Barranco D., Cimato A., Fiorino P., Rallo R., Touzani A., Castaneda C., Serafin F., Trujillo I. 2000. World Catalogue of Olive Varieties. Consejo Oleicola Internacional, Madrid: 360 p.

Bartolini G. 2008: The new 2008 web edition (data bank) of the "Olive Germplasm" http://www.oleadb.it/introduction.htm (Aug. 2014)

Bassi, D., Tura, D., Geuna, F., Failla, O. and Pedò, S. 2002. Characterisation of local olive (*olea europaea* l.) Accessions by oil composition, morphological and molecular markers methods. Acta Horticulturae, 586: 57-60

Bautista R., Crespillo R., Canovas F.M., Gonzalo Carlos M. 2002. Identification of olive-tree cultivars with Scar Markers. Euphytica, 129: 33-41

Beghè D., Ferrarini A., Ganino T., Fabbri A. 2011. Molecular characterization and identification of a group of local Olea europaea L. varieties. Tree Genetics & Genomes, 7: 1185-1198

Belaj A., Cipriani G., Testolin R., Rallo L., Trujillo I. 2004a. Characterization and identification of main Spanish and Italian olive cultivars by Simple-sequence –repeat markers. HortScience, 39, 7: 1557-1561

Belaj A., Dominguez-Gracía M.C., Atienza S.G., Urdíroz NM., De la Rosa R., Satovic Z., Martín A., Kilian A., Trujillo I., Valpuesta V., Del Río C. 2012. Developing core collection of olive (*Olea europaea L.*) based on molecular markers (DarTs, SSRs, SNPs) and agronomic traits. Tree Genetics & Genomes, 8: 365-378

Belaj A., Leon L., Satovic Z., De la Rosa R. 2011. Variability of wild olives (*Olea europaea* subsp. Europaea var. sylvestris) analysed by agro-morphological traits and SSR markers. Scienta Horticulturae, 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 the north-western Mediterranean assesses by SSR markers. Annals of Botany. 100, 3: 449-458

Belaj A., Muňoz-Diez C., Baldoni L., Satovic Z., Barranco D. 2010. Genetic diversity and relationships of wild and cultivated olives at regional level in Spain. Scientia Horticulturae, 124: 323-330

Belaj A., Satovic Z., Cipriani G., Baldoni L., Testolin R., Rallo L., Trujillo I. 2003b. Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and their effectiveness in establishing genetic relationships in olive. Theoretical and Applied Genetics, 107:736-744

Belaj A., Satovic Z., Ismaili H., Panajoti Dh., Rallo L., Trujillo I. 2003a. RAPD genetic diversity of Albanian olive germplasm and its relationships with other Mediterranean countries. Euphytica, 130: 381-395

Belaj A., Satovic Z., Trujillo I., Rallo L. 2004b. Genetic relationships of Spanish Olive cultivars using RAPD Markers. HortScience, 39, 5: 948-951

Besnard G and Bervillé A. 2000. Multiple origins for Mediterranean olive (Olea europaea L. ssp. Europaea) based upon mitochondrial DNA polymorphisms. Comptes Rendus de l' Academie des Scientes Paris, Sciences de la vie. Life Sciences, 323: 173-181

Besnard G and Bervillé A. 2002. ON chloroplast DNA variation in the olive (*Olea europaea* L.) complex: comparison of RFLP and PCR polymorphisms. Theoretical and Applied Genetics, 104: 1157-1163

Besnard G., Baradat P., Bervillé A. 2001a. Genetic relationships in the olive (*Olea europaea* L.) reflect multilocal selection of cultivars. Theoretical and Applied Genetics, 102: 251-258

Besnard G., Breton C., Baradat P., Khadari B., Bervillé A. 2001b. Cultivar identification in olive based on RAPD. Journal of the American Society of Horticultural Science, 126, 6: 668-675

Besnard G., Gracia-Verdugo C., Ruvio de Casas R., Treier U.A., Galland N., Vargas P. 2008. Polyploidy in olive complex (*Olea europaea*L): Evidence from flow cytometry and nuclear microsatellite analysis. Annals of Botany, 101: 25-30

Besnard G., Khadari B., Navascues M., Fernandez-Mazuesoc M., El Bakkali A., Arrigo N., Naali-Cherif D., Brunini-Bronzini de Caraffa V., Santoni S., Varkas P., Savolainen V. 2013. Proceeding of the Royal Sciety, 208: 1756

Blears MJ. De Grandis S.A, Lee H., and Trevors J.T. 1998. Amplified fragment length polymorphism (AFLP): e review of the procedure and its applications. Journal of industrial Microbiology & Biotechnology, 21: 99-114

Bogani P., Cavalieri D., Petruccelli R., Polsinelli L., Roselli G. 1994.Identification of olive tree cultivars by using random amplified polymorphic DNA. Acta Hort. 356: 98-101

Botstein D., White R. L., Skolnick M., Davis R. W., 1980. Construction of a genetic linkage map in man using Restriction Fragment Length Polymorphism. American Journal of Human Genetics, 32: 314-331

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

Bracci T., Sebastiani L., Busconi M., Fogher C., Belaj A., Trujillo I. 2009. SSR markers reveal the unique of olive cultivars from the Italian region of Liguria. Scientia Horticulturae, 122: 209-215

Brake M., Migdadi H., Al-Gharaibeh M., Ayoub S., Haddad N., El Oqlah A. 2014. Characterization of Jordanian olive cultivars (Olea europaea L.) using RAPD and ISSR molecular markers. Scientia Horticulturae, 176: 282-289

Breton C., Guerin J., Ducatillion C., Medail F., Kull A.C., Berville A. 2008. Taming the wild and 'wilding' the tame: Tree breeding and dispersal in Australia and the Mediterranean. Plant Science, 175: 197-205

Breton C., Terral J-F., Pinatel Ch., Medail F., Bonhomme F and Berville A. 2009. The origin of the domestication of the Olive tree. Biologies, 333: 1059-1064

Breton C., Tersac M., Bervillé A. 2006. SSR genetic diversity in wild olive (oleaster, *Olea europaea L.*) suggests several Plio-Pleistocene refuge zones in the Mediterranean basin and gene flow with olive. Journal of Biogeography, 33: 1916-1928

Brown AHD. 1989. Core collections: a practical approach to genetic resources management. Genome, 31, 2: 818-824

Busconi M., Sebastiani L., Fogher C. 2006. Development of SCAR markers for germplasm characterization in olive tree (*Olea europaea L*). Molecular Breeding. 17: 59-68.

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

Caroli S., Santoni S., Ronford J. 2011. AMaCAID: a useful tool for accurate marker choice for accession identification and discrimination. Molecular Ecology Resources, 11, 4: 733-738

Carriero F., Fontanazza G., Cellini F., Giorio G. 2002. Identification of simple sequence repeats (SSRs) in olive (Olea europaea L.). Theoretical and Applied Genetics, 104: 301-307

Cipriani G., Marrazzo M.T., Markoni R., Cimato A. 2002. Microsatellite markers isolated in olive (*Olea europaea L.)* are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars. Theoretical and Applied Genetics, 104: 223-228

D'Imperio M., Viscosi V., Scarano M-T., D'Andrea M., Zullo A.B., Pilla F. 2011. Integration between molecular and morphological markers for the exploitation of olive germplasm (Olea europaea). Scientia Horticulturae, 130: 229-240

De la Rosa R., Angiollillo A., Rallo L., Guerrero C., Pellegrini M., Besnard g., Berville a., Martin A., Baldoni L. 2003. A first genetic linkage map of olive (Olea europaea L) cultivars using RAPD , RFLP , AFLP and SSR markers. Theoretical and Applied Genetics, 106, 1273-1282

De la Rosa R., James C.M., 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: 265-267

De la Rosa R., James C.M., Tobutt K.R. 2004. Using microsatellites for paternity testing in olive progenies. HortSience, 39: 351-354

De la Rosa R., Belaj A., Munoz-Merida A., Trelles O., Ortiz-Martin I. 2013. Development of EST-derived SSR Markers with Long-core Repeat in Olive and Their Use for Paternity Testing. Journal of the American Society for Horticultural Science, 138, 4: 290-296

Delgado-Martinez F.J., Amaya I., Sanchez-Sevilla J.F. Gomez-Jimenez M.C. 2012. Microsatellite marker-based identification and genetic relationships of olive cultivars from Extremadura region of Spain. Genetics and Molecular Research, 11, 2: 918-932

Díaz A, Martín A , Barracado D., De la Rossa R., Rallo P. 2006b. self-incompatibility of Arbequina and Picual olive assessment by SSR markers. Journal of the American Society for Horticultural Science, 131; 250-255

Díaz A., De la Rossa R., Martín A., Rallo P. 2006a. Development, characterization and inheritance of new microsatellites in olive (Olea europaea L.) and evaluation of their usefulness in cultivar identification and genetic relationship studies. Tree Genetics & Genomes, 2: 165-175

Diaz A., Dela Rosa R., Rallo P., Munoz-Diez C., Trujillo I., Barranco D., Martin A., Belaj A. 2007. Selections of an Olive Breeding Program Identified by Microsatellite Markers. Crop Science, 47: 2317-2322

Dice L., R. 1945. Measure of the amount of ecologic association between species. Ecology, 26: 297-302

Díez C. M., Imperato A., Rallo L., Barranco D., Trujillo I. 2012. Worldwide core collection of olive cultivars based on simple sequence repeat and morphological markers. Crop Science, 52: 211-221

Díez C. M., Trujillo I., Barrio E., Belaj A., Barranco D., Rallo L. 2011. Centennial olive trees as reservoir of genetic diversity. Annals of Botany, 108: 797-807

do Val A.D.B., Ferreira J.L., Vierira Neto J., Pasqual M., de Oliviera A.F., Borém A., Cançado G.M.A. 2012. Genetic diversity of Brasilian and Introduced olive germplasm based on microsatellite markers. Genetics and Molecular Research, 11, 1: 556-571

Domínguez-García M. C., Belaj A, De la Rosa R., Satovic Z., Heller-Uszynska K., Kilian A., Martín A., Atienza A. G. 2012. Development of DarT markers in olive (Olea europaeaL.) and usefulness in variability studies and genome mapping. Scientia Horticulturae, 136: 50-60

Doveri S., Sabino Gil B., Díaz A., Reale S., Busconi M., Da Câmara Machado A., Martín A., Fogher C., Donini P., Lee D. 2008. Standardization of a set of microsatellite markers for use in cultivar identification studies in olive (Olea europaea L.). Scientia Horticulturae, 116: 367-373

Earl D., A, vonHoldt B., M. 2012. Structure Harvester: a website and program for visualizing structure output and implementing the Evanno method. Conservation Genetic Resources, 4, 2: 359-361

El Aabidine A.Z., Charafi J., Grout C., Doligez A., Santoni S., Moukhli A., Jay-Allemand C., El Modafar C and Khadari B. 2010. Cinstruction of genetic Linkage map for the olive based on AFLP and SSR markers. Crop Science, 50, 6 2291-2302

Ellis J.R. and Burke J.M. 2007. EST-SSRs as a resource for population genetic analyses. Heredity, 99: 125-132

Ercisli S., Barut E., Ipek A. 2009. Molecular characterization of olive cultivars using amplified fragment length polymorphism markers. Genetics and Molecular Research, 8, 2: 414-419

Ercisli S., Ipek A., Barut E. 2011. SSR marker-based DNA fingerprinting and cultivar identification of Olives (Olea europaea). Biochemical Genetics, 49: 555-561

Erre P., Chessa I., Munoz-Díez C., Belaj A., Rallo L., Trujillo I. 2010. Genetic diversity and relationship between wild and cultivated olives (Olea europaea L.) in Sardinia as assessed by SSR markers. Genetic Resources and Crop Evolution, 57: 41-54

Escribano P, Viruel M., A, Hormaza J., I. 2008. Comparison of different methods to construct a core germplasm collection in woody perennial species with simple sequence repeat markers. A case study in cherimoya (Annona cherimola, Annona ceae), an characterizat subtropical fruit tree species. Annales of Applied Biology, 153: 25-32

Essadki N., Ouzzani N., Lumaret R., Moumni M. 2006. ISSR variation in olive tree cultivars from Morocco and other western countries of the Mediterranean Basin. Genetic Resources and Crop Evolution, 53: 475-482

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software Structure, a simulation study. Molecular Ecology. 14: 2611–2620.

Fabbri A., Hormaza J.I., Polito V.S. 1995. Random amplified polymorphic DNA analysis of Olive (Olea europaea L.) cultivars. Journal of the American Society for Horticultural Science. 120, 3: 538-542

FAOSTAT. 2013. http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor (Aug. 2014)

Fendri M., Ferreira E., Caladi L., Abreu I., Rodriguez-Garcia M., Amorin M.I., Alche J.D. 2014. Discrimination of Portugese and Spanish olive cultivars using microsatellite markers. Academia Journal of Agricultural Research, 2, 3: 058-061

Fendri M., Trujillo I., Trigui A., Rodriguez-Gracía M.I., de Dioz Ramirez J. 2010. Simple sequence repeat identification and endocarp characterization of olive tree accessions in Tunisian germplasm collection. Horticultural Sciences, 45, 10: 1429-1436

Gadaleta A., Mangini G., Mule G., Bianco A. 2007. Characterization of dinucleotide and trinucleotide EST-derived microsatellites in the wheat genome. Euphytica, 153: 73-85

Gadaleta A., Giancaspro A., Zacheo S., Nigro D., Giove L.S., Colasuonni P., Blanco A. 2011. Comparison of genomic and EST-derived SSR markers in phylogenetic analysis of wheat. Plant Genetic Resources, 9, 2: 243-246

Ganino T., Bartolini G., Fabbri A. 2006. The classification of olive germplasm-A review. Journal of Horticultural Science & Biotechnology, 81, 3: 319-334

Gemas V.J.V, Almadanim M.C., Tenreiro R., Martins A., Fevereiro P. 2004. Genetic diversity in the olive tree (*Olea europaea* L subs. europaea) cultivated in Portugal revealed by RAPD and ISSR markers. Genetic Resources and Crop Evolution, 51: 501-511

Gomes S., Martins-Lopes P., Lopes L., Guedes-Pinto H. 2009. Assessing genetic diversity in *Olea europaea L*, using ISSR and SSR markers. Plant Molecular Biology Reports, 123: 82-89

Gomes S., Martins-Lopes P., Guedes-Pinto H. 2012. Olive tree genetic resources charachterization through molecular markers; Genetic Diversity in Plants. Mahut Çalişkan (eds). Rijeka, Intech: 27p.

Gouesnard B., Bataillon T., M, Decoux G., Rozale C., Schoen D., J, David J., L. 2001. MSTRAT: An Algorithm for Building Germ Plasm Core Collection by Maximizing Allelic or Phenotypic Richness. Journal of Heredity, 92: 93-94

Green P. S. 2002. A revision of Olea L. Kew Bulletin, 57: 91-140

Guichoux E., Lagache L., Wagner S., Chaumeil P., Leger P., Lepais O., Lepoittevin C., Malusa T., Revardel E., Salin F., Petit R.J. 2011. Current trents in microsatellite genotyping. Molecular Ecology Resources, 11: 591-661

Haidimitriou 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. Journal of the American Society of Horticultural Science, 130, 2: 211-217

Hakim I.R., Kammoun N.G., Makhloufi E., Rebai A. 2010. Discovery and Potential of SNP Markers in Characterization of Tunisian Olive Germplasm. Diversity, 2,1: 17-27

Hannachi H., Breton C., Msallem M., El Hadij S.B., El Gazzah M., 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 Horticulturae, 116: 280-290

Haouane H., Bakkali A.E., Moukhli A., Tollon C., Santoni S., Oukabli A., Modafar C.E., Khadari B. 2011. Genetic structure and core collection of the World Olive Germplasm

Bank of Marrakech: towards the management and use of Mediterranean olive genetic resources. Genetica, 139: 1083-1094

Hernández P., de la Rosa R., Rallo L., Dorado G., Martin A. 2001. Development of SCAR markers in olive *(Olea europaea*) by direct sequencing of RAPD products: applications in olive germplasm evaluation and mapping. Theoretical and Applied Genetics, 103, 5: 788-791

Hu J., Wang L., Li J. 2011. Comparison of genomic SSR and EST-SSR markers for estimating genetic diversity in cucumber. Biologia Plantarum, 55, 3: 577-580

Huang H., Lu J., Ren ZH., Hunter W., Dowd S.E., Dang P. 2011. Mining and validating grape (Vitis L.) ESTs to develop EST-SSR markers for genotyping and mapping. Molecular Breeding, 28: 241-254

Ibtissem L., Mouna M.A., Messaound M. 2014. Phenotypic diversity of some olive tree progenies issued from a Tunisian breeding program. European Scientific Journal, 10, 6: 292-313

Ismaili H., Çeloaliaj Q., 1995. Alcuni aspeti dell olivicultura Albanese. Atti Convegno L'Olivicoltura Mediterranea, Ist. Sper. Oliv. Rende (Cs), Italia: 85-92

Ismaili H., Gixhari B., Osmani R. 2013. The diversity of genetic resources in the Albanian olive. Agriculture and Forestry, 59, 3: 35-36

Ismaili H., Vorpsi V., Veshaj Z., Sevo R. 2010. Production and periodicity characteristics of some olive cultivars (*Olea europaea* L.) in Vlora environmental conditions. Albanian Journal of Agricultural Sciences, 3, 9: 32-38

Ismaili I. 2013. Gjermoplazma e ullirit: Katalogu i varieteteve autoktone te ullirit. Albania, Flesh press: 151 p.

Jaccoud D., Peng K., Feinstein D. M., Kilian A. 2001. Diversity Arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Research, 29, 4:e 25.

Jakše J., Štajner N., Tomić L., Javornik B. 2013. Application of Microsatellite Markers in Grapevine and Olives In: The Mediterranean Genetic Code- Grapevine and Olive. (Eds) Poljuha D., Sladonja B. (es.). Rijeka, Intech: 25-50

Kafazi N., Muço Th. 1984. Kultura e Ullirit. Albania, ILB press, 359p

Kalia R.K., Rai M.K., Kalia S., Singh R., Dhawan A.K. 2011. Microsatellite markers: an overview of recent progress in plants. Euphytica, 177: 309-334

Kalinowski S.T., Taper M.L., Marshall T.C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology, 16: 1099-1006

Kapaj A. M., and Kapaj I. 2012. Olive oil sector in Albania and its prespective. Olive oil-constituens, quality, health properties and bioconversions. (Ed) Boskou D. (ed.). Rijeka, Intech: 495-510

Kaya H. B., Cetin Oznur., Kaya H., Sahin M., Sefer F., Kahraman A. 2013. SNP discovery by Illumina-based Transcriptome sequencing of the olive and genetic characterization of Turkish olive genotypes revealed by AFLP, SSR and SNP markers. PloS ONE, 8, 9: e7367

Khadari B., Breton C., Mountier N., Roger J.P., Besnard G., Berville A., Dosba F.2003. The use of molecular markers for germplasm management in French olive collection. Theoretical and Applied Genetics, 106, 3: 521-529

Khadari B., El Aabidine A.Z. and Grout C. 2010. A genetic linkage map of olive based on Amplified Length Polymorphysm, Intersimple Sequence Repeat and Simple Sequence Repeat Markers. Journal of the American Society of Horticultural Science, 135, 6: 548-555

Koehmstedt A.M., Aradhy M.K., Soleri D., Smith J.L., Polito V.S. 2011. Molecular charachterization of genetic diversity, structure, and differentiation in the olive (Olea europaea L.) germplasm collection of United States Department of Agriculture. Genetic Resources and Crop Evolution, 58: 519-531

Kullaj E. 2012. Following olive footprints in Albania. Cultivation and culture, folklore and history traditions and uses. El-Kohly M (ed). Scripta Horticulturae, 13: 439 p.

Kumar P., Gupta V.K., Misra A.K., Modi D.R., Pandey B.K. 2009. Potential of molecular markers in plant biotechnology. Plant Omics Journal, 2, 4: 141-162

Kump B., Javornik B. 1996. Evaluation of genetic variability among common buckwheat (*Fagopyrum esculentum* Moench) populations by RAPD markers. Plant Science, 114: 149-158

Las Casas G., Scollo F., Distefano G., Continella A., Gentile A., La Malfa S. 2014. Molecular haracterization of olive *(Olea europaea* L) Sicilian cultivars using SSR markers. Biochemical Systematics and Ecology, 57: 15-19

Leoneti L., Imami D., Zhllima E and Stefanllari A. 2009. Report 'The olive and olive oil value chain in Albania' prepared by USAID-AAC (Albanian Agriculture Competitiveness) project.
http://www.eastagri.org/files/oil-albania.pdf (Avg. 2013)

Linos A., Nikoloudakis N., Katsiotis A., Hagidimitriou M. 2014. Genetic structure of the Greek olive germplasm revealed by RAPD, ISSR and SSR markers. Scientia Horticulturae, 175: 33-43

Lopes M. S., Mendonça D., Sefc K.M., Sabino-Gil F., de Camara-Machado A., 2004. Genetic evidence of intra-cultivar variability with Iberian olive cultivars. Hortscience, 39, 7: 1562-1565

Lumaret R., Ouzzani N., Michaud H., Vivier G., Deguilloux M.F., Di Giusto F. 2004. Allozyme variationof 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 generalized regression approach. Cancer Research, 27: 209-220

Mantia M.L. Lanin O., Caruso T., Testolin R., 2005. SSR-Based DNA fingerprints reveal the genetic diversity of Sicilian olive (Olea europaea L.) germplasm. Hournal of Horticultural Science and Biotechnology, 80, 5: 628-632

Marra F.P., 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 Genetics & 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. Scientia Horticulturae, 123: 82-89

Martins-Lopes P., Lima-Brito J., Gomes S., Meirinhos J., Santos L., Guedes-Pinto H. 2007.RAPD and ISSR molecular markers in Olea europaea L: Genetic Variability and molecular cultivar identification. Genetic Resources and Crop Evolution, 54: 117-128

Minch E., Ruiz-Linares A., Goldstein D., Feldman M., Cavalli-Sforza L.L. 1997. MICROSAT: A Computer program for Calculating Various Statistics on Microsatellite Allele Data, ver. 1.5.d. Stanford University, Stanford

Muzzalupo I., and Perri E. 2009. Genetic diversity in olive tree cultivars from Italy and other countries of the Mediterranean basin as revealed by RAPD and SSR marker. Advances in Horticultural Science, 23, 4: 263-275

Muzzalupo I., Lombardo N., Musachio A., Noce M.E., Pellegrino G., Perri E., Sajjad A. 2006. DNA sequence analysis of microsatellite markers enhances their efficiency for germplasm management in an Italian Olive collection. Journal of the American Society of Horticultural Science, 131, 3: 352-359

Muzzalupo I., Russo A., Chuappetta A.A., Benincasa C., Perri E. 2010. Evaluation of genetic diversity in Italian Olives (Olea europaea L) cultivars by SSR markers. Journal of Biotechnology, 150: 472-480

Muzzalupo I., Stefanizzi F., Perri E.2009a. Evaluation of olives cultivated in Southern Italy by Simple Sequence Repeat Markers. Horticultural Science, 44, 3: 582-588

Muzzalupo I., Stefanizzi F., Salimonti A., Fallabella R., Perri E. 2009b. Microsatellite markers for the identification of a group of Italian olive accessions. Scientia Agricola, 66, 5: 685-690

Muzzalupo I., Vendramin G.G., Chiappetta A. 2014. Genetic diversity of Italian olives (Olea europaea) germplasm analysed by SSR markers. The Scientific World Journal, 2014: 1-12

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 89, 3: 583-590

Nei M., Li W.H. 1979. Mathematical model for studying genetic variation in terms of endonucleases. Proceeding of the National Academy Sciences of the United States of America, 76: 5269-5273

Newton C., Lorre C., Sauvage C., Ivorra S., Terral J-F. 2014. On the origins and spread of Olea europaea L (olive) domestication: evidence for shape variation of olive stones at Ugaris Late Bronze Age Syria-a window on the Mediterranean basin and on the westward diffusion of olive variety. Vegetation history and Archaeobotany, 25, 5: 567-575.

Noormohammadi Z., Hosseini-Mazinani M., Trujillo I., Rallo L., Belaj A., Sadeghizadeh M., 2007. Identification and classification of Main Iranian olive cultivars using microsatellite markers. HortScience, 42, 7: 1545-1550

Noormohammadi Z., Hosseini-Mazinani M., Trujillo I., Belaj A. 2009. Study of intracultivar variation among main Iranian olive cultivars using SSR markers. Acta Biologica Szegediensis, 53, 1: 27-32

Omrani-Sabbaghi A., Shahriari M., Falahati-Anbaran M., Mohammadi S.A., Nankali A., Mardi M., Ghareyazie B. 2007. Microsatellite Markers based assessment of genetic diversity in Iranian olive (*Olea europaea* L.) collections. Scientia Horticulturae, 112: 439-4473

Osmani R. 1995. Asspeti del germplasma olivicolo Albanese. Frutticoltura, 3: 59-63.

Owen C.A., Bita E.C., Banilas G., Hajjar S.E., Sellianakis V., Aksoy U., Hepaksoy S., Chamoun R., Talhook S.N., Metzidakis I., Hatzoupoulos O., Kalaitzis P. 2005. AFLP reveals structural details of genetic within cultivated olive germplasm from the Eastern Mediterranean. Theoretical and Applied Genetics, 110, 7: 1169-76

Ozkaya M.T. Cakir E., Gokbayrak Z., Ercan H., Taskin N. 2006. Morphological and molecular haracterization of Derik Halhali olive (*Olea europaea* L.) accessions grown in Derik-Mardin province of Turkey. Scientia Horticulturae, 108: 205-209

Pashley C. H., Ellis J.R., McCauley D.E., Burke J.M. 2006. EST databased as a source for molecular markers: Lesons from Helianthus. Journal of Heredity, 97, 5: 381-388

Peakoll R, Smouse P., E. 2006. GenAlEx v.6: genetic analyses in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6: 288-295

Peakoll R, Smouse P., E. 2012. GenAlEx v.6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics, 28: 2537-2539

Poljuha D., Sladonja B., Šetić E., Milotić A., Bandelj D., Jakše J., Javornik B. 2008. DNA fingerprinting of olive varieties in Istria (Croatia) by microsatellite markers. Scientia Horticulturae, 115, 3: 223-230

Pontikis C.A, Loukas M., Kousounis G. 1980. The use of biochemical markers to distinguish olive cultivars. Journal of Horticultural Science, 55,4: 333-343

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population Structure Using Multilocus Data. Genetics, 155: 945-959

Rallo P., Dorado G., Martin A. 2000. Development of simple sequence repeats (SSRs) in olive tree (Olea europaea L.). Theoretical and Applied Genetics, 101: 984-989

Rao R., La Mura M., Corrado G., Ambros O., Foroni I., Perri E., Pugliano G. 2009. Journal of Horticultural Science & Biotechnology, 84, 3: 261-266

Reale S., Dovero S., Díaz A., Aniolillo A., Lucentini L., Pilla F., Martín A., Donini P., Lee D. 2006. SNP-based markers for discriminating olive (*Olea europaea* L.) cultivars Genome, 49: 1193-1205

Rekik I., Salimonti A., Kamonun N.G., Muzzalupo I., Lepais O., Gerber S., Perri E., Rebai A. 2008. Characterization and identification of Tunisian olive tree varieties by microsatellite markers. Horticultural Science, 43, 5: 1371-1376

Rohlf F.J. 2005. NTSYSpc Numerical Taxonomy and Multivariate Analysis System, version 2.2; Exeter Software: New York, NY USA

Rony C., Baalbaki R., Kalaitzis P., Talhouk S.N. 2009. Molecular characterization of Lebanese olive germplasm. Tree Genetics & Genomes, 5: 109-115

Rotondi A., Cultrera N.G.M., Mariotti R., Baldoni L. 2011. Genotyping and evaluation of local olive varieties of a climatically disfavoured region through molecular, morphological and oil quality parameters. Scientia Horticulturae, 130: 562-569

Rotondi A., Magli M., Ricciolini C., Balboni L. 2003. Morphological and molecular analyses for the characterization of a group of Italian olive cultivars. Euphytica, 132: 129-137

Roubos K., Moustakas M., Aravanopoulos F.A. 2010. Molecular identification of Greek olive (*Olea europaea* L) cultivars based on microsatellite loci. Genetics and Molecular Research, 9, 3: 1865-1876

Rozen S., Skaletsy J., 2000. Primer3 on www for general users and biologist programmers. In: Bioinformatic methods and protocols: Methods in Molecular Biology. Krawetz S., Misener I. (eds.). Humana Press: 365-386

Sabino Gil F., Busconi M., Da Câmara Machado A., Fogher C. 2006. Development and characterization of microsatellite loci from *Olea europaea*. Molecular Ecology Notes, 6: 1275-1277

Sanz-Cortés F., Badenes M.L., Paz S., Iniguez A., Llacer G. 2001. Molecular haracterization of olive cultivars using RAPD markers. Journal of the American Society of Horticultural Science, 126, 1: 7-12

Sanz-Cortés F., Parfitt D.E. Romero C., Struss D., Llacer G.and Badenes M.L 2003. Intraspecific olive diversity assessed with AFLP. Plant Breeding, 122: 173-177

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

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments.Nature Biotechnology, 18: 233-234

Sefc K.M., Lopes S. M., Mendoça D., Rodrigues Dos Santos M., Laimer Da Câmara Machado M., Da Câmara Machado A. 2000. Identification of microsatellite loci in olive (*Olea europaea*) their characterization in Italian and Iberian olive trees. Molecular Ecology, 9: 1171-1193

Sesli M and Yeğenoğlu E.D. 2010. Genetic relationships among and within wild and cultivated olives based on RAPDs.Genetics and Molecular Research, 9, 3: 1550-1556

Simko I. 2009. Development of EST-SSR markers for study of population structure in Letucce (Lactuca sativa L.) Journal of Heredity, 100, 2: 256-262

Soleri D., Koehmsted A., Aradhya M.K., Polito V., Pinney K. 2010. Comparing the historic olive trees (olea europaea L.) os Santa Cruz Island with contemporaneous trees in the Santa Barbara, CA area: a case study of diversity and structure in an introduced agricultural species conserved in situ. Genetic Resources and Crop Evolution, 57: 973-984

Soltis D.E, Soltis P. S. 1989. Isozymes in plant biology. London, Chapman and Hall: 268p.

Štajner N., Tomić L., Ivanišević D, Korać N., Cvetković-Jovanović T., Beleski K., Angelova E., Maraš V., Javornik B. 2014. Microsatellite inferred genetic diversity and structure of Western Balkan grapevines (Vitis vinifera L.) Tree Genetics & Genomes, 10: 127-140

Terral J-F. 1996. Wild and cultivated olive (*Olea europaea* L): a new approach to an old problem using inorganic analysis of modern wood and archaeological charcoal. Review of Palacobotany and Palynology, 91: 383-397

Terral J-F., Alonso A., Buxo I Capdevila R., Chatti N., Fabre L., Fiorentino G., Marinval Ph., Perez Jorda G., Pradat B., Rovira N., Alibert P. 2004. Historical biogeography of olive domestication (*Olea europaea L*) as revealed by geometrical morphometry applied to biological and archaeological material. Journal of Biogeography, 31, 1: 63-77

Tessier C., David J., This P., Boursiquot J.M., Charrier A. 1999. Optimization of the choice of molecular marker for varietal identification in Vitis vinifera L. Theoretical and Applied Genetics, 98: 171-177

Thomaj F., Panajoti Dh. 2003. Kultivarët e Ullirit. Tiranë, Pegi: 152 p.

Trentacoste E.R. and Puertas C.M. 2011. Preliminary characterization and morphological evaluation of the olive germplasm collection of Mendosa province (Argentina)

Trujillo I., Ojeda M.A., Urdiroz N.M, Potter D., Barranco D., Rallo L., Diez C.M. 2014. Identification of the Worldwide Olive Germplasm Bank of Cordoba (Spain) using SSR and morphological markers. Tree Genetics & Genomes, 10: 141-155

Trujillo I., Rallo I., Arus P. 1995. Identifying olive cultivars by isoenzyme analysis. Journal of the American Society of Horticultural Science, 120, 12: 318-325

Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research, 23, 21: 4407-4414

Wagner H.W., Sefc K.M. 1999. Identity1.0. Centre for Applied Genetics, University of Agricultural Sciences, Vienna

Wallander E and Albert A. V. 2000. Phylogeny and classification of Oleaceae based on RPS16 and TRNL-F sequence data. American journal of Botany, 87, 12: 1827-1841

Weber J.L. 1990. Informativeness of human (dC-dA) (dG-dT)n polymorphism. Genomics, 7: 524-530

Weising K., Nybom H., Wolff K., Kahl G. 2005. DNA fingerprinting in plants; Principles, methods and applications. 2nd edition. CRC press Taylor and Francis Group: 444 p.

Wen M., Wang H., Xia Z., Zou M., Lu C., Wenquan W. 2010. Development of EST-SSR and genomic-SSR markers to asses genetic diversity in *Jatropha curcas L.* BMC Research Notes, 3, 42: 1-8

Wiesman Z., Avidan N., Lavee S., Quebedeaux B. 1998. Molecular characterization of common olive varieties in Israel and the West Bank using Randomly Amplified Polymorphic DNA (RAPD) markers. Journal of the American Society of Horticultural Science, 123, 5: 837-841

Williams J. G.K. Kubelik A.R, Livak K., J., Rafalski J.A., Tingey S.V. 1990. Dna polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleid Acids Research, 18, 22: 6531-6535

Wittenberrg A. H.J D., van der Lee T., Cayla C., Kilian A., Visser R. G. F., Schouten H. J. 2005. Validation of the high-Throughput marker technology Dart using the model plant Arabidopsis thaliana. Molecular Genetics and Genomics, 274: 30-39

Woodhead M., Russell J., Squirrell J., Hollinghworth P.M., Mackenzie K., Gibby M., Powell W. 2005. Comparative analysis of population genetic structure in Athyrium distentifolium (Pteridophyta) using AFLs and SSRs from anonymous and transcribed gene regions. Molecular Ecology, 144, 6: 1681-1695

Wu S. B. Collins G., Sedgley M. 2004. A molecular lingage map of olive (Olea europaea L.) based on RAPD, microsatellite and SCAR markers. Genome, 47; 26-35

Zane I., Bargelloni L., Patarnello T. 2002. Strategies for microsatellite isolation: a review. Molecular Ecology, 11: 1-16

Zhang M., Mao W., Zhang G., Wu F. 2014. Development and characterization of polymorphic EST-SSR and genomic SSR markers for Tibetian annual wild barely. Plos ONE, 9 4: e94881

Zitoun B., de Caraffa V.G., Giannettini J., Breton C., Trigui A., Maury J., Gambotti C., Marzouk B., Berti L.2008. Genetic diversity in Tunisian olive accessions and their relatedness with other Mediterranean olive genotypes. Scienta Horticulturae, 115: 416-419

Zohary D., Hopf M. 2000. Domestication of plants in the Old World. 3rd edition. Oxford, University Press: 316 p.

Zohary D., Spiegel-Roy P. 1975. Beginnings of the fruit growing in old World. Science, 187: 319-327

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my supervisor Prof. Dr. Branka Javornik for giving me the great opportunity for my PhD study research, for her patience, motivation, and immense knowledge.

I am very grateful to the co-supervisor, Assoc. Prof. Dr. Nataša Štajner for her aspiring guidance, constructive criticism and friendly advice.

The completion of my research would have not been possible without the valuable help of Assoc. Prof. Dr. Jernej Jakše who assisted me even in the softwares domain.

I am deeply grateful to Dr. Hairi Ismaili for providing helpful suggestions and helped me in the collection of plant material.

A special thank to Mrs. Nataša Hren for her technical advice and assistance in laboratory work. I am grateful to all my labmates in Centre of Plant Biotechnology and Breeding, Biotechnical Faculty, University of Ljubljana, Slovenia for the stimulating discussions, and for everything we shared during our journey together.

Finally, I would like to thank my family and my husband for their constant support and belief in me.

ANNEX A

Identification code, given name, origin and end-use of olive genotypes

No	Name	Abbrevation	Origin/city	Location	End- use
1	Athena	AH-Ref	Slovenia	Reference Cultivar	Oil
2	Amfisse	Amfis-E	Italy	Ex-situ	Table
3	Arbequina	Arbequ-Ref	Spain	Reference Cultivar	Oil
4	Ascolana Tenera	AscoTen-Ref	Italy	Reference Cultivar	Table
5	Ascolana	Asco-E	Italy	Ex-situ	Dual
6	Bahuta	Bahut-I	Albania/Tirane	In-situ	Oil
7	Belica	Belica-Ref	Slovenia	Reference Cultivar	Oil
8	Bella di Spagna	BelDSp-E	Spain	Ex-situ	Table
9	Bllanic	Bllanic-I	Albania/Vlore	In-situ	Dual
10	Boç	Boç-E	Albania/Tirane	Ex-situ	Dual
11	Boç	Boç-I	Albania/Tirane	In-situ	Dual
12	Boç	Boç-R	Albania/Tirane	Repository	Dual
13	Buga	Buga-Ref	Slovenia	Reference Cultivar	Oil
14	Carboncella	Carbo-R	Italy	Repository	Oil
15	Carolea	Carol-R	Italy	Repository	Dual
16	Cazzinicchio	Cazz-E	Italy	Ex-situ	Dual/oil
17	Cellina di Nardo	CDNard-E	Italy	Ex-situ	Dual
18	Cerje	Cerje-I	Albania/Vlore	In-situ	Table
19	Cima di Melfi	CimDMe-E	Italy	Ex-situ	Oil
20	Cipressino	Cipress-E	Italy	Ex-situ	Oil
21	Cipressino	Cipress-Ref	Slovenia	Reference Cultivar	Oil
22	Coratina	Corat-I-E	Italy	Ex-situ	Oil
23	Coratina	Corat-II-R	Italy	Repository	Oil
24	Črnica	Cnica-S	Slovenia	Reference Cultivar	Oil
25	Cucurelia	Cucur-E	Italy	Ex-situ	Dual
26	Cucurelia	Cucur-R	Italy	Repository	Dual
27	Dolce di Cassano	DolDCa-E	Italy	Ex-situ	Oil
28	Frantoio	Franto-E	Italy	Ex-situ	Oil
29	Frantoio	Franto-R	Italy	Repository	Oil
30	Frantoio	Franto-Ref	Italy	Reference Cultivar	Oil
31	Frengu	Freng-I	Albania/Kruje/Tirane	In-situ	Oil
32	G44	G44-I	Albania/Shkoder	In-situ	Oil
33	G45	G45-I	Albania/Shkoder	In-situ	Oil
34	G46	G46-I	Albania/Shkoder	In-situ	Oil
35	Ganjoll	Ganjoll-I	Albania/Shkoder	In-situ	Dual
36	Gordal Sevillana	GordSev-E	Spain	Ex-situ	Table

No	Name	Abbrevation	Origin/city	Location	End- use
37	Grossa di Spagna	GrosDSp-R	Italy	Repository	Table
38	Gjykates	Gjykat-I	Albania/Tirana	In-situ	Dual
39	Halneiqis	Haln-E	Greece	Ex-situ	Table
40	Halneiqis	Haln-R	Greece	Repository	Table
41	Itrana	Itrana-E	Italy	Ex-situ	Dual
42	Itrana	Itrana-Ref	Slovenia	Reference Cultivar	Dual
43	Kala	Kala-I	Albania/Vlore	In-situ	Dual
44	Kalinjot Cilindrik	KalCil-I	Albania/Vlore	In-situ	Dual
45	Kalinjot Pobrat	KalPob-E	Albania/Vlore	Ex-situ	Dual
46	Kalinjot Vertikal	KalVert-I	Albania/vlore	In-situ	Dual
47	Kalinjot Vertikal	KalVert-R	Albania/Vlore	Repository	Dual
48	Kalinjot	Kal-E	Albania/Vlore	Ex-situ	Dual
49	Kalinjot	Kal-I	Albania/Vlore	In-situ	Dual
50	Kalinjot	Kal-R	Albania/Vlore	Repository	Dual
51	Kallmet Kokërr Madh	KallKM-I	Albania/Lezhe	In-situ	Dual
52	Kallmet Kokërr Vogël	KallKV-E	Albania/Lezhe	Ex-situ	Oil
53	Kallmet Kokërr Vogël	KallKV-I	Albania/Lezhe	In-situ	Oil
54	Karen	Karen-E	Albania/Tirane	Ex-situ	Oil
55	Karen	Karen-I	Albania/Tirane	In-situ	Oil
56	Klon 10/1	Kl I-E	Albania/Vlore	Ex-situ	Dual
57	Klon 13	Kl II-E	Albania/Vlore	Ex-situ	Dual
58	Klon 16/2	Kl III-E	Albania/Vlore	Ex-situ	Dual
59	Klon 17/2	Kl V-E	Albania/Vlore	Ex-situ	Dual
60	Klon 17/3	Kl XVII-E	Albania/Vlore	Ex-situ	Dual
61	Klon 17/4	Kl IV-E	Albania/Vlore	Ex-situ	Dual
62	Klon 19/2	Kl VI-E	Albania/Vlore	Ex-situ	Dual
63	Klon 20/2	Kl VII-E	Albania/Vlore	Ex-situ	Dual
64	Klon 21/1	Kl VIII-E	Albania/Vlore	Ex-situ	Dual
65	Klon 7/1	Kl IX-E	Albania/Vlore	Ex-situ	Dual
66	Klon 7/2	Kl X-E	Albania/Vlore	Ex-situ	Dual
67	Kloni 22	Kl XI-E	Albania/Vlore	Ex-situ	Dual
68	Kloni 23	Kl XII-E	Albania/Vlore	Ex-situ	Dual
69	Kloni 25	Kl XIII-E	Albania/Vlore	Ex-situ	Dual
70	Kloni 27	Kl XIV-E	Albania/Vlore	Ex-situ	Dual
71	Kloni 29	Kl XV-E	Albania/Vlore	Ex-situ	Dual
72	Kloni 7/3	Kl XVI-E	Albania/Vlore	Ex-situ	Dual
73	Kokërr Madh Berati-1	KMB I-R	Albania/Berat	Repository	Table
74	Kokërr Madh Berati-2	KMB II-R	Albania/Berat	Repository	Table
75	Kokërr Madh Berati-3	KMB III-R	Albania/Berat	Repository	Table
76	Kokërr Madh Berati	KMB-I	Albania/Berat	In-situ	Table
77	Kokërr Madh Elbasani-2	KME II-I	Albania/Elbasan	In-situ	Tabe

No	Name	Abbrevation	Origin/city	Location	End- use
78	Koran 33	Koran 33-E	Albania/Tirane	Ex-situ	Oil
79	Koran 69	Koran 69-E	Albania/Tirane	Ex-situ	Oil
80	Kotruvsi	Kotr-E	Albania/Berat/Fier/Mallakaster	Ex-situ	Oil
81	Kotruvsi	Kotr-I	Albania/Berat	In-situ	Oil
82	Kotruvsi	Kotr-R	Albania/Berat	Repository	Oil
83	Krypsi I Krujës-1	KrypKr I-I	Albania/Kruje	In-situ	Dual
84	Krypsi I Beratit	KrypBr-I	Albania/Berat	In-situ	Table
85	Krypsi I Elbasanit	KrypEl-I	Albania/Elbasan	In-situ	Dual
86	Krypsi I Krujës-2	KrypKr II-I	Albania/Tirane	In-situ	Dual
87	Krypsi I Krujës	KrypKr-R	Albania/Kruje	Repository	Dual
88	Krypsi I Shkodrës	KrypShk-I	Albania/Shkoder	In-situ	Dual
89	Kuleks	Kuleks-I	Albania/Tirana	In-situ	Oil
90	Kushan Kokërr Madh-1	KusKMa I-E	Albania/Tirane/Durres/Kruje	Ex-situ	Oil
91	Kushan Kokërr Madh-3	KusKMa III-E	Albania/Tirane/Durres/Kruje	Ex-situ	Oil
92	Kushan Kokërr Mesëm-1	KusKMe I-E	Albania/Tirane/Durres/Kruje	Ex-situ	Oil
93	Kushan-1	Kus I-I	Albania/Tirane	In-situ	Oil
94	Kushan-2	Kus II-I	Albania/Tirane	In-situ	Oil
95	Leccino	Leccino-E	Italy	Ex-situ	Oil
96	Leccino	Leccino-Ref	Slovenia	Reference Cultivar	Oil
97	Leccione	Leccione-Ref	Slovenia	Reference Cultivar	Oil
98	Lukova-2	Lukov-E	Albania/Sarande	Ex-situ	Dual/Oil
99	Lukova-2	Lukov-R	Albania/Sarande	Repository	Dual/Oil
100	Llashtrak-1	Llasht I-I	Albania/Shkoder	In-situ	Table
101	Llashtrak-2	Llasht II-I	Albania/Shkoder	In-situ	Table
102	Majatica	Majatic-E	Italy	Ex-situ	Dual/table
103	Managjel Formë Veçantë	ManaFV I-E	Albania/Kruje/Lezhe	Ex-situ	Dual
104	Managjel-1	Mana I-E	Albania/Kruje/Lezhe	Ex-situ	Dual
105	Managjel-2	Mana II-E	Albania/Kruje/Lezhe	Ex-situ	Dual
106	Managjel	Mana-I	Albania/Kruje	In-situ	Dual
107	Managjel	Mana-R	Albania/Kruje/Lezhe	Repository	Dual
108	Manzanilla	Manzan-E	Spain	Ex-situ	Dual
109	Manzanilla	Manzan-R	Spain	Repository	Dual
110	Marks	Mark-I	Albania/Berat	In-situ	Dual
111	Mastoidis	Mastoid-E	Greece	Ex-situ	Oil
112	Maurino	Maurino-Ref	Slovenia	Reference Cultivar	Oil
113	Micka	Micka-I	Albania/Tirane	In-situ	Oil
114	Mixan-1	Mixa I-E	Albania/Elbasan	Ex-situ	Oil
115	Mixan-2	Mixa II-E	Albania/Elbasan	Ex-situ	Oil
116	Mixan	Mixa-I	Albania/Elbasan	In-situ	Oil
117	Moraiolo	Moraiol-E	Italy	Ex-situ	Oil
118	Nisiot	Nisiot-I	Albania/Fier	In-situ	Dual

No	Name	Abbrevation	Origin/city	Location	End- use
119	Nivica	Nivic-E	Albania/Sarande	Ex-situ	Dual/Oil
120	Nivica	Nivic-R	Albania/Sarande	Repository	Dual/Oil
121	Noccellara Messinese	NocceMes-E	Italy	Ex-situ	Dual
122	Nocciara	Nociar-E	Italy	Ex-situ	Dual
123	Nocellara Dell Bellice	NocDBe-E	Italy	Ex-situ	Table
124	Nocellara Dell Bellice	NocDBel-Ref	Slovenia	Reference Cultivar	Oil
125	Nolca	Nolca-E	Italy	Ex-situ	Table
126	Oblonga	Oblong-R	USA	Repository	Oil
127	Ogliarola	Ogliaro-E	Italy	Ex-situ	Oil
128	Ogliastro	Ogliast-E	Italy	Ex-situ	Oil
129	Oleastër e Vlorës	OleasVl-I	Albania/Vlore	In-situ	Oil
130	Olivastër-4	O4-I	Albania/Tirane	In-situ	Oil
131	Olivastër cilindrike	Ocilind-E	Albania/	Ex-situ	Oil
132	Olivastër cilindrike	Ocilind-I	Albania/Tirane	In-situ	Oil
133	Olivastër e Krujës	Okr-I	Albania/Kruje	In-situ	Oil
134	Olivastër e Kuqe e Tiranës-1	OKTr I-E	Albania/Tirane	Ex-situ	Oil
135	Olivastër e Kuqe e Tiranës-2	OKTr II-E	Albania/Tirane	Ex-situ	Oil
136	Olivastër e Kuqe e Tiranës-3	OKTr III-E	Albania/Tirane	Ex-situ	Oil
137	Olivastër e Kuqe e Tiranës	OKTr-I	Albania/Tirane	In-situ	Oil
138	Olivastër e Zezë e Tiranës	OZTr-I	Albania/Tirane	In-situ	Oil
139	Olivastër e Zezë Ebasani-1	OZEl I-E	Albania/Elbasan	Ex-situ	Oil
140	Olivastër e Zezë Elbasani	OZEl-I	Albania/Elbasan	In-situ	Oil
141	Olivastër ILB-1	OILB I-E	Albania/	Ex-situ	Oil
142	Olivastër ILB-2	OILB II-E	Albania/	Ex-situ	Oil
143	Olivastër Kokërr Gjatë	OKGj-I	Albania/Tirane	In-situ	Oil
144	Olivastër Kokërr Hollë Elbasani	OKHEl-E	Albania/Elbasan	Ex-situ	Oil
145	Olivastër-10	O10-I	Albania/Durres	In-situ	Oil
146	Olivastër e Zezë Berati	OZBr-I	Albania/Tirane	In-situ	Oil
147	Passola di Andria	PassoDA-E	Italy	Ex-situ	Table
148	Passola	Passo-E	Italy	Ex-situ	Table
149	Pendolino	Pendo-R	Italy	Repository	Oil
150	Pendolino	Pendo-Ref	Slovenia	Reference Cultivar	Oil
151	Peranzana	Peranz-E	Italy	Ex-situ	Double
152	Perk	Perk-I	Albania/Tirane	In-situ	Dual
153	Picholine	Picholi-E	France	Ex-situ	Dual
154	Picholine	Picholi-Ref	Slovenia	Reference Cultivar	Dual
155	Picual	Picual-R	Spain	Repository	Oil
156	Pulazeqin	Pulaz-I	Albania/Vlore	In-situ	Oil
157	Pulazeqin	Pulaz-R	Albania/Vlore	Repository	Oil
158	Pulaz-20	Pulaz20-E	Albania/Vlore	Ex-situ	Oil
159	Qumështor	Qumësht-I	Albania/Tirane	In-situ	Dual

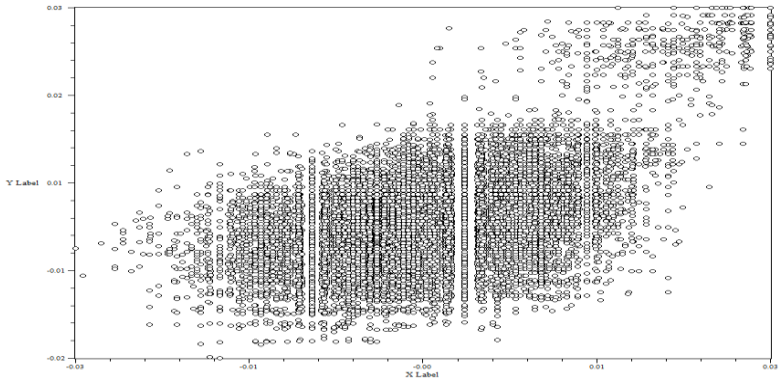
No	Name	Abbrevation	Origin/city	Location	End- use
160	Sant Agostino	Sagost-E	Italy	Ex-situ	Table
161	Santa Caterina	Scater-E	Italy	Ex-situ	Table
162	Santa Caterina	Scater-Ref	Slovenia	Reference Cultivar	Table
163	Simona	Simo-E	Italy	Ex-situ	Oil
164	S-Samo	Samo-Ref	Slovenia	Reference Cultivar	Table
165	Štorta	Storta-Ref	Slovenia	Reference Cultivar	Table
166	Termite di Bitteto	TerDBite-E	Italy	Ex-situ	Dual
167	Toscanina	Toscan-E	Italy	Ex-situ	Oil
168	Ulli I Bardhë I Krujës-1	UBKr I-E	Albania/Kruje	Ex-situ	Oil
169	Ulli I Bardhë I Krujës-2	UBKr II-E	Albania/Kruje	Ex-situ	Oil
170	Ulli I Bardhë I Lezhës	UBLe-I	Albania/Tirane	In-situ	Oil
171	Ulli I Bardhë I Shkodrës	UBShk-I	Albania/Shkoder	In-situ	Oil
172	Ulli I Bardhë I Tiranës-1	UBTr I-I	Albania/Tirane	In-situ	Oil
173	Ulli I Bardhë I Tiranës-2	UBTr II-I	Albania/Tirane	In-situ	Oil
174	Ulli I Bardhë I Tiranës-4	UBTr IV-I	Albania/Tirane	In-situ	Oil
175	Ulli I Bardhë I Tiranës-5	UBTr V-I	Albania/Tirane	In-situ	Oil
176	Ulli I Bardhë I Tiranës	UBTr-E	Albania/Tirane	Ex-situ	Oil
177	Ulli I Bardhë Pobrati	UBPobr-E	Albania/Tirane	Ex-situ	Oil
178	Ulli I Hollë I Himarës-1	UHH I-E	Albania/vlore	Ex-situ	Oil
179	Ulli I Hollë I Himarës-2	UHH II-E	Albania/vlore	Ex-situ	Oil
180	Ulli I Hollë I Himarës-3	UHH III-E	Albania/vlore	Ex-situ	Oil
181	Ulli I Hollë I Himarës-4	UHH IV-E	Albania/vlore	Ex-situ	Oil
182	Ulli I Hollë I Himarës	UHH-I	Albania/Vlore	In-situ	Oil
183	Ulli I Kuq I Damsit	UKDams-I	Albania/Mallakaster	In-situ	Oil
184	Ulli I Kuq UBT	UKUBT-R	Albania/Durres/Kruje	Repository	Oil
185	Ulli I Kuq-1	UK I-E	Albania/Durres/Kruje	Ex-situ	Oil
186	Ulli I Kuq-1	UK I-I	Albania/Tirane	In-situ	Oil
187	Ulli I Kuq-2	UK II-E	Albania/Durres/Kruje	Ex-situ	Oil
188	Ulli I Kuq-2	UK II-I	Albania/Karpan	In-situ	Oil
189	Ulli I Zi -1	UZ I-E	Albania/Durres	Ex-situ	Oil
190	Ulli I Zi -2	UZ II-E	Albania/Durres	Ex-situ	Oil
191	Ulli I Zi	UZ-I	Albania/Tirane	In-situ	Oil
192	Ulli I Zi I Tiranës -1	UZTr I-I	Albania/Tirane	In-situ	Oil
193	Ulli I Zi I Tiranës-2	UZTr II-I	Albania/Tirane	In-situ	Oil
194	Unafka	Unafk-E	Albania/Berat	Ex-situ	Oil
195	Unafka	Unafk-I	Albania/Berat	In-situ	Oil
196	Unafka	Unafk-R	Albania/Berat	Repository	Oil
197	Unknown-1	Unk I-I	Albania/shkoder	In-situ	Dual
198	Unknown-2	Unk II-I	Albania/shkoder	In-situ	Dual
199	Unknown-3	Unk III-I	Albania/Shkoder	In-situ	Dual
200	Vajs Peqini	VajPeq-E	Albania/Elbasan	Ex-situ	Oil

No	Name	Abbrevation	Origin/city	Location	End- use
201	Vajs Peqini	VajPeq-I	Albania/Elbasan	In-situ	Oil
202	ZV-Zelenjak	ZV_Ref	Slovenia	Reference Cultivar	Oil

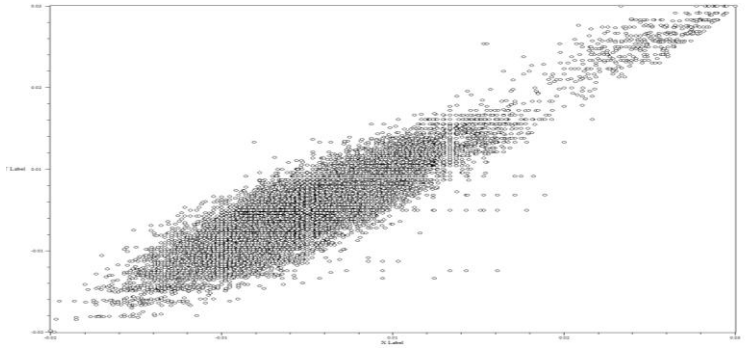
*Each accession is also designated according to an identification code, as a short indication of the given name, the roman number of different sample trees with the same name, and the acronym of the collection location (I=In situ; E=Ex situ; R=Repository; Ref = Reference cultivars

ANNEX D

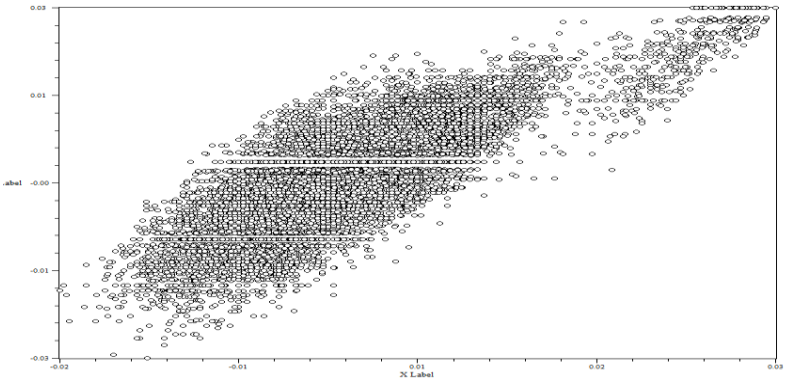
Graphic presentation of correlations between g-SSR and EST-SSR



Comparison of similarity matrices of genomic-derived SSR and EST-SSR



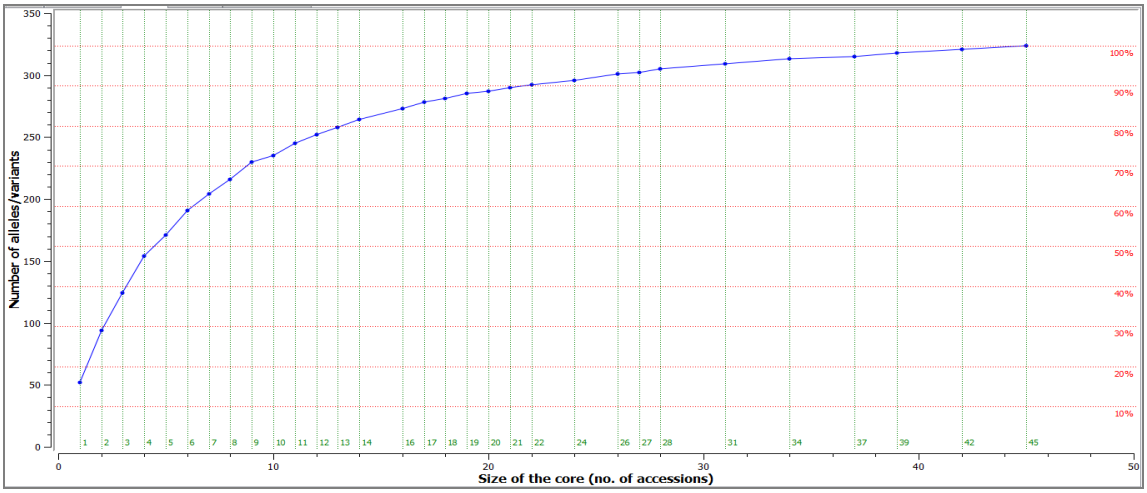
Comparison of similarity matrices of genomic-derived SSR and that generated by using both types of markers (g-SSR and EST-SSR)



Comparison of similarity matrices of EST-derived SSR and that generated by using both types of markers (g-SSR and EST-SSR)

ANNEX E

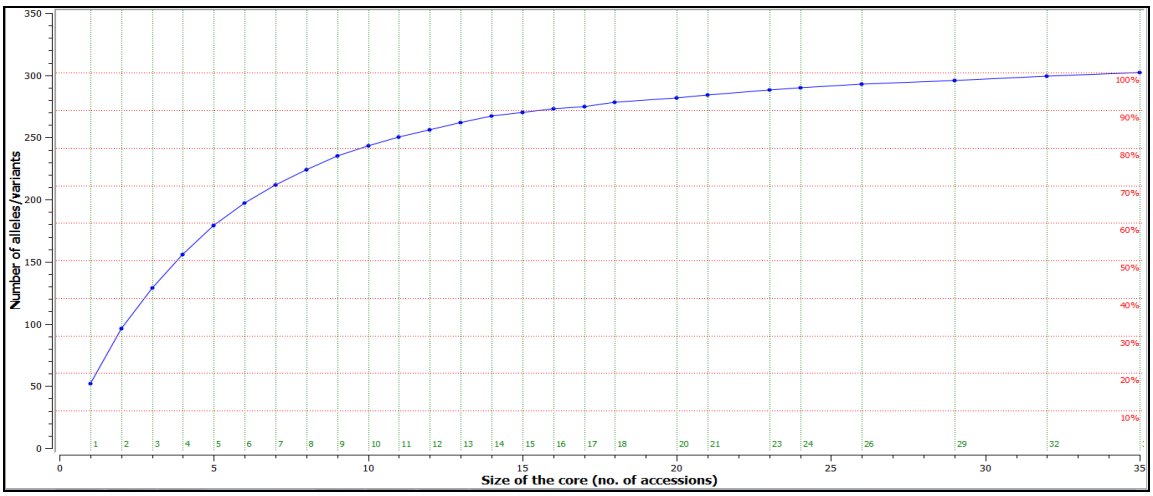
List of genotypes of the *ex situ* and *in situ* core collections



Genetic diversity as a function of the number of accessions included in the *ex situ* core collection

Ex situ core collection representative genotypes	**Origin of olive genotypes**	
	Albanian olive genotypes	**Foreign olive genotypes**
Carboncella-R, Olivastër Kokërr Hollë Elbasani -E, Olivastër e Kuqe Tiranë I-E, Mastoidis-E, Klon16 E, Cellina di Nardo″ E, Managjel Formë Veçantë I-E, Kallmet Kokërr Vogël E, Nolca E, Lukova2 E, Termite di Bitteto E, Sant Agostino E, Ulli i Hollë Himare III-E, Oblonga R, Santa Caterina E, Cucurelia R, Klon13 E, Picholine E, Amfisse E, Ogliarola E, Dolce di Cassano E, Karen E, Nocellara Dell Belice E, Majatica E, Lukova2 R, Passola di Andria E, Itrana E, Cipressino E, Pulazeqin R, Klon4 E, Manzanilla E, Klon3 E, Klon12 E, Unafka R, Toscanina E, Kalinjot Pobrat E, Gordall Sevillana E, Ulli i Bardhë Pobrati E, Cima di Melfi E, Noccellara Messinese E, Boç E, Cazzinicchio E, Klon10 E, Simona E, Managjel I-E	**20**	**25**
	45 genotypes	

Accessions included in the *ex situ* core collection and their origin



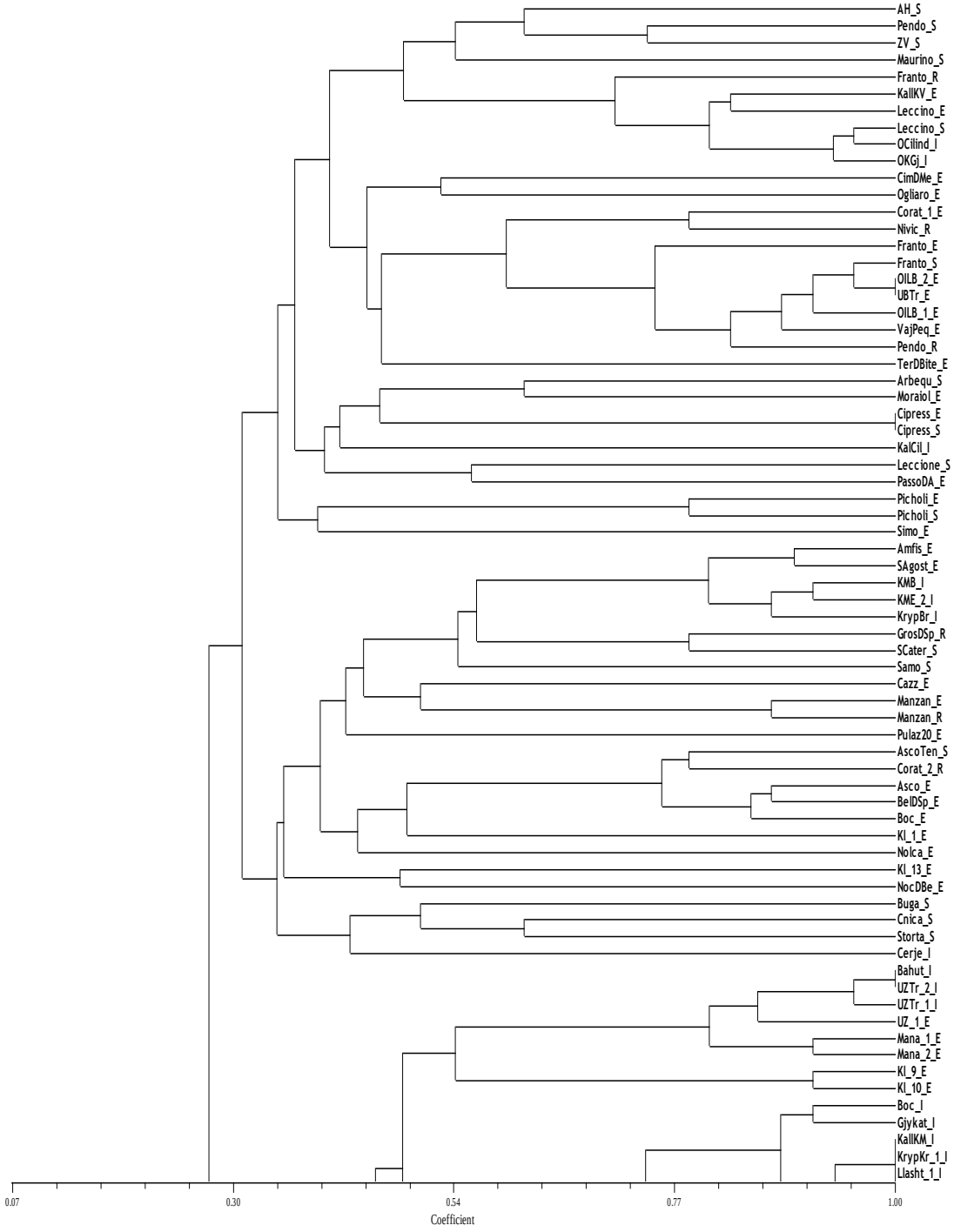
Genetic diversity as a function of the number of accessions included in the *in situ* core collection

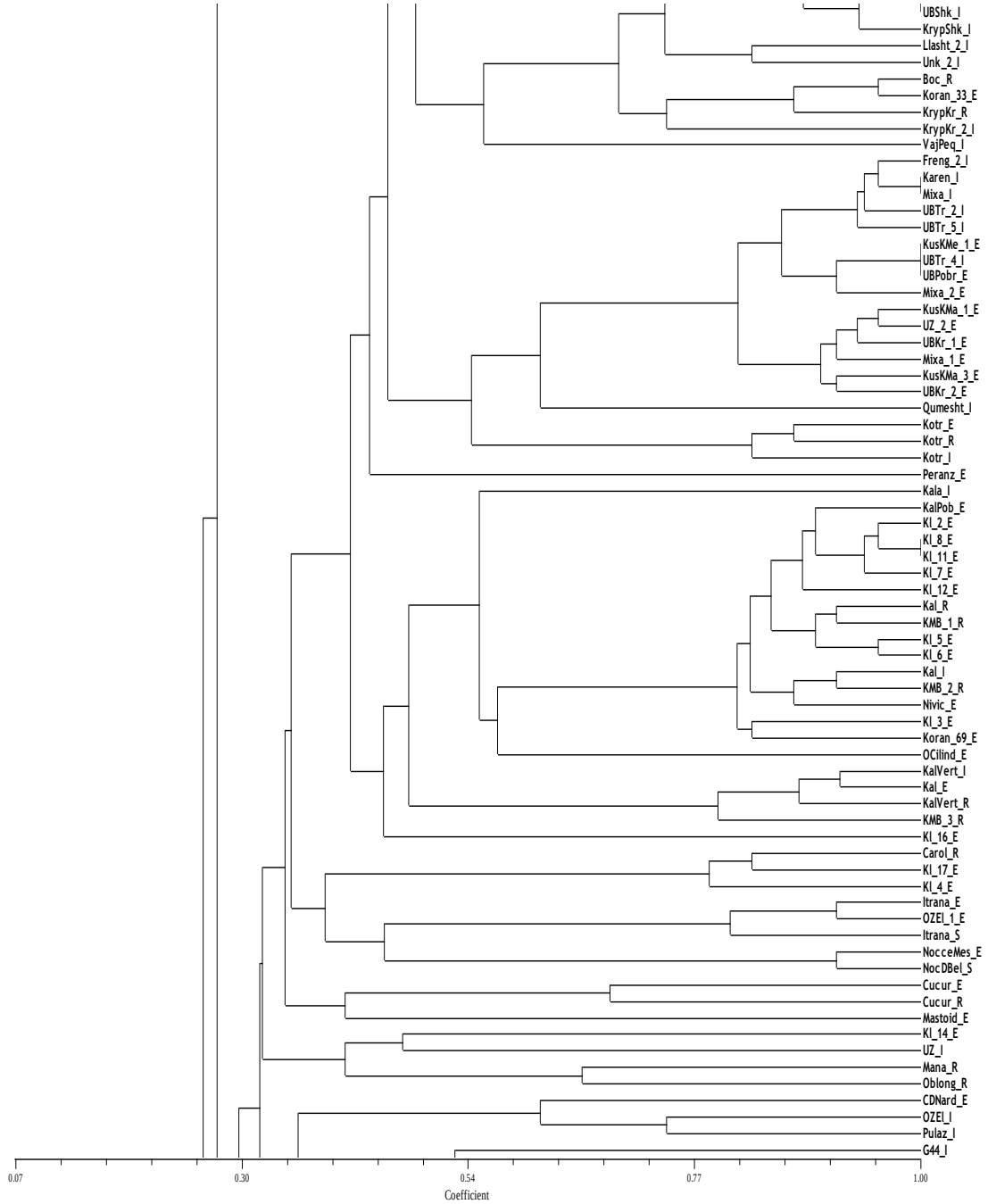
***In situ* core collection representative genotypes**	**Origin of olive genotypes**	
	Albanian olive genotypes	**Foreign olive genotypes**
Oleastër e Vlorës I, Kushan I-I, Kala I, Olivastër 4 I, Unafk I, Micka I,, Olivastër Cilindrike I, G44 I, Çerje I, Kallmet Kokërr Madh I, Perk I, Kallmet Kokërr Vogël I, Unknown II-I, Unknown I-I, Olivastër e Zezë Berati I, Ulli i Hollë Himare I, Pulazeqin I, Ulli i Kuq II-I, Olivastër e Kuqe Tiranë I, Nisiot I, Ulli i Zi I, Marks I, Vajs Peqini I, G45 I,, Kokërr Madh Berati I, Ulli i Bardhë i Tiranës IV-I, Kryps Elbasani I, Olivastër 10 I, Managjel I, Kalinjot Vertikal I, Olivastër Kruje I, Kalinjot Cilindrik I	**32**	**0**

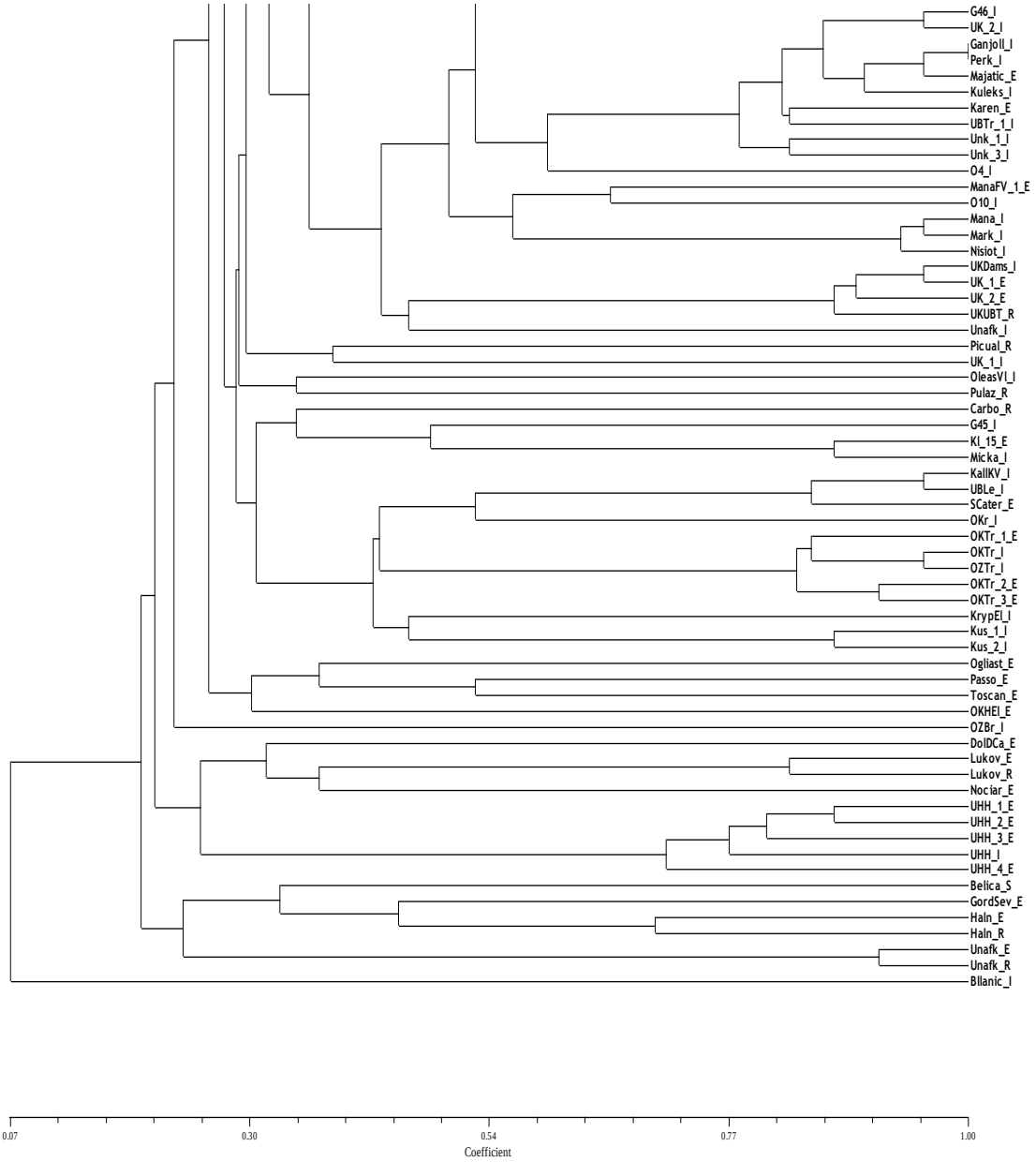
Accessions included in the *in situ* core collection and their origin

ANNEX F

Dendrogram of Albanian olive genotypes and olive cultivars from Slovenian collection based on Dice coefficient and the UPGMA clustering method




UBShk_I
KrypShk_I
Llasht_2_I
Unk_2_I
Boc_R
Koran_33_E
KrypKr_R
KrypKr_2_I
VajPeq_I
Freng_2_I
Karen_I
Mixa_I
UBTr_2_I
UBTr_5_I
KusKMe_1_E
UBTr_4_I
UBPobr_E
Mixa_2_E
KusKMa_1_E
UZ_2_E
UBKr_1_E
Mixa_1_E
KusKMa_3_E
UBKr_2_E
Qumesht_I
Kotr_E
Kotr_R
Kotr_I
Peranz_E
Kala_I
KalPob_E
Kl_2_E
Kl_8_E
Kl_11_E
Kl_7_E
Kl_12_E
Kal_R
KMB_1_R
Kl_5_E
Kl_6_E
Kal_I
KMB_2_R
Nivic_E
Kl_3_E
Koran_69_E
OCilind_E
KalVert_I
Kal_E
KalVert_R
KMB_3_R
Kl_16_E
Carol_R
Kl_17_E
Kl_4_E
Itrana_E
OZEl_1_E
Itrana_S
NocceMes_E
NocDBel_S
Cucur_E
Cucur_R
Mastoid_E
Kl_14_E
UZ_I
Mana_R
Oblong_R
CDNard_E
OZEl_I
Pulaz_I
G44_I
0.07
0.30
0.54
0.77
1.00
Coefficient


G46_I
UK_2_I
GanjolI_I
Perk_I
Majatic_E
Kuleks_I
Karen_E
UBTr_1_I
Unk_1_I
Unk_3_I
O4_I
ManaFV_1_E
O10_I
Mana_I
Mark_I
Nisiot_I
UKDams_I
UK_1_E
UK_2_E
UKUBT_R
Unafk_I
Picual_R
UK_1_I
OleasVI_I
Pulaz_R
Carbo_R
G45_I
KI_15_E
Micka_I
KallKV_I
UBLe_I
SCater_E
OKr_I
OKTr_1_E
OKTr_I
OZTr_I
OKTr_2_E
OKTr_3_E
KrypEl_I
Kus_1_I
Kus_2_I
Ogliast_E
Passo_E
Toscan_E
OKHEl_E
OZBr_I
DolDCa_E
Lukov_E
Lukov_R
Nociar_E
UHH_1_E
UHH_2_E
UHH_3_E
UHH_I
UHH_4_E
Belica_S
GordSev_E
Haln_E
Haln_R
Unafk_E
Unafk_R
Bllanic_I
0.07
0.30
0.54
0.77
1.00
Coefficient