DETERMINATION OF GENETIC DIVERSITY IN NATIONAL OLIVE COLLECTION (*Olea europaea L.*) USING SSR AND SRAP MARKERS AND DEVELOPMENT OF SNP MARKER FOR TRACEABILITY OF MEMECIK OLIVE CULTIVAR

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Molecular Biology and Genetics

by
Neslihan IŞIK

June 2010
İZMİR

We approve the thesis of **Neslihan IŞIK**

Prof. Dr. Anne FRARY

Supervisor

Assoc. Prof. Dr. Bahattin TANYOLAÇ

Committee Member

Assoc. Prof. Dr. Sami Doğanlar

Committee Member

28 June 2010

Assoc. Prof. Dr. Sami Doğanlar
Head of Molecular Biology And Genetics

Assoc. Prof. Dr. Talat Yalçın
Dean of the Graduate
School of Engineering and Science

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor, Prof. Dr. Anne FRARY, for her guidance during my research and study.

I would like to thank all the members of the department of Molecular Biology and Genetics, Institute of Technology and my colleagues for their kindness during my study.

My deepest heartfelt gratitude to my family, my father, my mother and brother for their unflagging love and support throughout my life.

ABSTRACT

DETERMINATION OF GENETIC DIVERSITY IN NATIONAL OLIVE COLLECTION (*Olea europaea L.*) USING SSR AND SRAP MARKERS AND DEVELOPMENT OF SNP MARKER FOR TRACEABILITY OF MEMECIK OLIVE CULTIVAR

In this study genetic diversity of the Turkish olive collection was successfully characterized by using 13 SSR and 12 SRAP markers. For SSR marker analysis, we also added 3 European accessions to the national cultivars and binominal microsatellite data were detected by Qiaxcel software and clustered by NTSYS program. The UPGMA dendrogram showed good fit with the distance matrix (r = 0.85). While the outgroups were clustered with 0.27 similarity to national accessions, Turkish accessions clustered with 0.55 minimum similarity. Using SRAP markers we found 0.66 minimum similarity among Turkish accessions with a good fit between the distance data and UPGMA dendrogram (r = 0.83). The SSR and SRAP distance matrices were compared and they were found to correlated at r =0.23. To support the tree, ordination tests (PCA) for each marker system were performed and similar clustering was seen as in the trees. The other aim of this research was to develop SNP markers for traceability of Memecik accessions. We sequenced the *anthocyanidin synthase* gene of 11 exported olive accessions and aligned using Biolign program. There was no unique SNP that could be used to separate Memecik from the other accessions so we used a combination of SNPs. By using two probes at the same time, Memecik accession was distinguished from the other 10 exported accessions. In addition to find the SNP marker to trace Memecik accession, we developed the SNP marker to trace Trabzon Yağlık accessions.

ÖZET

ULUSAL ZEYTİN KOLEKSİYONUNDAKİ (*Olea europaea L.*) GENETİK ÇEŞİTLİLİĞİN SSR VE SRAP MARKÖRLERİ KULLANILARAK BELİRLENMESİ VE MEMECİK ZEYTİN ÇEŞİDİNİN İZLENEBİLİRLİĞİ İÇİN SNP MARKÖRLERİNİN GELİŞTİRİLMESİ

Bu çalışmada, 13 SSR ve 12 SRAP işaretleyicisi kullanılarak Türk zeytin koleksiyonunu tanımlanmıştır. SSR işaretleyicilerine dayalı çeşitlilik çalışmada 3 yabancı çeşitte eklenmiştir ve genetik çeşitlilik için iki terimli data Qiaxcel programında tespit edilmiştir ve NTSYS programında gruplanmıştır. Yabancı çeşitler, ulusal çeşitlerle 0.27 benzerlik ile gruplanırken Türk çeşitleri 0.55 benzerlikle gruplanmıstır ve 0.85 R değeri vermiştir. SRAP işaretleyicileri kullanılarak, Türk çeşitleri arasında 0.83 korelasyon ile minimum benzerlik 0.66 bulduk. Elde edilen 2 matrix (SSR ve SRAP matrixleri) karşılaştırıldı ve benzerlikleri 0.23 olarak bulundu. Çizilen filogenetik ağacı desteklemesi için ordination testleri (PCA) her iki marker sistemi için de gerçekleştirilmiştir ve ağaçlarında görülen ile aynı gruplanmıştır. Bir diğer amacımız Memecik çeşidinin izlenebilirlik testi için SNP işaretleyicisi geliştirmekti. İhraç edilen 11 zeytin çeşidinin *Anthocyanidin synthase* genleri dizilendi ve Bioligm program ile sıralandı. Memecik çeşidini bir kere de ayıran tek bir SNP bulunamadı bu yuzden ard arda kullanılmak üzere SNP kombinasyonu yapıldı. Bu sebeple aynı anda 2 prob kullanılarak Memecik çeşidi diğer ihraç edilen 10 çeşitten ayrıldı. Memecik çeşidini izlemek için SNP markörü bulmanın yanısıra Trabzon Yağlık Çeşidini izlemek için SNP markörü geliştirdik.

TABLE OF CONTENTS

LIST OF FIGURES viii

LIST OF TABLES ix

CHAPTER 1. INTRODUCTION 1

1.1. General Information Olea europapaea L. Trees in Turkey 1

1.2. Distribution of Olea europapaea L. Trees in Turkey 2

1.3. Differences in Turkish Olive Accessions 4

1.3.1. Morphological Characteristics of Turkish Olive cultivars 4

1.3.2. Molecular Genetic Approaches to Discriminate Turkish Olive Accession 9

1.3.2.1. Simple Sequence Repeat Approaches 9

1.3.2.2. SRAP Molecular Markers 10

1.3.2.3. Single Nucleotide Polymorphisms, SNPs 11

1.3.2.3.1. Application of SNPs on Plant Breeding 13

1.3.2.3.2. Allelic Discrimination Methods, for SNPs 14

1.4. Goals 15

CHAPTER 2. MATERIALS AND METHODS 17

2.1. Materials 17

2.1.1. Olive Materials 17

2.1.2. SSR Primers 19

2.1.3. SRAP Primers 19

2.1.4. Ant and Cyc Primers for SNP Detection 20

2.2. Methods 20

2.2.1. DNA Extraction 20

2.2.1.1. DNA Isolation from Olive Oil 20

2.2.1.2. DNA Extraction from Leaves 21

2.2.2. Data Analysis 21

2.2.2.1. PCR Analysis of SSR Primers and Capillary Electrophoresis 21

2.2.2.2. SRAP Marker Based Analysis 22
2.2.2.3. Sequence Analysis 22
2.2.2.4. Real time-PCR Optimization 24

CHAPTER 3. RESULTS AND DISCUSSION 25
3.1. DNA extraction 25
3.2. Genotyping Using Molecular Markers 26
3.2.1. SSR Primers Based Characterization on Capillary
Electrophoresis 26
3.2.1.1. Genetic Similarity Test Using SSR Primers 28
3.2.2. Genotyping Using ORF-Based Markers, SRAPs 32
3.2.2.1. Genetic Similarity Test Using SRAP Markers 33
3.2.3. Genotyping Using Both SRAP and SSR Markers 36
3.2.3.1. Genome Searching for Memecik-Specific SNPs 39
3.2.3.2. Genotyping Using the Designed Four Probes on
LightCycler 480 47

4. CONCLUSIONS 49

REFERENCES 54

LIST OF FIGURES

Figure **Page**

Figure 1.1. HPLC chromatogram of EVOO of harvest year at 280 nm........................... 8
Figure 1.2. SNP analysis methods on the LightCycler®480 System.............................. 15
Figure 3.1. Capillary electrophoresis result for udo24 marker of one individual........... 27
Figure 3.2. Gel image of polymorphic SSR marker on capillary electrophoresis 27
Figure 3.3. Dendrogram was drawn based on SSR markers of Turkish olive cultivars. 31
Figure 3.4. PCA analysis of SSR based data 32
Figure 3.5. Dendrogram of Turkish olive cultivars, based on SRAP primers 34
Figure 3.6. PCA analysis, 2-D plot of SRAP primer based data 35
Figure 3.7. 2-D of SSR and SRAP based data 37
Figure 3.8. Dendrogram of Turkish olive cultivar using SSR and SRAP markers 38
Figure 3.9. Complete illusturation of SNPs on ant. Gene 40
Figure 3.10. Probe&primer sequences and illustration on the gene 45
Figure 3.11. Real time PCR results of probe42 47

LIST OF TABLES

Table **Page**

Table 1.1. Leaf features of Turkish olive accessions 5

Table 1.2. Flower features of Turkish olive accessions 6

Table 1.3. Fruit features of Turkish olive accessions 7

Table 1.4. SNPs functional classes 12

Table 1.5. SNP databases 13

Table 2.1. Turkish Olive Cultivars and outgroup genotype use and localization list 17

Table 2.2. SSR primers used to characterize genetic diversity in olive 19

Table 2.3. Sample names and their regions, used for sequence analysis 23

Table 2.4. Conditions of Melting Curve Anaysis 24

Table 3.1. DNA concentration of samples 25

Table 3.2. PIC values for each SSR primers 28

Table 3.3. SRAP data PIC value table 33

Table 3.4. Sequence results based – nucleotidic substitutions and their
localization for each genotype 43

Table 3.5. Probe-Tm based nucleotide analysis 48

Table 3.6. Real time PCR results and sequence results of the probe
binding regions 49

Table 3.7. Genotypes whcih have the same alleles .... .50

CHAPTER 1

INTRODUCTION

1.1. General Characteristics of *Olea europaea* L.

Olive, *Olea europaea* L., is one of the oldest known cultivated fruit tree crops and according to archeological evidence, it was domesticated 5.000 years ago in Palestine. By the third millenium B.C., olive had become a significant crop in the Eastern Mediterranean (Zohary and Spiegel-Roy 1975). Olive oil was valued during ancient times and is still values today because of its health benefits. Olive oil is preferred by many people for its high contents of monounsaturated fatty acids (oleic acid) (Grigg 2001), antioxidants such as carotenoids and phenolic compounds (Visioli, et al. 2002). Researches are also determined the reducing affect of first myocardial (Panagiotakos, et al. 2000). Today, approximately 17.5 million metric tons (Mt) of olives and 2.8 million Mt of oil are produced yearly on 9.3 million hectares of land worldwide (Faostat 2007) mostly in the Mediterranean Region but also in Australia, Canada, China, Peru, Chile, Argentina and United States.

Olive is an outcrossing diploid species (2n=46) which is self incompatible depending on the cultivar (Cuevas 2001). Olive cultivars are widely produced by vegetative propagation so they are uniform but very heterozygous (Wu, et al. 2004). Therefore, traditional cultivars are still grown in most locations and very few new olive varieties have been bred.

Olive cultivars and wild accesions are usually distinguished by their morphological and agronomic features such as fruit size, oil content and growth habit. However, such features can be affected by the environment. Utilization of these characters also requires that the plant reach reproductive age. Moreover, many wild and feral types are morphologically very similar to cultivars. Vast numbers of olive trees are grown in the world without true labelling because of the difficulties of characterization of accessions. This results in problems such as mislabelling in nomenclature; homonyms, synonyms and misidentity in olive (Barranco and Rallo 2000). Such types of problems can be resolved accurately by using DNA markers.

Accurate and easy identification of cultivars is necessary to manage breeding, to create variability and identify Turkish cultivar to the world for trade objectives and variety protection. Therefore, identification of olive cultivars is important not only to increase biodiversity in olive but also it is commercially important for export or import of olive cultivars and their olive oil. Labelling is an important issue in trade. Labelling has to be done correctly especially for export. Labelling matters include; location of olive trees, their accessions and their acidity. To determine such characteristics, we first need to know the features correctly. In Turkey, labelling is currently verified by morphological testing. However, in Europe, molecular marker systems such as SSR (simple sequence repeat) and SRAP (sequence-related amplified polymorphism) marker systems (fingerprinting) are used for identification of olive accessions. Both of them are important in terms of easy to use and cheap. However neither SRAP nor SSR makers are as sensitive as new generation marker system like SNPs. SNPs are single nucleotide polymorphisms. These mutations are inherited from parents to individuals and SNP polymorphisms are analyzed to distinguish accessions for labelling. Also separation of accessions is important so that the desired cultivation of the olive can be done to increase quality and yield.

1.2. Distribution of Olea europaea L. Trees in Turkey

In our country, olive trees can grow in 35 locations in the Karadeniz, Marmara, Aegean, Mediterranean and Southeast Anatolia Regions (Canozer 1991). Producers in Turkey and the world grow *Olea europea subsp. sativa* of Olea species. Each region of Turkey where olives are grown must have certain climatic and environmental features. For example, olives are only found in regions of the Karadeniz which are not very windy and have microclimate features suitable for production of olive trees. Therefore, only 0.6% of olive trees in Turkey are grown in this region and 0.5% of total yield comes from subcultivars such as Butko, Görele, Marantelli, Pastos, Otur, Salamuralık, Tuzlamalık, and Yağlık which are cultivated in this region.

In the Marmara Region, olive cultivation can only be done sheltered places. In this region, people prefer to grow and trade olive for its fruit instead of producing olive oil, so this region is important for the edible olive. Thus, 84% of total production in this region is for eating while the remaining low quality olives are processed for olive oil.

The Marmara Region has 9% of total trees and it provides 7% of the total olive production in Turkey. The most important accession in the Marmara Region is known as Gemlik which has black fruit and this accession constitutes 80% of all accessions. The other important accessions are Çelebi, Edincik Su, Karamürsel Su, and Samanlı with other cultivars including Beyaz Yağlık, Çizmelik, Erdek Yağlık, Eşek Zeytini, Şam and Siyah Salamuralık.

The Aegean Region is the most important area for olive production and it has 14% of the olive trees in Turkey. This region, supplies 75% of all olives in Turkey with 86% of olives used in oil production and the rest as table olives. Important accessions in the Aegean Region are Ayvalık, Çakır, Çekişte, Çilli, Domat, Erkence, İzmir Sofralık, Kiraz, Memecik, Memeli and Uslu. Other olive accessions which are cultivated in this region are Ak Zeytin, Aşı Yeli, Dilmit, Eşek Zeytini, Girit Zeytini, Hurma Kaba, Hurma Karaca, Kara Yaprak, Taşarası, Yağ Zeytini and Yerli Yağlık.

In the Akdeniz Region, vegetable cultivation is more valuable than olive and olive oil production, nevertheless 10% of total Turkey's olive trees are found in this region and also it provides 12% of table olives. In this region, 68% of produced yield is used for olive oil production. Büyük Topak Ulak, Halhalı, Sarı Haşebi, Sarı Ulak, Saurani, Karamani, Küçük Topak Ulak and Sayfi are the names of olive accessions commonly grown in the Akdeniz Region.

The Southeastern Region of Turkey is exposed to the Akdeniz climate and 73% of total olive production in this region is used for oil production. This region is also known as a motherland of olive and it has a rich number of accessions. Significant accessions are Eğriburun, Kalembezi, Kan Çelebi, Kilis Yağlık, Nizip Yağlık and Yağ Çelebi. Other important accessions are Belluti, Halhalı Çelebi, Hamza Çelebi, Hırmalı Çelebi, Hursuki, İri Round, Mavi, Melkabazı, Tesbih Çelebi, Yağlık Çelebi, Yağlık Sarı Zeytin, Round Çelebi, Round Halhalı, Yün Çelebi and Zoncuk.

1.3. Differences in Turkish Olive Accessions

1.3.1. Morphological Characteristics of Turkish Olive Cultivars

Although olive trees are called the 'wealthy plant of poor land', it requires good conditions to get high yield in olive cultivation. For economical growth, soil characteristics, climate demands and temperature features have to be proper and meet the olive trees' needs (Canozer 1991). In terms of soil features some factors limit olive cultivation such as bedrock level, soil without oxygen and heavy soil. Ideal soil for olive breeding should be deep, calcareous, sandy-loamy, loamy or argilaceous-loamy. These kinds of soils are good for root development because they supply water permeability, water retention and aeration to plant. Purely sandy or argilaceous and humid soil is risky for olive breeding. Climate requirements of olive trees are also discriminating. Although olive trees can survive in an undesired climate, they cannot give adequate yield in such conditions. The natural cultivation environment of olive trees is at 30-45° northern-southern latitude. The Mediterranean climate zone in this latitude is best for olive growth, especially when olive is grown inshore of this region. Olives can also survive at more than 500m above sea level (Canozer 1991). Leaf, flower and fruit features of Turkish olive accessions have been determined and these data give one way to discriminate the Turkish olive accessions (Tables 1, 2 & 3) (Cirik 1988). In comparison with other Turkish olives, Memecik is better than Domat, which is another popular local accession, in terms of oil content. Also, Memecik olives have the best flesh value among commonly grown olives such as Ayvalık, Domat, Erkence, Gemlik and Uslu. Oil content and value of flesh of Turkish olive cultuvars is illustrated in Table 1.3.

Table 1.1. Leaf features of Turkish olive accessions
(Source: Cirik 1988)

Accession name	Leaf length(mm)	Leaf width(mm)	Leaf ratio (w/l)	Stalk length (mm)	Leaf shape
Ayvalık	62.54	11.84	5.28	4.50	Long, narrow
Büyük Topak Ulak	65.21	12.01	5.42	4.70	Long, narrow, elliptic
Çakır	52.14	11.64	4.47	3.50	Mid-long, mid-broad, elliptic
Çekişte	54.02	9.18	5.88	3.70	Mid-long, narrow, elliptic
Çelebi	52.58	12.56	4.18	4.70	Mid-long, broad, elliptic
Çilli	63.06	11.98	5.26	4.00	Long, narrow, elliptic
Domat	68.58	12.20	5.61	4.00	Extra-long, narrow, elliptic
Edincik Su	62.98	11.72	5.37	3.80	Long, narrow, elliptic
Eğriburun	55.81	12.88	4.33	3.90	Mid-long, broad, elliptic
Erkence	45.56	11.86	3.84	3.20	Short, broad, elliptic
Gemlik	50.68	11.84	4.28	3.40	Short, broad, elliptic
Halhalı	63.18	11.82	5.34	3.21	Long, narrow, elliptic
İzmir Sofralık	71.00	11.60	6.12	6.50	Extra -long, narrow, elliptic
Kalembezi	58.50	12.86	4.54	4.50	Mid-long, midbroad, elips
Kan Çelebi	67.62	10.96	6.16	5.30	Extra-long, extra-narrow, elliptic
Karamursel Su	66.52	12.10	5.49	3.80	Extra-long, narrow, elliptic
Kilis Yağlık	57.46	13.18	4.35	3.80	Mid-long, mid-broad, elliptic
Kiraz	61.42	12.32	4.98	4.00	Long, mid-broad, elliptic
Memecik	53.70	10.84	4.95	3.90	Mid-long, mid-broad, elliptic
Memeli	57.38	11.08	5.17	3.60	Mid-long, mid-broad, elliptic
Nizip Yağlık	58.48	14.06	4.25	3.70	Mid-long, broad, elliptic
Samanlı	66.44	15.32	4.33	5.50	Extra-long, broad, elliptic
Sarı Haşebi	52.22	12.50	4.17	3.70	Mid-long, broad, elliptic
Sarı Ulak	53.38	13.38	3.98	3.10	Mid-long, broad, elliptic
Saurani	52.73	12.32	4.28	3.80	Mid-long, broad, elliptic
Tavşan Yüreği	59.60	9.20	6.47	4.30	Long, extra-narrow, elliptic
Uslu	59.60	12.40	4.80	3.90	Long, mid-broad, elliptic
Yağ Çelebi	52.84	12.10	4.36	3.60	Mid-long, broad, elliptic

Table 1.2. Flower features of Turkish olive accessions
(Source: Cirik 1988)

Accession name	Inflorescence length(m)	Aver.inflorescence length(mm)	Inflorescence flower number	Aver.flower number	Flowering time
Ayvalık	20-42	29.9	14-34	20	16.may- 09.jun
Büyük Topak Ulak	25-32	26.8	7-18	12	12.may- 09.jun
Çakır	7-26	20.20	5-15	8	13.may- 09.jun
Çekişte	22-35	26.70	8-29	16	13.may- 09.jun
Çelebi	21-42	30.50	16-42	24	12.may- 09.jun
Çilli	16-25	20.60	8-15	11	13.may- 04.jun
Domat	14-36	24.6	8-27	16	16.may- 06.jun
Edincik Su	14-23	20.50	7-21	12	12.may- 12.jun
Eğriburun	26-33	20.20	10-28	20	16.may- 04.jun
Erkence	14-28	19.90	6-20	11	13.may- 09.jun
Gemlik	15-26	20.30	10-23	14	12.may- 09.jun
Halhalı	21-35	26.20	9-25	12	16.may- 03.jun
İzmir Sofralık	18-33	24.30	1-17	10	16.may- 09.jun
Kalembezi	18-27	24.00	11-23	15	16.may- 14.jun
Kan Çelebi	24-45	38.00	11-28	19	15.may- 04.jun
Karamursel Su	21-41	31.90	10-37	24	16.may- 12.jun
Kilis Yağlık	20-40	29.70	14-30	20	20.may- 04.jun
Kiraz	33-45	37.90	14-27	20	20.may- 09.jun
Memecik	24-35	29.90	6-15	11	16.may- 06.jun
Memeli	17-28	22	10-24	17	21.may- 09.jun
Nizip Yağlık	14-31	24.40	10-30	17	15.may- 15.jun
Samanlı	22-40	30.40	5-40	22	16.may- 09.jun
Sarı Haşebi	15-21	18.60	6-16	11	13.may- 15.jun
Sarı Ulak	16-27	22	1-28	17	15.may- 14.jun
Saurani	13-18	14.80	8-15	10	16.may- 09.jun
Tavşan Yüreği	14-22	22.70	8-15	11	20.may- 04.jun
Uslu	24-34	28.90	6-19	11	12.may- 04.jun
Yağ Çelebi	24-36	31.50	6-36	18	16.may- 09.jun

Table 1.3. Fruit features of Turkish olive accessions
(Source: Cirik 1988)

Accession Name	100 fruit weight (gr)	100 fruit volume (cm3)	Fruit number per kg.	% Flesh	% Oil	% Humidity	Fruit shape
Ayvalık	363.8	360	274	85.26	24.72	55.74	Nearly-round, cylindrical
Büyük Topak Ulak	484.3	480.5	206	88.31	20.20	52.02	Round
Çakır	284.3	271.6	352	86.88	23.62	50.28	Pear-shaped
Çekişte	542	530	185	85.23	26.89	50.40	Oval
Çelebi	710.8	698.2	141	86.95	28.38	41.05	Long, cylindrical
Çilli	490	470	204	88.97	20.55	53.11	Nearly-round, oval
Domat	530.3	525.8	189	83.76	20.57	55.89	Cylindrical
Edincik Su	494.2	476.8	202	89.41	16.71	61.16	Round
Eğriburun	258	250	388	86.17	20.84	50.23	Oval
Erkence	303.6	283	329	86.17	25.36	54	Oval
Gemlik	372.8	370	268	85.86	29.98	45.05	Nearly-round, oval
Halhalı	383.01	370.5	261	82.79	21.11	49.03	Nearly-round, oval
İzmir Sofralık	750.2	1010	133	87.55	20.16	52.07	Oval
Kalembezi	222	225	450	84.29	31.50	46.16	Nearly-round
Kan Çelebi	615.1	620	163	88.90	16.90	52.03	Round
Karamursel Su	710	680	141	87.07	18.60	59.20	Oval
Kilis Yağlık	176.8	180	566	82.25	31.82	40.79	Round
Kiraz	412.5	392	242	85.76	19.76	56.90	Round
Memecik	478	465	209	88.28	24.50	52.60	Oval
Memeli	463.8	460	216	88.57	20.20	58.55	Oval
Nizip Yağlık	217.6	200	460	81.31	27.36	40.49	Nearly-round, cylindrical
Samanlı	396.4	390.2	252	84.51	20.77	52.59	Round
Sarı Haşebi	293.8	274.6	340	85.53	24.72	50.37	Oval
Sarı Ulak	376.52	330	266	71.85	18.84	52.86	Long, cylindrical
Saurani	295.9	275.6	338	86.61	29.18	57.51	Cylindrical
Tavşan Yüreği	608.17	595.8	164	86.43	20.20	56.50	Heard-shape
Uslu	353.4	340	283	85.17	21.50	60.61	Oval
Yağ Çelebi	442.15	430	226	84.57	21.10	53.40	Long, oval

In research, olives and olive oil have been found to have a large variety of chemical compounds such as gallic acid, hydroxytyrosol, 2,3dihydroxybenzoic acid, tyrosol, 4hydroxybenzoic acid, 4hydroxyphenylacetic acid, vanillic acid, caffeic acid, vanillin, p-coumaric acid, ferulic acid, cinnamic acid, luteolin and apigenin (Ocakoglu

2007). All of these compounds are antioxidants which are essential for human health in terms of protecting humans from cancer (Visioli, et al. 2002) and aging (Tuck, et al. 2002). Memecik olives contain high levels of hydroxytyrosol, luteolin, apigenin, p-coumaric acid, cinnamic acid (Ocakoglu 2007). In these respects, Memecik is clearly beter than Gemlik, Ayvalık, Edremit and Nizip. Figure 1 illustrates the chemical composition of common Turkish olive oils. Approximate chemical composition values of Memecik accession can also be similarly obtained with Erkence (Ocakoglu 2007). The same increased luteolin and apigenin levels were detected in most of the Spanish, Italian and Portugues virgin olive oils (Vinha, et al. 2005)

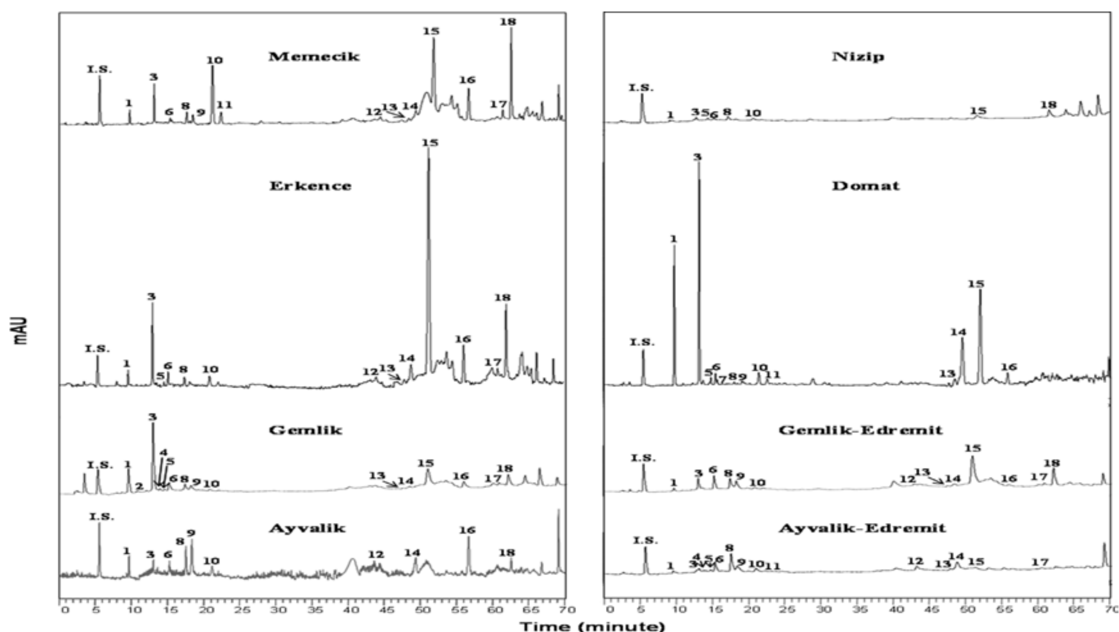

Figure 1.1 HPLC chromatogram of EVOO of harvest year at 280 nm. (IS) gallic acid; (1) hydroxytyrosol (Hyt); (2) 2,3dihydroxybenzoic acid (Dba); (3) tyrosol (Tyr); (4) 4 hydroxybenzoic acid (Hdba); (5) 4 hydroxyphenylacetic acid (Hpha); (6) vanillic acid(Va); (7) caffeic acid (Ca); (8) vanillin (Val); (9) unidentified; (10) p-coumaric acid (Pcoa); (11) ferulic acid (Fa); (12) unidentified; (13) unidentified; (14)unidentified; (15) cinnamic acid (Cina);(16) luteolin (Lut); (17) unidentified; (18) apigenin (Apg) (Source: Ocakoglu 2007).

1.3.2. Molecular Genetic Approaches to Discriminate Turkish Olive Accessions

Molecular markers which reveal polymorphisms at the DNA level, have been playing an increasing role in molecular breeding. Types of DNA markers include bi-allelic dominant markers such as RAPDs (random amplification of polymorphic DNA) (Fabbri, et al. 1995) and AFLPs (amplified fragment polymorphism); bi-allelic co-dominant markers, such as RFLPs (restriction fragment length polymorphisms); and multiallelic co-dominant markers, such as microsatellites and SRAPs (sequence-related amplified polymorphism), (Li and Quiros 2001). DNA markers are the most widely used markers in molecular breeding. In addition to the above markers, a sequence-based marker system, Single Nucleotide Polymorphisms (SNPs) is very informative. SNPs are in most cases bi-allelic, co-dominant markers (Kwok 2001).

1.3.2.1. Simple Sequence Repeat Approaches

SSR primers are evolutionary conserved sequences which are designed to amplify simple sequence repeat or microsatellite regions. These microsatellites are composed of 2 to 5 bp repeated nucleotides and these main motifs are repeated 9-30 times. A small genomic library of *Olea europaea* L., immensely enriched in (GA/CT)*n* repeats was used to identify SSRs in olive (Weising, et al. 1998).

SSR markers are efficient markers which can be used to identify and differentiate accessions (Carriero, et al. 2002). SSR markers' efficiency is based on their high polymorphism contents, codominance, ease of detection and repeatability among researchers (Aranzana, et al. 2003). Therefore, to summarize the advantages of SSR markers, they are easy to use, show Mendelian inheritance and are highly informative. In addition, a great number of SSR primer pairs are publicly available and SSRs are cost effective per genotype.

Primer pairs are generally 15-31 DNA bases in length and PCR products are 90-300 base pairs, depending on the SSR primers. After amplification, separation using capillary electrophoresis is more effective because of its highly sensitive resolution which can be used to detect 3-5 bp differences which cannot be seen using agarose gel electophoresis.

SSR markers are used for comparing germplasm with other competitors by seed companies to protect their yield specificity, to develop core germplasm collections to avoid high storage costs and to check homogeneity (uniformity) among individuals or to increase biodiversity and retain the resistance capacity of plants. In olive, SSR markers have been used to differentiate some olive varieties. According to Tunisian researchers, SSR markers were useful for finding mislabelling and eliminating redundacy in molecular breeding studies (Taamalli, et al. 2008).

Hitherto, classical agarose electrophoresis was used to determine SSR locus size. Recently, capillary electrophoresis systems are used for size separation and quantification of DNA fragments. Thanks to the negative charge of nucleic acids, each sample is loaded automatically into individual capillary by positively charging the capillary terminus. The working procedure of capilary systems is closely related to agarose gel electrophoresis. They both separate the alleles based on the principle that low molecular weight molecules migrate faster than high molecular weight molecules. In contradiction to agarose electrophoresis, the capillary system has a detector that detects and measure the signal of the DNA molecules passing through the capillary. The emission signal data collected from the photomultiplier detector is then converted to electronic data (electropherogram and a gel image) by BioCalculator software. We preferred the capillary system by reason of its a number of advantages over traditional gel electrophoresis such as higher detection sensitivity, fast and automated analysis of up to 96 samples.

1.3.2.2. SRAP Molecular Markers

SRAP markers were developed (Li and Quiros 2001) to amplify the open reading frames (ORFs) in *Brassica oleracea*. This system is based on a two primer system, each of which is 17-21 nucleotide in length. Primers consist of core sequences, three selective nucleotides and the core sequences contain ‘filler’ sequences (non-specific constitution) and specific sequences. The forward primer is 17 bp (10 bp filler + CCGG + three selective nucleotides) while the reverse primer is 18 bp (11 bp filler + AATT + three selective nucleotides). They can pair with exons and promoters or introns, respectively. Polymorphisms are the result of different lengths of introns or promoters among individuals. SRAPs generate a large number of polymorphic

fragments in each reaction, are simple to use, are applicable to any species and are highly reproducible. These markers have been applied to several crops (Li and Quiros 2001) such as potato (*Solanum tuberosum*), rice, lettuce (*Lactuca sativa*), Chinese cabbage (*Brassica rapa*), rapeseed (*Brassica napus*), garlic (*Allium sativum*), apple (*Malus domestica*), *Citrus*, and celery (*Apium graveolens*).

1.3.2.3. Single Nucleotide Polymorphisms, SNPs

A SNP is technically defined as a single nucleotide variation at a specific location which is found in more than 1% of the population. Therefore, SNPs do not include insertion or deletion polymorphisms. However, in practice this definition is not applied correctly, some biallelic variations, including insertions, deletions and variations with less than 1% allele frequency, are called SNPs (Brookes, et al. 1999).
Both SNP markers and SNP marker sets are developed using a single base change in DNA sequences, which is usually an alternative of two possible nucleotides at a defined position. These kinds of single base differences between homologous DNA fragments can be near or in a gene, so their effect on function may be difficult to illustrate. SNPs can change the function of DNA, RNA and proteins. Therefore, they can affect protein amino acid sequence synonymously or non-synonymously.

SNPs are divided into 2 main group SNPs can result in change of protein function or change only in structure if they occur in regulatory regions which affect the gene expression level or translation of a gene product. Also, SNPs are not only found in regulatory regions which result in vital change in gene or protein level but also another kind of SNPs are called as a non-synonnymous SNPs found in intronic region which are informative and important evolutionary and the results are used in traceability or origin studies. SNPs are given different names based on their effect on protein function. SNP classes are shown in Table 1.5 (Mooney 2005).

Table 1.4. SNPs functional classes

Coding SNPs	cSNP	Positions that fall within the coding regions of genes
Regulatory SNPs	rSNP	Positions that fall in regulatory region of genes
Synonymous SNPs	sSNP	Positions in exon that do not change the codon to substitude an amino acid
Non-synonymous SNPs	nsSNP	Positions that incur an amino acid substitution
Intronic SNPs	iSNP	Positions that fall within introns

After the discovery of millions of DNA sequence variants of the human genome in the Human Genome Project, SNPs were noticed as an important source of genetic variation (Collins 1997). At the end of 2000, over 1,5 million SNPs were found as a result of this project. These resources have been commonly used in population genetics, evolutionary genomics, pharmacogenomics and association studies. In this way the molecular basis of many genetic diseases were illustrated (Reale, et al. 2006). Several examples of SNP databases are shown in Table 1.4. According to NIH, there are more than 24,5 million candidate SNPs in the human genome (www.ncbi.nlm.nih.gov) Because SNP markers are abundant in the genome and inherited, SNPs are a very popular and powerful DNA marker system (Chakravarti 1998). Understanding the effects of genetic differences such as amino acid changes in the genome is possible using single nucleotide polymorphisms (SNP). Finding the polymorphisms among accessions is based on sequence comparison and after that a detection strategy is used.

Table 1.5. SNP databases

URL		Comments
Genome resources		
dbSNP	http://wwww.ncbi.nlm.nih.gov/projects/SNP	SNPs from the complete genome
HapMap Consortium	http://hapmap.org/cgi-perl/gbrowse	Whole genome SNPs in four population
GoldenPath	http://genome.ucsc.edu/	Genome database
JSNP	http://snp.ims.u-tokyo.ac.jp/	Common SNPs within Japanese population
Mutation repositories		
HGMD	http://www.hgmd.org/	Mutation database with many annotations
Swiss-Prot	http://us.expasy.org/	Protein database with extensive variant annotations
HGVbase	http://gvsbase.cgb.ki.se/	SNPs from the complete genome
	PharmGKB http://www.pharmgkb.org	Genes involved in drug metabolism
SNP500Cancer	http://snp500cancer.nci.nih.gov/home..1.cfm	Genes involved in cancer
Tools		
SNPper	http://snpper.chip.org/	Novel software for SNP analysis
BioPerl	http://www.bioperl.org/	A programming application program interface (API) for bioinformatics analysis

1.3.2.3.1. Application of SNPs in Plant Breeding

SNPs are becoming of more interest in agricultural breeding programmes as high-throughput methods (Gupta, et al 2008). SNP method is used in plants for many molecular genetic marker applications. These applications include high-resolution genetic map construction, linkage disequilibrium-based association mapping, genetic diagnostics, genetic diversity analysis, cultivar identification, phylogenetic analysis, and characterization of genetic resources (Rafalski 2002). The use of SNPs will become more widespread with the increasing availability of crop genome sequence, the reduction in cost, and the increased throughput of SNP assays. Genetic diversity studies in maize were performed using SNPs at 21 loci along chromosome 1 (Tenaillon, et al. 2002). This study facilitated an understanding of the forces contributing to genetic diversity in maize. Similar approaches have been used for cultivar identification in

malting barley (Sato and Takeda 2009) and wheat cultivars (Somers, et al. 2003). These assays have given rise to distinctness, uniformity and stability testing and assessment of plant breeding studies (Chiapparino, et al. 2004).
In addition to human genetic databases, some databases for important plants have been developed like (http://bioinf.scri.sari.ac.uk/barleysnpdb/contact.html) as Barley SNP Database and Wheat SNP Database (http://wheat.pw.usda.gov/ITMI/WheatSNP/), and Potato Database (https://www.gabipd.org/projects/Pomamo/).

1.3.2.3.2. Allelic Discrimination Method for SNPs

There are a great many methods used to detect SNPs including both sequence specific and sequence-nonspecific (by measuring molecular weight based- mass spectrometry) ways. Both methods are utilized for detection of polymorphisms or mutations and all mechanisms are effective but each has its pros and cons. Sequence-nonspecific detection is depend on the capture, cleavage or mobility change during electrophoresis or liquid chromotography of mismatched heteroduplexes formed among allelic DNA molecules which has different size and/or conformation (Kwok and Chen 1998). However, the reliability of sequence-specific detection methods is higher. Mainly four general mechanisms are used for allelic discrimination: allele-specific hybridization, allele-specific nucleotide incorporation, allele-specific oligonucleotide ligation, and allele-specific invasive cleavage (Kwok and Chen 2003).

Melting curve genotyping is useful for single nucleotide polymorphism (SNPs) detection and uses real time PCR. Different thermal characteristics of mutant or wild type PCR amplicons are used by the Light Cycler 480 System (Roche Applied Science). Sequence knowledge is needed to design specific probes to a specific region and detection is implemented based on the theory that labelled probes that bind with different alleles or allele combinations in an SNP-containing region. Region contained mismatches of probe and target and fully matched regions give the different profile on melting curve analysis by SimpleProbes. SimpleProbes are used as the hybridization probes because they hybridize to a target sequence containing the SNP of interest. Hybridized probes emit a fluorescent signal that is used to discriminate fully bound probes (wild-type) or floppily bound (mutant-type). The ability to detecting the fluorescent signal change resulting from mismatches between mutant and probe allow

sensitivity to detect one nucleotide differences. Thus, SimpleProbe probes are an efficient tool for SNP genotyping and useful for identifying heterozygous samples in addition to determining the wild types and mutant types. SimpleProbes have an advantage because of their highly efficient labeling and single step nature.

Heterozygote, homozygote mutant and wild type samples have been separated in melting curve analysis as shown in Figure 1.2. In this figure, the data are based on Melting Curve Analysis (MCA), the gene polymorphism was analyzed with SimpleProbe probes and the three genotypes (homozygous C/C and T/T, heterozygous C/T) are shown. As to the working principle of MCA, the best match between probe and target requires a higher melting temperature because of the difficulty of breaking hydrogen bonds. So, the samples in the figure which are dyed with green, represent the homozygous wild type target and red ones show the mutant samples (the lowest H bonds interaction between sample and probes), blue marked samples have both strength (heterozygous samples).

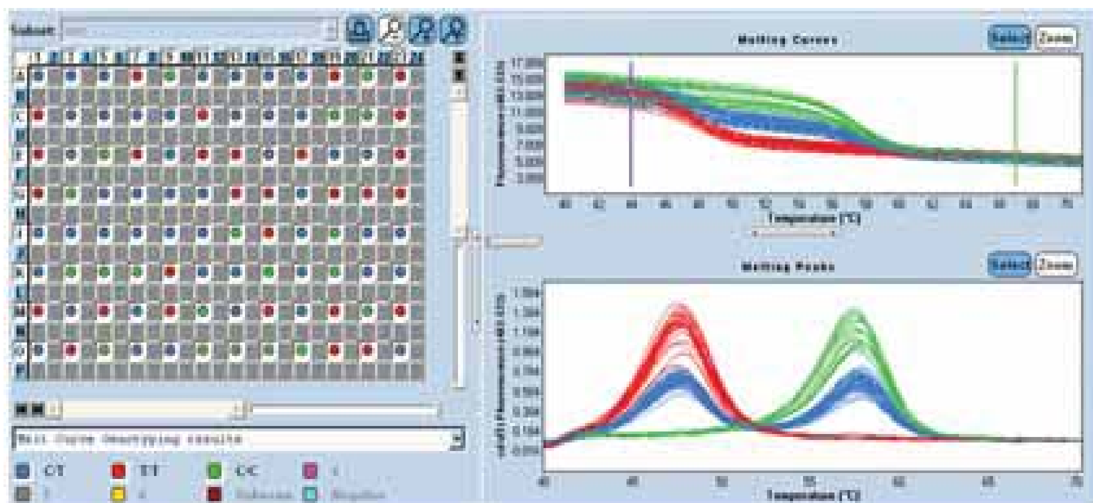

Figure 1.2. SNP analysis methods on the LightCycler®480 System.

1.4. Goals

The one of the aim of us was to make characterization of Turkish Olive cultivars by using SSR and SRAP markers. Morphological characterization and patrimony selection may make Turkish olive accessions genetically closer/identical that can in a

bottleneck effect for breeding efforts. Therefore, before such breeding efforts can be begun, the genetic diversity of Turkish accessions should be determined. The data can then be used in choosing new lines for breeding or for preservation of established varieties.

The other aim was to use SNPs in olive accessions to find unique polymorphism(s) to distinguish the Memecik accession among other Turkish olive accessions. In another word we tried to make traceability test for Memecik olive accessions. We chosed the Memecik accessions because approximately, 45% of total Turkish olive oil production depends on Memecik accessions. Moreover, when this oil is exported to Europe, some registration procedures are necessarily needed like a traceability test to verify purity. Thus, it was hoped that SNP(s) could be used for a traceability test of Memecik olive using its oil.

CHAPTER 2

MATERIALS AND METHODS

2.1. Materials

2.1.1. Olive Materials

The 66 olive accessions' leaf samples and their oil were collected from Alata and Ataturk Central Horticultural Research Institutes, Yusufeli, Milas Agricultural Office, Hatay, Kilis City Office of Agriculture. Six individuals were sampled from Black Sea Region and 9 from Marmara, 10 from Mediterranean, 18 from Southeast Anatolia Region and the last 23 from Aegean Region. Three samples were also chosen as outgroups which are Manzanilla (Spain), Ascolana (Italy) and Lucques (France) to determine their genetic diversity using DNA-based markers. The morphological features of the varieties are described in the introduction and their localization in Turkey is provided in Table 2.1.

Table 2.1. Turkish Olive Cultivars and outgroup genotype use and localization list

Genotype number	Genotype	Use	City, Region	Genotype number	Genotype	Use	City, Region
1	Trabzon yağlamalık	Table&oil	Trabzon, Blacksea	36	İri yuvaklak	Table	Tatayn (Urfa), South Anatolia
2	Samsun yağlık	Oil	Samsun, Blacksea	37	Yağ çelebi	Oil	Tatayn (Urfa), South Anatolia
3	Pastos	Table	Trabzon, Blacksea	38	Zoncuk	Table	Derik (Mardin), South Anatolia
4	Otur	Table&oil	Artvin, Blacksea	39	Halhalı	Table&oil	Derik (Mardin), South Anatolia
5	Satı	Table	Artvin, Blacksea	40	Hursuki	Oil	Derik (Mardin), South Anatolia
6	Siyah salamuralık	Table	Tekirdağ, Marmara	41	Belluti	Table&oil	Derik (Mardin), South Anatolia
7	Samsun tuzlamalık	Table	Samsun, Blacksea	42	Melkabazı	Oil	Derik (Mardin), South Anatolia

(cont. on next page)

Table 2.1. (cont.)

8	Beyaz yağlık	Oil	Tekirdağ, Marmara	43	Mavi	Table	Derik (Mardin), South Anatolia
9	Çizmelik	Table	Tekirdağ, Marmara	44	Ayvalık yağlık	Oil	Ayvalık (Balıkesir), Marmara
10	Erdek yağlık	Oil	Erdek (Balıkesir), Marmara	45	Hurma karaca	Table	İzmir, Aegean
11	Edincik	Table	Edincik (Balıkesir), Marmara	46	Hurma kaba	Table	İzmir, Aegean
12	Eşek zeytini	Table&oil	Ödemiş (İzmir), Aegean	47	Erkence	Oil	İzmir, Aegean
13	Trilye	Table&oil	Bursa, Marmara	48	Çilli	Table	İzmir, Aegean
14	Samanlı	Table	İznik, Marmara	49	İzmir sofralık	Table	İzmir, Aegean
15	Çelebi	Table&oil	İznik, Marmara	50	Çakır	Oil	İzmir, Aegean
16	Büyük topak ulak	Table	Tarsus, Mediterranean	51	Memeli	Table&oil	İzmir, Aegean
17	Sarı ulak	Table	Tarsus, Mediterranean	52	Dilmit	Oil	Bodrum (Muğla), Aegean
18	Çelebi	Table	Silifke(Mersin), Mediterranean	53	Girit zeytini	Table&oil	Bodrum (Muğla), Aegean
19	Halhalı	Table&oil	Hatay, Mediterranean	54	Tavşan yüreği	Table	Milas (Muğla), Aegean
20	Sarı çelebi	Table&oil	Hatay, Mediterranean	55	Ak zeytin	Table&oil	Ödemiş (İzmir), Aegean
21	Saurani	Oil	Hatay, Mediterranean	56	Çekişte	Table &oil	Ödemiş (İzmir), Aegean
22	Sayfi	Oil	Hatay, Mediterranean	57	Kara yaprak	Oil	Kuşadası (Aydın), Aegean
23	Karamani	Oil	Hatay, Mediterranean	58	Yağ zeytini	Oil	Kuşadası (Aydın), Aegean
24	Elmacık	Oil	Hatay, Mediterranean	59	Yerli Yağlık	Oil	Kuşadası (Aydın), Aegean
25	Yağlık sarı zeytin	Oil	K. Maraş, Mediterranean	60	Aşı yeli	Table&oil	Aydın, Aegean
26	Kilis yağlık	Oil	Kilis, Mediterranean	61	Taş arası	Table&oil	Aydın, Aegean
27	Nizip yağlık	Table&oil	Nizip (Antep), South Anatolia	62	Taş arası	Table&oil	Kuşadası (Aydın), Aegean
28	Kan çelebi	Table	Nizip (Antep), South Anatolia	63	Memecik	Table&oil	Milas (Muğla), Aegean
29	Halhalı çelebi	Table&oil	Hatay, Mediterranean	64	Domat	Table	Manisa, Aegean
30	Hamza çelebi	Oil	Nizip (Antep), South Anatolia	65	Kiraz	Table&oil	Manisa, Aegean
31	Yuvaklak halhalı	Oil	Nizip (Antep), South Anatolia	66	Uslu	Table	Manisa, Aegean
32	Yağlık çelebi	Oil	Nizip (Antep), South Anatolia	67	Manzanilla	Table	Spain
33	Yün çelebi	Oil	Nizip (Antep), South Anatolia	68	Ascollana	Oil	Italy
34	Eğriburun	Table	Nizip (Antep), South Anatolia	69	Lucques	Table&oil	France
35	Eğriburun	Table	Tatayn (Urfa), South Anatolia				

2.1.2. SSR Primers

PCR amplification primers and conditions were performed by taking cognizance of previously published articles. In this research, 13 polymorphic SSR primers were used to determine to diversity of Turkish Olive accessions and draw the phylogenetic tree. These 13 primers and reference articles were illustrated in Table 2.2

Table 2.2. SSR primers used to characterize genetic diversity in olive

Locus	Repeat sequence	Annealing Temperature	Ref. primers
Dca3	(GA)19	60 °C	Sefc.et.al. 2000
Dca4	(GA)16	55 °C	Sefc.et.al. 2000
Dca7	(AG)19	60°C	Sefc.et.al. 2000
Dca11	(GA)26(GGGA)4	58°C	Sefc.et.al. 2000
Dca14	(CA)18A6(TAA)7	60 °C	Sefc.et.al. 2000
Dca18	(CA)4CT(CA)3(GA)19	55 °C	Sefc.et.al. 2000
Emo90	(CA)10	60 °C	De la Rosa R. et al.2002
Gapu71b	GA(AG)6(AAG)8	60 °C	F.Carriero et al 2002
Gapu 101	(GA)8(G)3(AG)3	60 °C	F.Carriero et al 2002
Udo9	(AG)16	55 °C	Cipriani G.et al.2002
Udo24	(CA)11(TA)2(CA)4	56 °C	Cipriani G.et al.2002
Udo28	(CA)23(TA)3	67.9 °C	Cipriani G.et al.2002
Udo43	(GT)12	58°C	Cipriani G.et al.2002

2.1.3. SRAP Primers

For SRAP analysis, all combinations of EM1 to EM17 forward primers and ME1 to Me14 reverse primers were surveyed using the PCR conditions: 94 °C for 5 min for denaturation followed by 5 cycles at 94°C for 1 min, 35 °C for 1 min, 72°C for 1 min and followed by 35 cycles at 94°C for 1 min, 55 °C for 1 min, 72°C for 1 min followed by elongation at 72°C for 10 min. Then 12 highly polymorphic SRAP primer combinations: EM1-ME4, EM3-ME13, EM3-ME14, EM6-ME13, EM7-ME1, EM8-ME8, EM8-EM6, EM11-ME2, EM12-ME8, EM12-ME9, EM12-ME13, EM13-ME7 were tested on Turkish olive cultivars. For 20 volume PCR, 100 ng of genomic DNA, 2 pmol of forward and reverse primers, 1x PCR buffer, 3 mM Mg^{++} , 0.7 µM dNTPs, 1 U *Taq* Polymerase were used (Mukhlesur 2007). The amplified DNA fragments were

separated using 4% agarose gel and stained by ethidium bromide (1:20.000 v:v) and electrophoresed at 1 V cm^{-1} for 180 min and photographed under a UV transilluminator (BIO-RAD, The discovery series).

2.1.4. Ant and Cyc Primers for SNP Detection

For SNP discovery in our olive cultivars, the *cycloartenol synthase* and *anthocyanidin synthase* loci were chosen (Reale, et.al.2006). NCBI accession numbers of *cycloartenol synthase* and *anthocyanidin synthase* are AB025344 and AF384050, respectively. For *cycloartenol synthase* gene, the sequences of forward and reverse primers are GCCCATTTCAGATTGCAC and GGGATTCTCAGGTCAGGA, respectively. Forward and reverse primers of *anthocyanidin synthase* gene are GCCCAGCAACAAGTGAGTATGCAAAAC,
AACCCAATTTTTCAACTCATTTTTCTTC ACC, respectively.

2.2. Methods

2.2.1. DNA Extraction

Common methods of DNA isolations which are used successfully in plant research resulted in dirty, yellow and highly viscose DNA for olive leaves and oil. The organic content of olive such as phenolics, polysaccharides and other organic substance has to be separated from the DNA. To successfully isolate DNA some modifications were tried.

2.2.1.1. DNA Isolation from Olive Oil

Olive fruit of the 69 accessions were collected and their oil extracted without hexane treatment by the Ozaltin Oil Company, Aydın. Cold-press extraction was preferred to avoid degradation of the DNA. To isolate DNA, a cetyltrimethyl-ammonium bromide (CTAB)-based method was performed by using 1 ml unfiltered cold-press olive oil. The same method was also repeated with the addition of 1% PVP.

Promega Wizard Kit was tested and Qiagen QIAamp DNA stool Kit was successfully used (Testolin, et al. 2005).

2.2.1.2. DNA Extraction from Leaves

Total DNA extraction from leaf materials was performed with previously published methods (Doyle and Doyle 1990) with some modifications included grinding with liquid nitrogen and using 1.5 g ground tissue instead of 400 mg. The chloroform step of extraction was done twice because of high phenolic content. After DNA extraction, quality and quantity of DNA for each sample were measured by using a Nanodrop spectrophotometer and the quantities obtained are given in Table 2.1. Although the Nanodrop measurements suggested that the DNA was not of excellent quality, no problems were encountered with PCR and sequencing. The 11 samples which were used in sequence analysis were also cleaned up by ethanol precipitation method (as described in the Beckman Coulter QuickStart sequence analysis kit).

2.2.2. Data Analysis

2.2.2.1. PCR Analysis of SSR Primers and Capillary Electrophoresis

PCR reactions were carried out in 25 µl volume containing: 40 ng of DNA, 1 pmol of forward and reverse primers, 1x PCR buffer, 3 mM Mg^{++}, 0.125 mM dNTPs, 1 U *Taq* Polymerase and PCR conditions at 94 ^{0}C for 3 min followed by 36 cycles at 94^{0}C for 30s, 45 s at annealing temperature (Table2.2) and 72 ^{0}C for 45 s, finally 7 min at 72 ^{0}C. PCR reactions were performed in Eppendorf Mastercycler Epgradient S. Before the ultimate fragment analyses PCR products (only 5 µl) were checked on 2% agarose gels containing ethidium bromide (1:20.000 v:v) and electrophoresed at 1 V cm^{-1} for 30 min and photographed under a UV transilluminator (BIO-RAD, the discovery series).
After PCR products were seen on the agarose gel, the fragments were separated using the high resolution kit of a Qiagen Capillary electrophoresis device. OL500 method was chosen according to size and concentration of target PCR products. To determine the size of PCR fragments, QX Alignment Marker 15 bp/3 kb was preferred and QX DNA

Size Marker 100 bp/3kb which correlated with the alignment marker was used for per reaction

2.2.2.2. SRAP Marker Based Analysis

Twelve SRAP primer combinations were used to amplify the 66 Turkish genotypes. The PCRs were performed (Ferriol, et al. 2003) with some modifications. PCR was performed in10 µl reaction volume containing 0.75 ml of 1mM each of forward primer and reverse primer, 1µl of 25mM $MgCl_2$, 1µl of 10X PCR buffer, 1µl of 2.5mM dNTPs (Promega, Madison, WI), 0.2ml of 5U Taq DNA polymerase and 2µl of 100 ng of genomic DNA. The thermal cycler profile for PCR amplifications in the Eppendorf Mastercycler Epgradient S was as follows: denaturation at 94˚C for 4 min, followed by five cycles of denaturing at 94 ˚C for 1 min, annealing temperature at 35˚C for 1 min and elongation at 72˚C for 1 min. In the remaining 30 cycles, the annealing temperature was increased to 50 ˚C for 1 min with a final elongation step at 72˚C for 7 min (Ferriol et al., 2003). The amplified fragments were separated on 4 % agarose gel with ethidium bromide (1:20.000 v:v). Digital images of the gel were saved onto a computer and scored manually.

2.2.2.3. Sequence Analysis

To determine the SNPs to use in traceability test of Memecik accessions, 11 accessions were chosen among 66 Turkish olive accessions for sequence analysis. These 11 accessions were chosen based on region. Sequence analysis was performed three times for each DNA sample. Chosen accessions are given in the Table 2.3.

Table 2.3. Sample names and their regions, used for sequence analysis

Genotype number	Genotype name	City, Region
1	Trabzon Yağlık	Trabzon, Blacksea
19	Samsun Tuzlamalık	Samsun, Blacksea
23	Erdek Yağlık	Balıkesir, Marmara
24	Edincik	Balıkesir, Marmara
58	Eğriburun	Antep,Southeast Anatolia.
75	Ayvalık Yağlık	Balıkesir, Marmara
78	Erkence	İzmir, Aegean
94	Memecik	Aydın, Aegean
95	Domat	Manisa, Aegean
96	Kiraz	Manisa, Aegean
97	Uslu	Manisa, Aegean

Each reaction was repeated to prevent sequence errors from Taq polymerase-based mistakes and reading errors of the sequencer detector. All reactions were performed using both forward and reverse primers to avoid the misnomer of nucleotides by detector reading false.

Anthocyanidin and *cycloartenol synthase* genes were amplified by PCR reactions with 100 µl volume which included 100 ng of DNA, 0.3 pmol of forward and reverse primers, 1x PCR buffer, 1.5 mM Mg^{++}, 0.125 mM dNTPs, 1 U *Taq* Polymerase and PCR conditions at 94 ^{0}C for 3 min followed by 36 cycles at 94 ^{0}C for 30s, 45 s at 58 ^{0}C annealing temperature and 72 ^{0}C for 45 s, 7 min at 72 ^{0}C. PCR reactions were completed in Eppendorf Mastercycler Epgradient S. PCR products were purified by High Pure PCR Product Purification Kit of Nucleospin Extract II Kit. Cycle-sequence PCR included, 8 µl Master mix, 1 µl primer, 0.6- 6 µl template (for 10 ng/ µl) and 5-10 µl dH_2O per reactions, reagents were supplied by Beckman Coulter Quickstart Sequence Kit. After the reactions were completed, samples were loaded with 40 µl sample loading solution to CEQ^{TM} 8800 Genetic Analysis System and LFR-1 method was chosen based on PCR product size. Also the same reactions were performed at the Biotechnology Center of İzmir Institute of Techonology, İzmir. Sequences were aligned and compared using the BioLign Programme. Probes were designed for allelic discrimination using LCPDS2 programme support from Roche Company.

2.2.2.4. Real Time-PCR Optimization

To find the working concentration of $MgCl^{++}$, two concentration: 1.5 mM and 3 mM were used. Each 20 µl reaction volume contained 1X master mix (supported by Light Cycler 480 probe master kit), 0.2 mM primer forward and reverse, 1.5 / 3 mM $MgCl^{++}$, 75 ng DNA. SimpleProbe option was chosen as the software detection format. The real-time PCR profile for PCR amplifications were as follows: denaturation for 1 cycle with no analysis, for 45 cycles with quantification, melting for 1 cycles with melting curve analysis and cooling for machinery utility. Condition details of the melting curve analysis are given in Table 2.4. Acquisition mode means the frequency with which fluorescence data was acquired. Single option means it acquired fluorescence data once at the end of this segment in each cycle (recommended for quantification tests). Hold options depicts the length of time to hold the target temperature in hours: minutes: seconds format. Ramp Rate (˚C/s) describes as the rate at which the instrument heats up or cools down to target temperature. And the last option was acquisitions per ˚C , so it was vital to arrange the number of data measurements taken per temperature point.

Table 2.4. conditions of Melting Curve Anaysis

	Target (˚C)	Acquisition mode	Hold (hh:mm:ss)	Ramp rate (˚C/s)	Acquisitions (per ˚C)
Denaturation	95	None	00:10:00	4.4	-
Cycling	95	None	00:00:10	4.4.	-
	57	Single	00:00:15	2.2	-
	72	None	00:00:10	4.4	-
Melting Curve	95	None	00:00:01	4.4	-
	40	None	00:00:30	2.2	-
	85	Continuous	-	0.10	6
Cooling	40	None	00:00:30	2.2	-

CHAPTER 3

RESULTS AND DISCUSSION

3.1. DNA Extraction

DNA extraction of 66 Turkish accessions and 3 outgoups were performed and their DNA concentrations are listed in Table 3.1. Also, their 230/260 and 260/280 ratios were checked to see the protein or phenol/ethanol contamination, ratios were in the proper interval (data not shown). However, 11 samples which were used in sequence analysis were cleaned up by ethanol precipitation to get an absorbance value of 2.00 for measurements at 230/260 and 260/280.

Table 3.1. DNA concentration of samples

Genotype	Location	DNA amount ng/µl	Genoype	Location	DNA amount ng/µl	Genotype	Location	DNA amount2
Trabzon yağlamalık	Trabzon	267.1	Çelebi	İznik	1156	İri yuvaklak	Tatayn	500.4
Samsun yağlık	Samsun	420.8	Büyük topak ulak	Tarsus	228.5	Yağ çelebi	Tatayn	730.4
Pastos	Trabzon	1032	Sarı ulak	Tarsus	491.3	Zoncuk	Derik	408.2
Tuzlamalık	Samsun	713.7	Çelebi	Silifke	765.2	Halhalı	Derik	414.6
Otur	Artvin	618.4	Halhalı	Hatay	476.8	Hursuki	Derik	327.8
Satı	Artvin	1215	Sarı çelebi	Hatay	1162.8	Belluti	Derik	590.6
Tuzlamalık	Samsun	229.7	Saurani	Hatay	820.5	Melkabazı	Derik	797.5
Siyah salamuralık	Tekirdağ	724.8	Sayfi	Hatay	443.3	Mavi	Derik	847.8
Samsun tuzlamalık	Samsun	851.1	Karamani	Hatay	1031.7	Samsun tuzlamalık	Samsun	713.7
Beyaz yağlık	Tekirdağ	487.2	Hurma kara	İzmir	470.9	Ayvalık yağlık	Ayvalık	911.3
Çizmelik	Tekirdağ	664.7	Yağlık sarı zeytin	K.Maraş	405.4	Hurma karaca	İzmir	846.4
Erdek yağlık	Erdek	653.9	Kilis yağlık	Kilis	574.3	Elmacık	Hatay	582.4
Edincik	Edincik	548.6	Nizip yağlık	Nizip	645.6	Erkence	İzmir	527.1

(cont. on next page)

Table 3.1. (cont.)

Eşek zeytini	Ödemiş	474.5	Kan çelebi	Nizip	1032.3	Çilli	İzmir	706.3
Trilya	İznik	412.2	Halhalı çelebi	Hatay	945.2	İzmir sofralık	İzmir	655.7
Samanlı	İznik	723.8	Hamza çelebi	Nizip	343.5	Çakır	İzmir	520.5
Taş arası	Aydın	464.4	Yuvaklak halhalı	Nizip	466.3	Memeli	İzmir	984.4
Taş arası	Kuşadası	924.0	Yağlık çelebi	Nizip	720.3	Dilmit	Bodrum	674.4
Memecik	Milas	767.2	Yün çelebi	Nizip	720.6	Girit zeytini	Bodrum	698.9
Domat	Akhisar	622.7	Eğriburun	Nizip	530.9	Tavşan yüreği	Milas	587.7
Kiraz	Akhisar	816.5	Eğriburun	Tatayn	487.5	Ak zeytin	Ödemiş	466.2
Uslu	Akhisar	848.1	Yerli yağlık	Kuşadası	1060.4	Çekişte	Ödemiş	700
Yağ zeytini	Kuşadası	663.0	Aşı yeli	Aydın	728.6	Kara yaprak	Kuşadası	699.5
Manzanilla	Spain	520	Lucques	France	900	Ascollana	Italy	1010

After DNAs of the Turkish olive cultivars were obtained, the SSR and SRAP marker systems were used to determine the diversity. Qiaxcel Capillary Electrophoresis system was used to genotype the PCR products which were amplified by SSR primers. Also SRAP primers were used to amplify ORFs (open reading frames) of our lines and the amplified fragments were analysed on agarose gel system.

3.2. Genotyping Using Molecular Markers

3.2.1. SSR Primers Based Characterization on Capillary Electrophoresis

A total of 13 polymorphic SSR markers were used to determine genetic diversity of the 69 olive accessions. After each locus was amplified by PCR, capillary electrophoresis was used to assign the sizes of the fragments. A total of 119 polymorphic regions were detected by Qiaxcel software. The polymorphic regions were analysed on the Qiaxcel software, peaks for each fragments were converted the binomial data first by software and then given peaks and gel images were checked visually. An example of these peaks is illustrated for one SSR locus of a individual in figure 3.1. For this locus and individual, there are two fragments which were sized as 164 bp and 182 bp, respectively as shown in figure 3.1. Binomial data were not only

obtained by program using peaks but also the data were analysed by examination of gel images. One of our gel images can be seen in figure 3.2. Thanks to this system our olive population was, characterized for marker polymorphism as shown in figure 3.2.

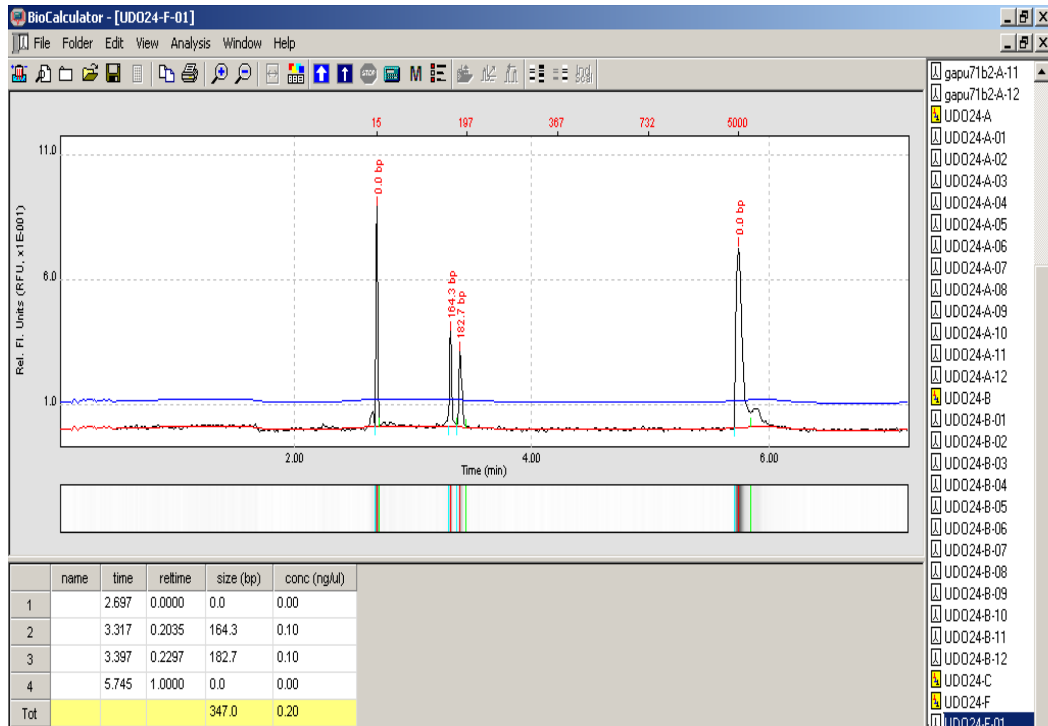

Figure3.1. Capillary electrophoresis result for udo24 marker of one individual. Separation range was defined by using first and last peaks which are called alignment marker and occur at 15 bp and 5000 bp. respectively.

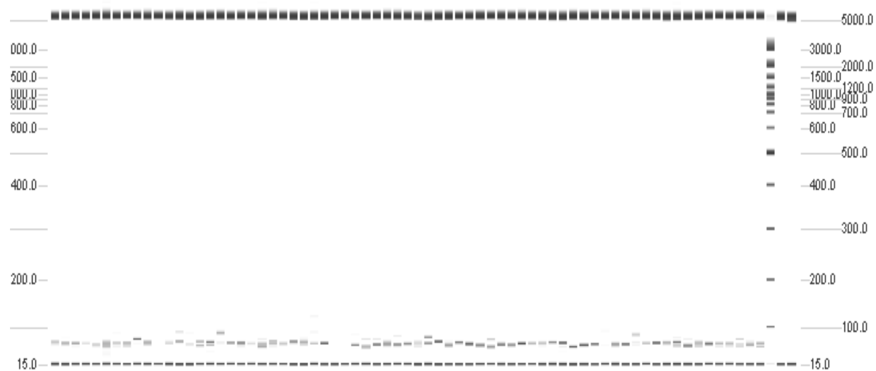

Figure 3.2. Gel image of polymorphic SSR marker on capillary electrophoresis The last cloumn was loaded with size marker to measure the size of fragments.

3.2.1.1. Genetic Similarity Test Using SSR Primers

The 13 SSR markers and their alleles are listed in Table3.1. All 13 SSR markers were polymorphic and produced 119 alleles with an average number of 9.15 alleles. The average PIC value was 0.235. For individual alleles, PIC values ranges from 0.074 to 0.391. DCA7, DCA14, EMO90 and GAPU71B had the highest average PIC values ranging from 0.288 for DCA14 to 0.391 for EMO90. UDO9 and UDO24 had the lowest PIC values 0.063 and 0.150, respectively.

Table 3.2. PIC values for each SSR primer

Locus	No.alleles	PIC±standard error	Range	Ref. primers
DCA3	9	0.318±0.054	0.083-0.493	1
DCA4	13	0.171±0.045	0.029-0.493	1
DCA7	8	0.289±0.061	0.227-0.497	1
DCA11	8	0.187±0.047	0.059-0.454	1
DCA14	10	0.288±0.053	0.029-0.499	1
DCA18	6	0.248±0.076	0.056-0.498	1
EMO90	8	0.391±0.027	0.277-0.498	2
GAPU71B	9	0.294±0.056	0.056-0.499	3
GAPU101	20	0.233±0.033	0.028-0.464	3
UDO9	8	0.074±0.020	0.028-0.161	4
UDO24	6	0.150±0.071	0.029-0.499	4
UDO28	6	0.276±0.062	0.110-0.487	4
UDO43	8	0.279±0.069	0.029-0.487	4

The SSR data for the 66 Turkish olive accessions and three outgroups was used to draw a dendrogram based on DICE matrix and UPGMA arithmetical averages in SHAN module using NTSYS-pc version 2.2. Programme. The dendrogram is shown in Figure.3.2.

Based on a Mantel test, the correlation between the dendrogram and distance matrix was r = 0.85 (P = 1.0) indicating a good fit between the tree and distance data (Rohlf 1994). This dendrogram revealed a minimum genetic similarity coefficient of 0.45 for the Turkish olive accessions. Turkish accessions were mainly grouped into

eight clusters with ~56% similarity and the outgroups from Italy, Spain and France were separated in a different cluster.

The European accessions clustered outside of the Turkish ones with 0.27 similarity among outgroups and the Turkish olives. The largest cluster of Turkish accessions, B, contained 43 accessions which subdivided into two main and three smaller groups at 0.59 similarity. Cluster B1 contained 13 accessions with 'Halhali Celebi and 'Yuvarlak Halhali' the most similar (0.90). Cluster B2 contained 23 accessions with three pairs of accessions having genetic similarity of 0.88 to 0.89 ('Celebi Silifke' and 'Halhali Hatay', 'Hurma Kara' and 'Yerli Yaglik', 'Asi Yeli' and 'Memecik'). The second largest cluster was cluster A with 11 accessions. These accessions were more genetically distinct than those in clusters B1 and B2. The remaining five clusters of accessions (from cluster C to H) had one to three accessions each with a maximum similarity of 0.71 in these clusters. Interestingly cluster F consisted of accessions from the Aegean region and these three accessions ('Memeli', 'Domat' and 'Kiraz') had nearly as much genetic diversity (0.57) as the three European accessions which originated from widely dispersed locations (Italy, Spain and France).
As with other crop plants, local olive varieties are often named based on their morphological appearance. Thus, 'Kiraz' which means cherry in Turkish was probably named because its fruit is round and chery-red when ripe. Others are named based on their use and/or location such as 'Trabzon Yaglik' and 'Samsun Yaglik,' both of which are used for their oil and are grown in two different Black Sea provinces. Although these names may reflect morphological or other similarities, our results show that such similarities do not correspond to molecular genetic similarity. For example, two 'Egriburun' accessions were analyzed, both of which are from the South Anatolia region. Despite their identical names and similar geographical origins, these accessions were only 45% genetically similar. Similarly, the two 'Celebi' and two 'Tasarasi' accessions were only 54 and 52% similar, respectively.

We saw localization-specific separation in our SSR based phylogenetic tree. In group A most of the accessions were from the Black Sea and Marmara Regions. Both of these regions are in the north of the Turkey and their climate and agricultural geology is not same but closer to each other than to the other regions. Sub-clusters of B1 were divided into 3 groups as the South Anatolia and Mediterranean-based accessions, Aegean-based and South Anatolian-based accessions. Similar sub-grouping was seen in group F which includes 3 İzmir accessions..

Relationships among the 69 olive accessions was also examined by principal component analysis (PCA). PCA indicated that the first Eigen vector explained 49.5% of the genotypic variance while the second and third vectors each accounted for only ~3% of the variance. The PCA plot showed that the European outgroups clustered away from the Turkish accessions. The Turkish varieties formed a fairly uniform cluster with a few accessions plotted away from the main cluster. These accessions included Erdek Yağlık (10), Otur (4), Eğriburun (35), Domat (64).and Kiraz (65). These accessions also clustered away from the main groups in the dendrogram. The ordination test supported the UPGMA-based tree, accessions from Black Sea Region and Marmara Region accounted for several accessions plotted above the main cluster and the other accessions that plotted away from the main group were Aegean and South Anatolian accessions.

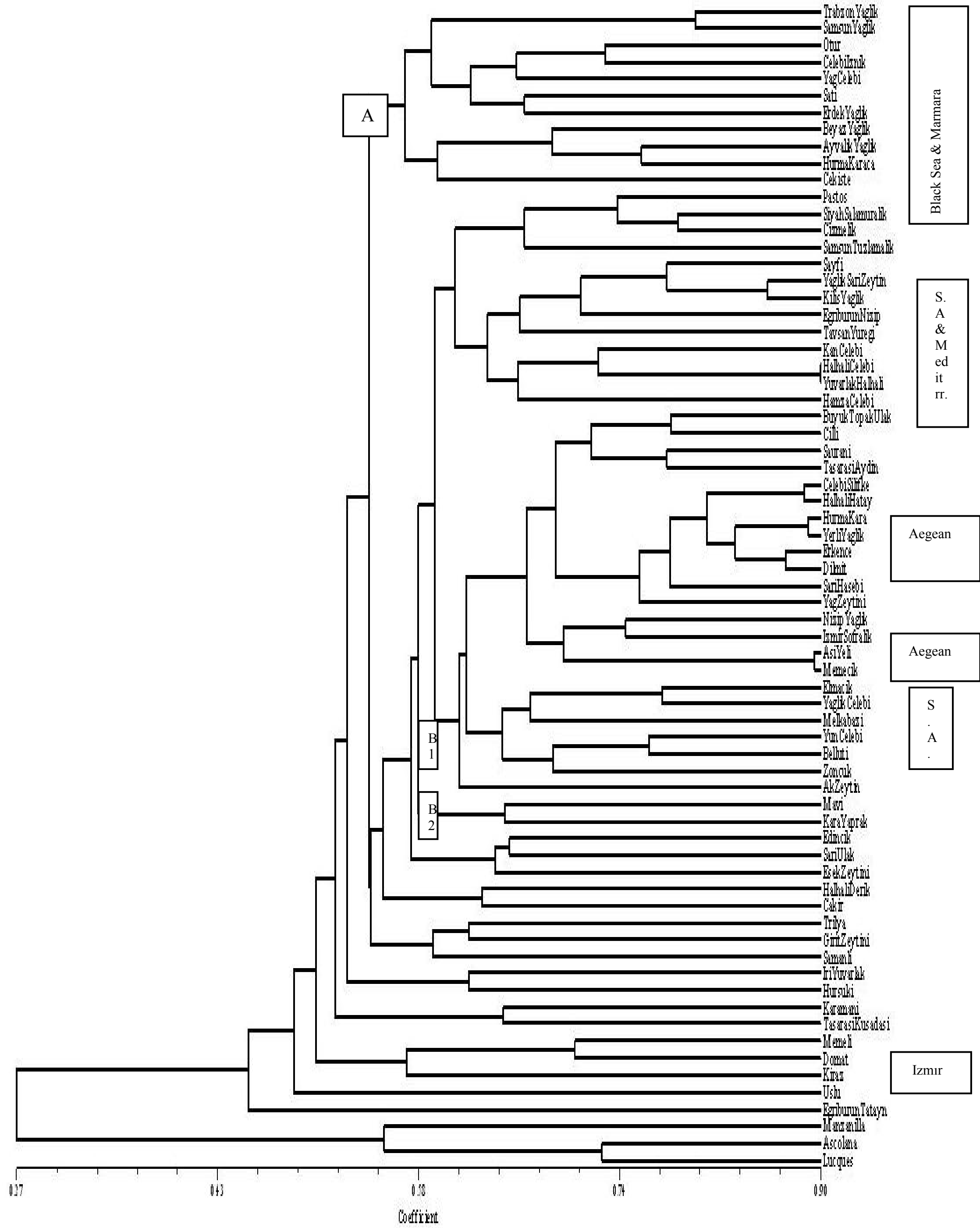

Figure 3.3. Dendrogram was drawn based on SSR markers of Turkish olive cultivars. Outgroups and Turkish cultivars are clearly separated.

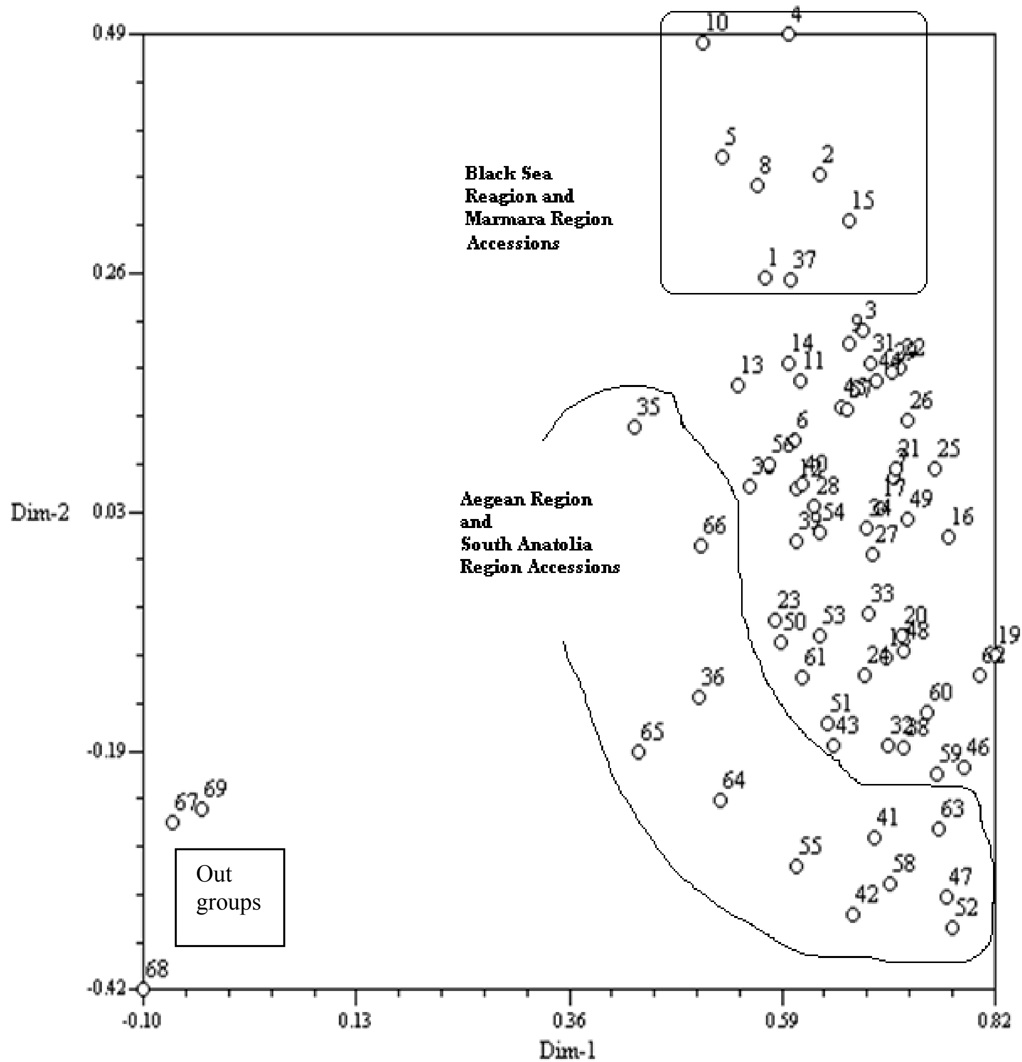

Figure 3.4. PCA analysis of SSR based data

3.2.2. Genotyping Using ORF-Based Markers, SRAPs

To determine the diversity of Turkish olive population we combined 17 EM forward primers and 14 ME reverse primers and then the most efficient 12 primers were chosen to apply to the population.

3.2.2.1. Genetic Similarity Test Using SRAP Markers

Totally 185 alleles were detected and 103 of them were polymorphic. Polymorphic bands were converted to1 - 0 binomial data, by visual inspection. Binomial data were used to draw a dendrogram of our olive cultivars. For each locus PIC values were calculated and PIC ranges are shown in Table 3.2. EM12-ME8, EM8-ME8, EM3-ME13 primer combinations gave the maximum PIC values 0.401, 0.332, 0.328, respectively, while the EM11-ME2 primer combination gave the lowest PIC value, 0.063.

Table 3.3. SRAP data PIC value

Locus	No. Alleles	PIC	Range
EM1-ME4	14	0,228 ± 0.04	0.03-0.43
EM3-ME13	9	0,328 ± 0.05	0.03-0.49
EM3-ME14	5	0,235 ± 0.08	0.08-0.49
EM6-ME13	21	0,294 ± 0.03	0.02-0.49
EM7-ME1	10	0,215 ± 0.04	0.05-0.43
EM8-ME8	7	0,332 ± 0.05	0.02-0.48
EM9-ME16	7	0,217 ± 0.07	0.02-0.47
EM11-ME2	4	0,063 ± 0.06	0.03-0.16
EM11-ME11	1	0,204 ± 0	0.20
EM12-ME8	6	0,401 ± 0.21	0.28-0.48
EM12-ME9	9	0,156 ± 0.06	0.26-0.49
EM12-ME13	8	0,27 ± 0.03	0.17-0.45
EM13-ME7	2	0,245 ± 0.13	0.49-0.56

The tree was drawn on the basis of DICE matrix and then UPGMA (Unweighted Pair Group Method) arithmetical averages in SHAN module using NTSYS-pc version 2.2. Correlation among the binomial data produced with the SRAP markers' alleles was carried out using the Mantel test. Correlation between the data matrix and the tree was found to be 0.83, a good fit. The dendrogram was scaled from 0.66 to 0.96. According to the dendrogram Turkish olive accessions divided into 12 groups with 0.82 similarity (shown in Figure 3.6). The largest group, group A, was composed of 3 subgroups which were called A1,A2 and A3. A1 included accessions from Black Sea Region and Marmara Region in one group, Mediterranean accessions were found in a different

cluster and one cluster carried the Aegean accessions. A2 included Aegean accessions in general. Geologically closer types also occurred in other small clusters. Supported by the SSR- based tree, in SRAP-based tree, the same accessions were found as the most diverse group. We also performed the ordination test with the SRAP data and the most diverged accessions (gathered from South Anatolia and Aegean accessions) were seen to plot away from the others.

Figure 3.5. Dendrogram of Turkish olive cultivars, based on SRAP primers

The first three PCA axes was extracted and the first eigen value was 56.49% and 4.03% and 3.65% were seen as the second and third eigen values. Cumulative eigen value of the first three axes was calculated as 64.17% A 2-dimentional plot of the first two dimensions was compared to separation of the UPGMA tree (figure 3.7.). According to both SSR and SRAP marker systems, Black Sea and Marmara regions' accessions were gathered into the same branches of clusters or in unique clusters while Mediterranean Region-specific accessions were found closer to South Anatolian accessions. Generally, specific accessions for South Anatolia and Aegean were found in separate clusters and distal for both marker system results.

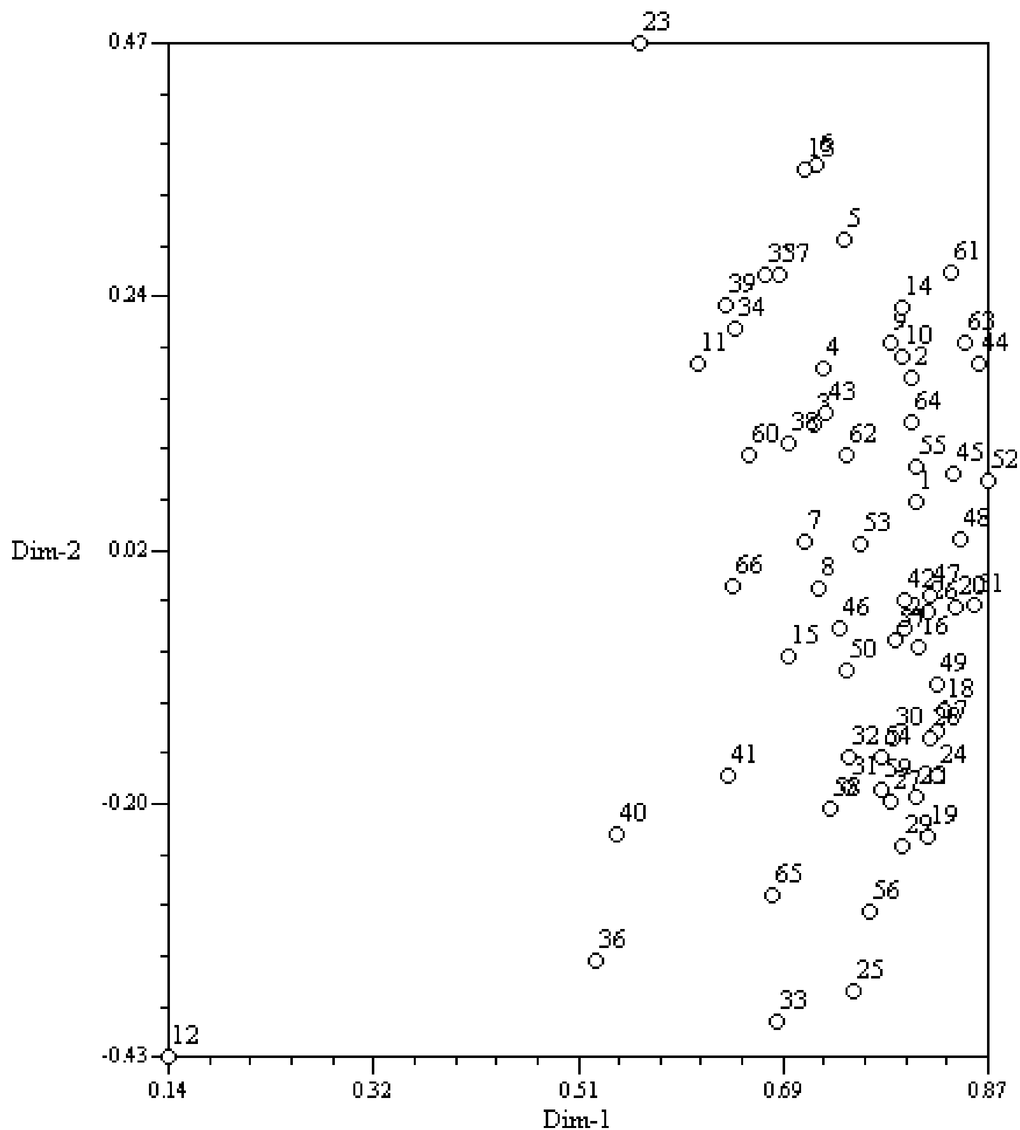

Figure 3.6. PCA analysis, 2-D plot of SRAP primer based data

3.2.3. Genotyping Using Both SRAP and SSR Markers

The13 SSR primer and 12 SRAP primer data were combined and used to consider the entire genome from intronic and non-intronic but conservative regions to draw the tree of Turkish olive cultivars.

Using SM matrix and UPGMA (Unweighted Pair Group Method) arithmetical averages in SHAN module the tree was drawn with NTSYS-pc version 2.2. Tree is shown in Figure 3.9. The tree clustered into 12 arms with 0.82 similarity. Each cluster included nearly the same accessions with the most diverged ones seen to be the same as in the separate SSR and SRAP analyses. Mantel test was done to illustrate the correlation among the binomial data produced by both SSR and SRAP markers. A good correlation was found ($r = 0.80121$), for the combined SSR-SRAP based data. Therefore these two types of marker work efficiently together and comprehend more of the genome in terms of including both ORFs and microsatellite regions. Also to compare both marker systems, Mantel test was performed for the two different distance matrices. The R value was calculated as 0.227 ($p=0.999$). This r value was found as what we expected because of working principle of these two markers are very different. SSR markers amplify the microsatellite regions of genome so it is possible to see both insertion-deletion and other mutations in noncoding of the genome. However, SRAP markers were amplified from the open reading frames so some of the mutations (or insertion/deletion) may have more vital consequences and passed to the next generation. To sum up, less diversity was seen as expected in more conserved SRAP marker. Our results were confirmed with this common knowledge, while the SSR-based dendrogram scaled from 0.45 to 0.90 similarity, the SRAP-based dendrogram had 0.66 minimum similary and 0.96 maximum similarity. Indicating that the SSR (microsatellite) based marker had better separation power because of polymorphic nature. Nevertheless, these two types of markers separated the cultivars into nearly the same clusters with region specific accessions.

Principle component analysis was performed for assembled data (SSR and SRAP marker based) as the ordination test. The first Eigen vector explained 78.1% of the genotypic variances while the second and third vectors accounted for 1.5% and 1.3% respectively. Cumulatively, the first 3 vectors accounted for 80.9% of variance. 2-

Dimensional and 3-Dimensional matrix plots were constructed as seen in Figure 3.8 for 2-D one. The most diverged accessions were marked on the 2-D illustration of ordination test and geographical region of sample were compared to see that Aegean and South Anatolian accessions had special diversity with supported the first two solitary trees of SSR's and SRAP's. Therefore the ordination test supported the identity of the most divergent accessions as determined by the SSR and SRAP marker results.

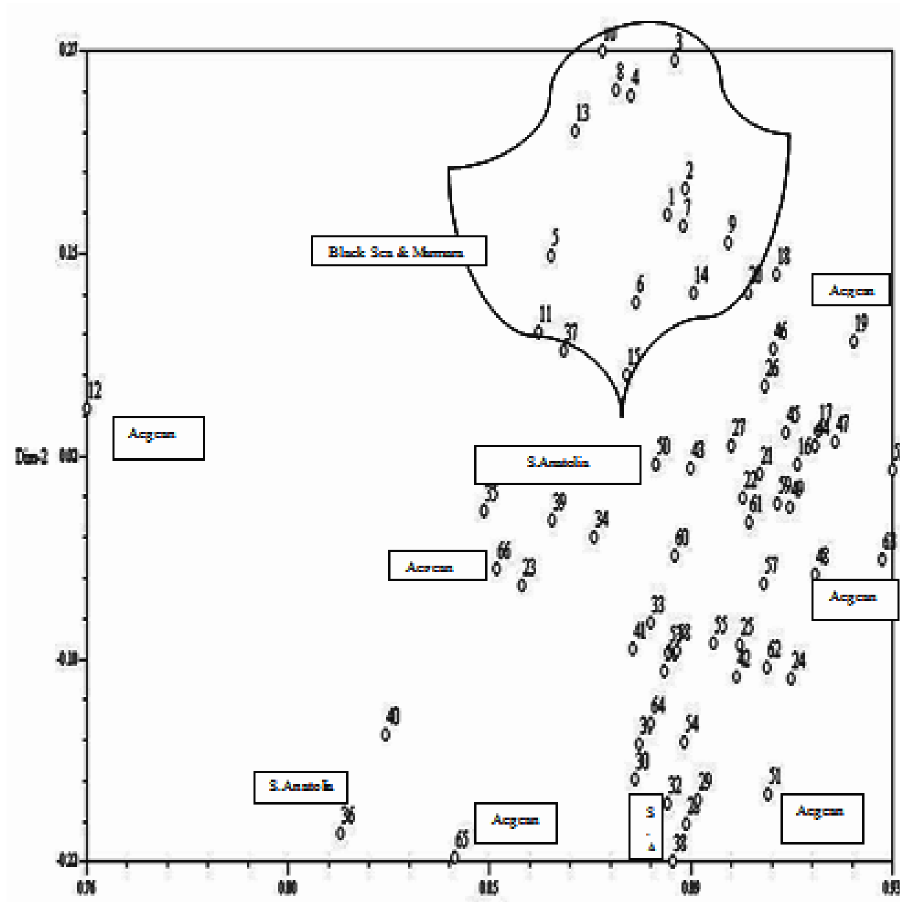

Figure 3.7. 2-D of SSR and SRAP based data

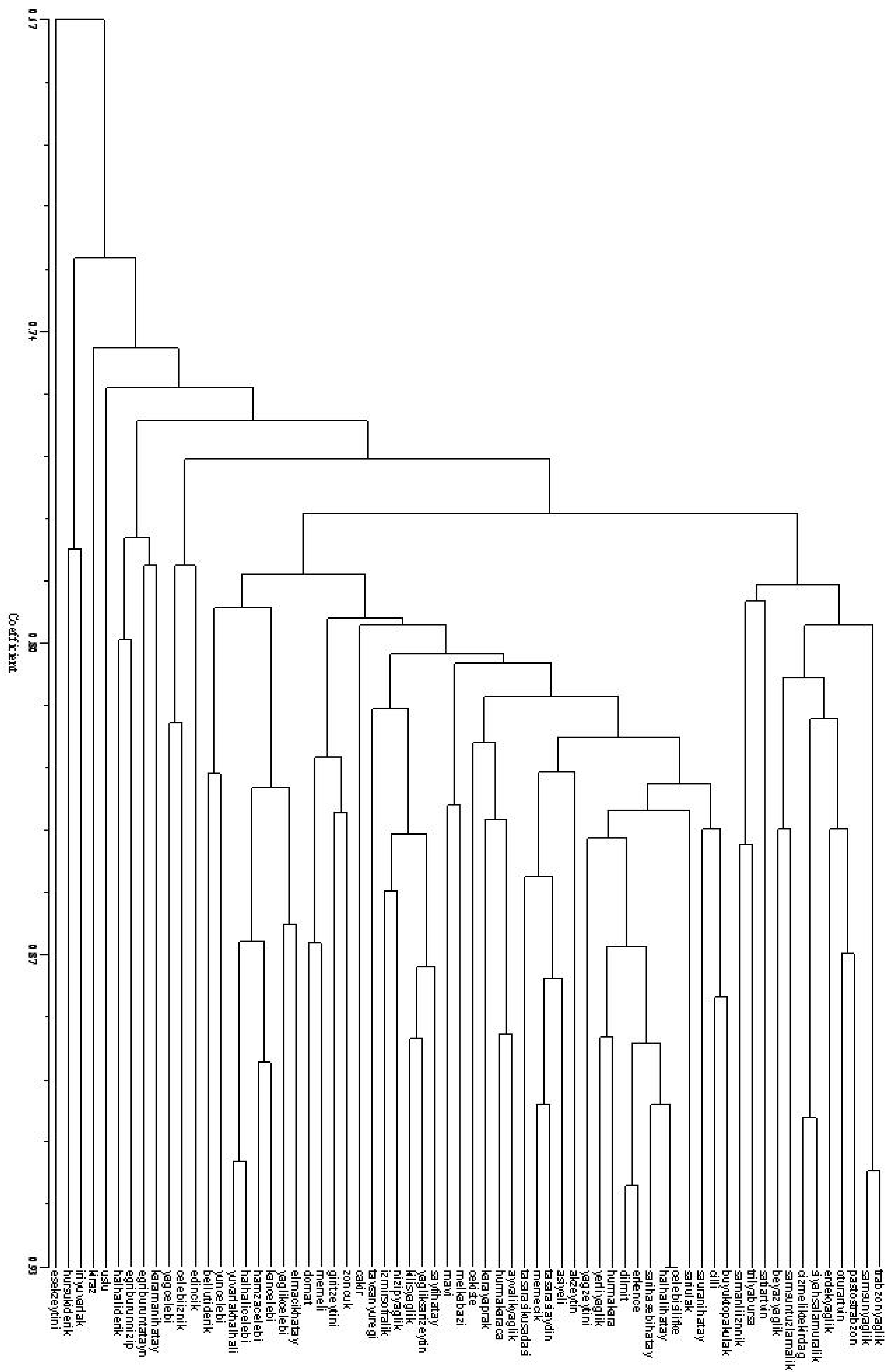

Figure 3.8. Dendrogram of Turkish olive cultivar using SSR and SRAP markers

3.2.3.1. Genome Searching for Memecik-Specific SNPs

Sequence analysis was performed successfully for each concentration of the 11 chosen samples using both *cycloartenol synthase* and *anthocyanidin synthase* genes. Electropherograms for each sample were checked by visually by Beckman Coulter software and also Finch Tv programme software. Obtained sequences were compared with each other and potential SNPs were determined using sequence result of 11 samples on alignment programme, BioLign. Using National Center for Biotechnology Information database, gathered sequences were blasted with previous sequences and each other. Sequences belong to Memecik accessions sequenced with forward and reverse primers were compared to the published *anthocyanidin synthase* gene. In figure 3.10, sequence of Memecik olive accessions by forward primer (8F), (gi|14550121|gb|AF384050.1|) and reverse primer(8R) are illustrated.

```
8F                         --------------------------------------------------
gi|14550121|gb|AF384050.1|
AGTGGTCAACTTGAGTGGCAGGACTATTTCTTCCACTGTATTTACCCAGA       50
8R                         --------------------------------------------------


8F                         --------------------------------------------------
gi|14550121|gb|AF384050.1|
GGAGAAGAGGGACATTTGGCCCAAGACTCCAACTGATTACAGCCCAGCAA    100
8R                         -----------------------------------------TCCCAGCAA 9

                                                      11T/C
8F                         -----------------TTACATATGTGCAGTGAGTTGTATTCTTTCG  31
gi|14550121|gb|AF384050.1|
CAAGTGAGTATGCAAAACAACTGAGAGGTCTAACAAGCAAAATACTATCG 1  50
8R
CAAGTGAGTATGCAAAACAACTGAGAGGTCTAACAAGCAAAATACTATCG     59
                                  * * *T/T *   **    ** ** ***

        42A/C
8F                         ---
CACTCTCTTGGACTAGGGTTAGAACAAGAAAGACTGGAAAAAGAAGT 78
gi|14550121|gb|AF384050.1|
GCACTCTCTCTTGGACTAGGGTTAGAACAAGAAAGACTGGAAAAAGAAGT   200
8R
GCACTCTCTCTTGGACTAGGGTTAGAACAAGAAAGACTGGAAAAAGAAGT  109
                           * ****C/C******************************************

                               96A/C     105T/C
8F
TGGTGGCATGGAAGAACTCCTCCTTCAATTCAAGATAAACTATTACCCAA    128
gi|14550121|gb|AF384050.1|
TGGTGGCATGGAAGAACTCCTCCTTCAATTCAAGATAAACTATTACCCAA    250
8R
TGGTGGCATGGAAGAACTCCTCCTTCAATTCAAGATAAACTATTACCCAA    159
                        ***************A/A******T/T*************************

                          144G/A    151A/G      159G/T
8F
AATGCCCTCAGCCGGAACTCGCCCTCGGGGTCGAAGCCCACACCGACGTC    178
gi|14550121|gb|AF384050.1|
AATGCCCTCAGCCGGAACTCGCCCTCGGGGTCGAAGCCCACACCGACGTC    300
8R
AATGCCCTCAGCCGGAACTCGCCCTCGGGGTCGAAGCCCACACCGACGTC    209
                        *************G/G****G/G*****G/G********************
```

(cont. on next page)

Figure 3.9. Complete illusturation of SNPs on ant. gene. Memecik accession's forward sequence 8R_sample Memecik accession's reverse sequence (in reverse complement format), and the gi|14550121|gb|AF384050.1 | reference sequence obtained from blast,NCBI (Olea europaea anthocyanidin synthase gene). Primers are marked by purple color, green and yellow markes show the homozygous alleles and sequence borders are shown by grey line.

8F
AGTGCACTCACTTTCATCCTCCACAACATGGTGCCTGGCCTGCAACTCTT 228
gi|14550121|gb|AF384050.1|
AGTGCACTCACTTTCATCCTCCACAACATGGTGCCTGGCCTGCAACTCTT 350
8R
AGTGCACTCACTTTCATCCTCCACAACATGGTGCCTGGCCTGCAACTCTT 259
****

233G/C 280A/C
8F
CTGTGAGGGAAAATGGGTCACAGCAAAATGTGTTCCCAACTCCATTATCA 278
gi|14550121|gb|AF384050.1|
CTGTGAGGGAAAATGGGTCACAGCAAAATGTGTTCCCAACTCCATTATCA 400
8R
CTGTGAGGGAAAATGGGTCACAGCAAAATGTGTTCCCAACTCCATTATCA 309
G/GA/A

306T/A 312T/C 321G/C
8F
TGCACATTGGAGATACTATTGAGATATTGTCTAATGGGAAGTATAAAAGT 328
gi|14550121|gb|AF384050.1|
TGCACATTGGAGATACTATTGAGATATTGTCTAATGGGAAGTATAAAAGT 450
8R
TGCACATTGGAGATACTATTGAGATATTGTCTAATGGGAAGTATAAAAGT 359
************************T/A**** T/C*****G/G or C****

377G/T
8F
GCTCTGCACAGAGGGCTTGTGAATAAAGAGAAGGTGAGGATTTCTTGGGC 378
gi|14550121|gb|AF384050.1|
GCTCTGCACAGAGGGCTTGTGAATAAAGAGAAGGTGAGGATTTCTTGGGC 500
8R
GCTCTGCACAGAGGGCTTGTGAATAAAGAGAAGGTGAGGATTTCTTGGGC 409
G/G**

415C/T
8F
AGTTTTCTGTGAGCCACCCAAGGAGAAGATTGTGCTGAAGCCGCTGCCGG 428
gi|14550121|gb|AF384050.1|
AGTTTTCTGTGAGCCACCCAAGGAGAAGATTGTGCTGAAGCCGCTGCCGG 550
8R
AGTTTTCTGTGAGCCACCCAAGGAGAAGATTGTGCTGAAGCCGCTGCCGG 459
**********************************C/T*************

Figure 3.9. (cont.) **(cont. on next page)**

```
                              450G/A   458G/A
8F
AGACAGTTTCTGAGGCTGAGCCACCACGCTTCCCGCCCCGTACCTTTTCT     478
gi|14550121|gb|AF384050.1|
AGACAGTTTCTGAGGCTGAGCCACCACGCTTCCCGCCCCGTACCTTTTCT     600
8R
AGACAGTTTCTGAGGCTGAGCCACCACGCTTCCCGCCCCGTACCTTTTCT     509
                   *******************G/A*****G/G*******************

                     495C/T
8F
CAGCATATCATGCACAAGCTGTTCAGGAAGAGTGAGGATTCCGGTGAAGA 528
gi|14550121|gb|AF384050.1|
CAGCATATCATGCACAAGCTGTTCAGGAAGAGTGAGGATTCCGGTGAAGA 650
8R                    CAGCATAT-
GCATCTCAGCAACAATGCGAGAAGCAAAAGCTCCGTG----              554
        ********    C/T ***    *  ***   *  *   * ***

8F                    AAAATGAGTTGAAAAATTGGGTA-------------------------     551
gi|14550121|gb|AF384050.1|
AAAATGAGTTGAAAAATTGGGTTATTGGTGGGGTATTTTCTTAGCATGTT     700
8R                    -----------------------------------------------


8F                    -----------------------------------------------
gi|14550121|gb|AF384050.1|
CTTCTGGATTTTTGAGCCTAAATGCATGGTGAATGGGATTGTGTATTCTG     750
8R                    -----------------------------------------------


8F                    -------------------------------------
gi|14550121|gb|AF384050.1|
GATATTAATGGTGATTGAATAATATCAGTATCATTTATT                789
8R                    -------------------------------------
```

Figure 3.9. cont

Forward 5'→3'and reverse 3'→5'primer of ant. gene are marked as purple line. SNPs are marked either green or yellow, green ones are genotypically homozygous with wild type (or ancestor allele), yellow ones are heterozygous genotype with altering allele. Below the alignment the possible alleles of the SNPs are written and the localization of SNP on gene is written above the alignment. Grey marked bases show the borders of the sequences, beyond these points sequences were not well determined from the chromatogram files. By looking the ' * ' signs of alignment, sequence of 2 were matched perfectly.

Using the Biolign programme both forward and reverse sequences for each sample were compared with forward/ reverse sequence and SNPs were detected among 11 samples. Putative SNPs are shown in Table 3.3. Sample numbers of accessions are given in Table 3.3, the name of numbered accessions and their localization can be seen in Table 2.3. The 17 identified SNPs and their localization in the genome are given in Table3.3. The numbers before the nucleotide change refer to localization of SNPs which were written in the downstream region of primers.

Table 3.4. Sequence result based-nucleotidic substitutions and their localization for each genotype

	Genotype numbere	Nucleotidic Substitutions										
	SNPs	S1	S19	S23	S24	S58	S75	S78	S94	S95	S96	S97
1	11 T/C	T/C	T/T	T/C	T/T	T/C	T/T	T/T	T/T	T/C	T/C	T/T
2 Probe1	42 A/C	C/C	C/C	A/C	A/C	C/C	A/A	A/C	C/C	A/C	A/C	A/C
3	96 A/C	A/C	A/A	A/C	A/A	A/C	A/A	A/A	A/A	A/C	A/C	A/A
4	105 T/C	C/C	T/T	T/T	T/T	T/C	T/C	T/T	T/T	T/C	T/C	T/T
5	144 A/G	A/G	A/A	G/A	G/G	G/A	G/G	G/G	G/G?	G/A	G/A	G/G
6	151 A/G	G/G	G/G	G/A	G/A	G/G	A/A	A/G	G/G	G/A	G/A	G/A
7 Probe2	159 G/T	G/G	G/T	G/G	G/G	G/G?	G/G	G/T	G/G	G/T	G/G	G/G
8	280 A/C	G/G	G/G	G/C	G/G	G/G	G/G	G/G	G/G	G/C	G/C	G/G
9	233 C/G	C/A	A/A	A/C	A/A	A/C	A/A	A/A	A/A	A/C	A/C	A/A
10	306 T/A	T/A	T/T	T/A	A/A	T/A	A/A	A/T	A/T	A/T	A/T	A/A
11	312 T/C	T/C	C/C	T/C	T/T	T/C	T/T	T/C	T/C	T/C	T/C	T/T
12 Probe3	321 G/C	G/G	C/C	G/G	G/G	G/G	G/G	G/G	G/C	G/G	G/G	G/G
13	377 G/T	G/G	G/G	G/T	G/G	G/T	G/G	G/G	G/G	G/T	G/T	G/G
14	415 C/T	C/C	T/T	C/C	C/C	C/C	C/C	C/T	C/T	C/C	C/C	C/C
15 Probe4	450 A/G	G/G	A/A	G/G	G/G	G/G	G/G	G/A	G/A	G/G	G/G	G/G
16	458 G/A	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G
17	495 C/T	C/C	T/T	C/C	C/C	C/C	C/C	C/T	C/T	C/C	C/C	C/C

Ambiguous nucleotides were written with red colour in Table3.3 using results of sequence alignment. We could not find one SNP to separate the Memecik accession from others. For this reason, we designed 4 different SimpleProbe which were located on the 42 nd 159th 321th 450nd nucleotide of the gene to use in combination.

The probes were designed with regard to the Tm of probe and primer and possible loop formation by Roche Company using the LCPDS2 program. Probes and primer localization on the Ant. gene region are illustrated in Figure 3.11.

945958	pl Olive 94f		target	Tm
Ola F	CAgCAACAAgTgAgTATgCAAAAC	S	4-27	57,6°C
Ola R	gCTgAgggCATTTTgggTAAT	A	171-151	58,8°C
OL 1 spA	AACCCT XI AgTCCAATAgAgAgTgCC p	A	27-4	59,4°C
Olb F	AAgAAAgACTggAAAAAgAAgTTggT	S	88-113	58,5°C
Olb R	TgCACTgACgTCgCTgTg	A	215-198	58,6°C
OL 2 spT	CCCTCggTgTCgA XI AgCCC p	S	181-198	61,4°C
Olc F	CAAAATgTgTTCCCAACTCCATTATC	S	283-308	59,2°C
Olc R	TgggTggCTCACAgAAAACTg	A	429-409	59,9°C
OL 3 spG	gCAgAgCACT XI TTTATACTTCCCATT p	A	366-342	59,1°C
Old F	TgCACAgAgggCTTgTgAATAAA	S	364-386	59,8°C
Old R	CAACTCATTTTTCTTCACCggAATC	A	570-546	59,7°C
OL 4 spG	ggTggCTCAgCCTCAgA XI AACTg p	A	484-463	61,0°C

(cont. on next page)

Figure 3.10. Probe & primer sequences and illustration on the gene

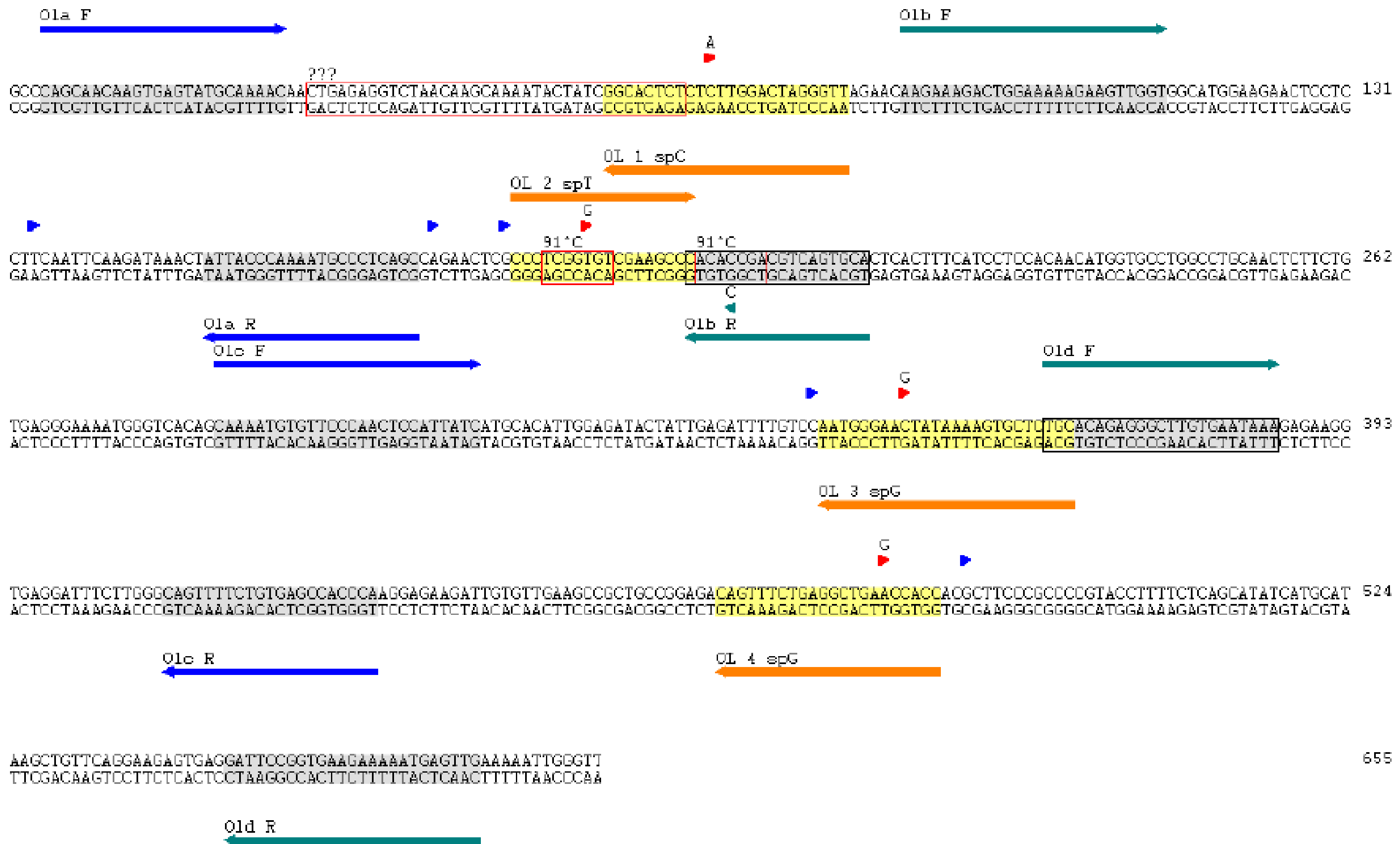

Figure 3.10. (cont.)

On the gene illustration, deep blue arrows and grey shading mark the primers while orange arrows and yellow shading mark 3'-FL labeled probes and blue arrows and blue shading were adduced the 5'-LC labeled probes. Red arrows, yellow shaded or red boxes were used to locate the position of TaqMan probes (SimpleProbe).

3.2.3.2. Genotyping Using the Designed Four Probes on LightCycler 480

After we performed the detection of SNPs based on sequence results and designed the proper probes, the genotyping was done using the probes on real-time-PCR to separate Memecik accession among Turkish olive cultivars. According to the real time-PCR working principle, strong interaction between probe and PCR product result in higher Tm ˚C, so we saw the high Tm for specific interaction of probes with PCR products and we assumed that samples which had peaks for higher melting ˚C were wild type. The peaks for probe1 (on the 42nd nucleotide) are illustrated in the Figure 3.12.

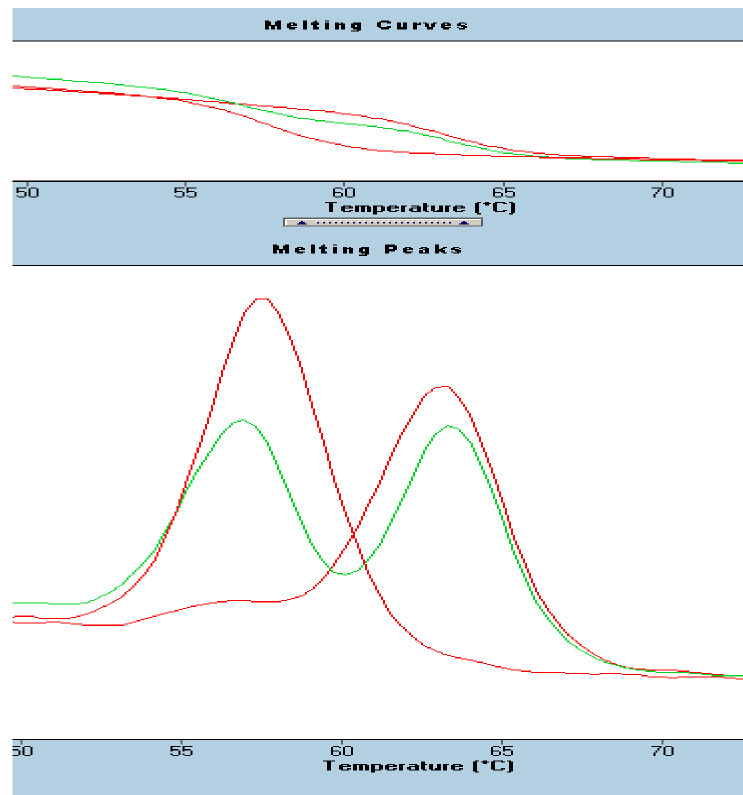

Figure 3.11. Real time PCR results of probe42

Three genotypes are illustrated in Figure 3.12. Green line shows the Memecik accessions which has 2 peaks at 56˚C and 63˚C. This means Memecik accession has A allele (at 63 ˚C) and C allele (at 56 ˚C). The 2 red lines show a different sample, one of them has only A allele and the other one has only C allele.

As I mentioned before, the higher melting temperature is seen at higher temperature. Therefore, for genotyping, melting peaks were used. Based on the principle of this analysis, the higher temperature peak was wild types and the lower peaks were from the mutants and if there were two peaks for one sample they were classified as heterozygous because one of DNA peaks was from the wild type and the other one was mutant allele. In Table 3.4., probes, the nucleotide assumed as wild type, complementary nucleotides, and their melting temperatures (˚C) are listed.

Table 3.5. Probe-Tm based nucleotide analysis

Probe localization, nucleotide substitution and its name	Arbiter nucleotide for designing probe	Complementer nucleotide (probe binds this nucleotide)	Melting temperatures and alleles
42, A/C , OL1SpA	T	A	High temp.(63 ˚C)→A (wild type) Low temp.(56 ˚C)→C (mutant)
159, G/T, OL2SpT	T	A	High temp.(62˚C) →T. (wild type) Low temp.(49˚C)→ G (mutant)
321, G/C, OL3Sp G	C	G	High temp.(59˚C)→G (wild type) Low temp. (49˚C)→C (mutant)
450, A/G, OL4SpG	C	G	High temp.(65˚C) →G (wild type) Low temp.(57˚C)→A (Mutant)

We tried our four probes on all accessions and probe-based results are given in Table 3.5. Also the expected results are included in this table and marked with '•' symbol.

Table 3.6. Real time PCR results and sequence results of the probe binding regions

SNPs	1	• 2	2	3	4	5	6	• 7	7	8	9	10	11	• 12	12	13	14	• 15	15	16	17
	11 T/C	42 A/C	42 A/C	96 A/C	105 T/C	144 A/G	151 A/G	159 G/T	159 G/T	233 C/G	280 A/C	306 T/A	312 T/C	321 G/C	321 G/C	377 G/T	415 C/T	450 A/G	450 A/G	458 G/A	495 C/T
S1	T/C	C/C	C/C	A/C	C/C	A/G	G/G	G/G	G/G	G/G	C/A	T/A	T/C	G/G	G/G	G/G	C/C	G/G	G/G	G/G	C/C
S19	T/T	C/C	C/C	A/A	T/T	A/A	G/G	G/T*2	T/T	G/G	A/A	T/T	C/C	C/C	C/C	G/G	T/T	A/A	A/A	G/G	T/T
S23	T/C	A/C	C/A	A/C	T/T	G/A	G/A	G/G	G/G	G/C	A/C	T/A	T/C	G/G	G/G	G/T	C/C	G/G	G/G	G/G	C/C
S24	T/T	A/C	C/A	A/A	T/T	G/G	G/A	G/G	G/G	G/G	A/A	A/A	T/T	G/G	G/G	G/G	C/C	G/G	G/G	G/G	C/C
S58	T/C	C/C	C/C	A/C	T/C	G/A	G/G	G/G*3	T/T	G/G	A/C	T/A	T/C	G/G*5	C/G	G/T	C/C	G/G*8	A/G	G/A	C/C
S75	T/T	A/A	A/A	A/A	T/C	G/G	A/A	G/G	G/G	G/G	A/A	A/A	T/T	G/G	G/G	G/G	C/C	G/G	G/G	G/G	C/C
S78	T/T	A/C	C/A	A/A	T/T	G/G	A/G	G/T	G/T	G/G	A/A	A/T	T/C	G/G*6	C/G	G/G	C/T	G/A	G/A	G/G	C/T
S94	T/T	C/C	C/C	A/A	T/T	G/G	G/G	G/G*4	T/T	G/G	A/A	A/T	T/C	G/C	G/C	G/G	C/T	G/A	G/A	G/G	C/T
S95	T/C	A/C	C/A	A/C	T/C	G/A	G/A	G/G	G/G	G/C	A/C	A/T	T/C	G/G	G/G	G/T	C/C	G/G	G/G	G/G	C/C
S96	T/C	A/C*1	C/A?	A/C	T/C	G/A	G/A	G/G	G/G	G/C	A/C	A/T	T/C	G/G*7	C/G	G/T	C/C	G/G	G/G	G/G	C/C
S97	T/T	A/C	A/C	A/A	T/T	G/G	G/A	G/G	G/G	G/G	A/A	A/A	T/T	G/G	G/G	G/G	C/C	G/G	G/G	G/G	C/C

To exhibit nucleotides which were detected with sequence analysis and compare with real-time PCR results, Table 3.5 was prepared can be explained by the following comments.

*1→ As to sequence analysis, there could be C/C but there was T with noise doubt on reverse sequence..

*2→ T/T on real time PCR but there was doubt on sequence analysis with noise doubt.

*3→ G/G without artifact on sequence analysis.

*4→ There was G peak clearly on sequence analysis with artifact T base doubt.

*5→ Clearly G/G there were A and T base as noise but not C.

*6 → As to forward sequence there was C base like noise

*7 → It was clearly G/G.

*8 → G/G sequences with A and C artifacts.

Therefore, we have tried four probes. We separated the accessions which were numbered 1(Trabzon Yağlık), 19 (Samsun Tuzlamalık), 58 (Eğriburun-Antep) and 94 (Memecik) from the others by probe1 (on 42th nucleotide).

Using probe2 (on the 159 th nucleotide), 94 (Memecik), 78 (Erkence) were determined to fall into the same group. Probe3 located on 312nd separated 94(Memecik), 23(Erdek Yağlık), 58 (Eğriburun-Antep), 78 (Erkence), 96 (Kiraz) accessions from the rest and the 4th probe separated genotype 94 (Memecik), 58 (Eğriburun-Antep) 78 (Erkence) which carried G/A alleles. Our first aim was to find a single probe to determine Memecik accessions among the others.

However we could not find the Memecik-specific SNP so we designed more than one probe to make a diagnostic combination. According to our result, if we used 159th probe and 42nd probe one after another only Memecik accessions would give C/C alleles for 42th probe and G/T alleles for 159th probe as shown in Table 3.6.

Table 3.7. Genotypes whcih have the same alleles

Genotype	Probe 42	Genotype	Probe 159	Genotype	Probe 321	Genotype	Probe 450
1	C/C	78	G/T	23	G/C	58	G/A
19	C/C	94	G/T	58	G/C	78	G/A
58	C/C			78	G/C	94	G/A
94	C/C			94	G/C		
				96	G/C		

According to the results we determined Memecik accessions among the sequenced accessions using two probes simultaneously. But when we examined all

cultivars (66 Turkish accessions and 3 outgroups) without previous sequence knowledge and comparison, we found more than one sample which had the same mutations at the same SNP, so the probe combination was not useful to trace the Memecik accessions among all Turkish cultivars (data not shown).

Also, Trabzon Yağlık accession gave the A/A alleles by probe450 so, it is separated from the rest sequenced accessions by one step and then, Probe450 was applied to 66 Turkish accessions and 3 outgroups as a result, only Trabzon Yağlık accession gave the A/A alleles among 69 accessions. In other words, we developed SNP marker to trace Trabzon Yağlık accessions in olive oil.

CHAPTER 4

CONCLUSIONS

Olive accessions in Turkey are named based on their phenotypic features but we know the environment effect on phenotype so visually selection and breeding may possibly cause mislabeling of them. In addition, there is no concern to breed different olive accessions which have special taste or low acidity by Turkish agriculturists. Anatolia is called the origin of olive trees however no Turkish accession stands in the forefront for olive oil in the world. Maybe this is because most producers mix the olive before pressing and there is no specific taste for one accessions.

To resolve some of these issues firstly we tried to characterize the Turkish olive cultivars by using SSR and SRAP markers. We drew trees for each marker system and nearly the same clusters were found. 13 SSR marker were used to determine the diversity and the tree was scaled from 0.45 (min.) to 0.90 (max.) similarity while the SRAP marker–based tree had 0.66. minimum similarity. This difference is explained by the conservative nature of SRAP markers which amplify open reading frames instead of microsatellite regions like SSR markers. The distance matrix of two marker was compared to see the correlation and it was found to be low as expected because SRAP markers calculate the closer relationship among cultivar. In both trees, nearly the same accessions were found to be most diverged from the majority of Turkish olive accessions. To support the data, ordination test was performed and it gavee 2-D matrix comparison results similar to what we expected based on dendrogram results.

To develop a traceability test for Memecik oil, we sequenced the 11 olives which are exported as oil and found the single nucleotide polymorphisms and then designed the SimpleProb to use melting curve analysis. We separated the Memecik accession from the other important exported olives. Also we performed the analysis for all cultivars but there were other accessions which had the same single nucleotide polymorphisms as Memecik. This was expected because it is known that growers may carry olives from region to another regions where a different phenotype may be seen because of environment effects. As a result the olive may be renamed over time but still retain high genetic similarity with the original cultivar. Also, we found Trabzon Yağlık accession

specific single nucleotide substitution called as probe4 (probe450) for 11 exported accessions and also we applied this probe to all cultivar and only Trabzon Yağlık gave us the A/A alleles. Therefore, probe4 (probe450) can be used to determine Trabzon Yağlık accession as one step SNP marker.

REFERENCES

Aranzana, M.J., Pineda, A., Cosson, P., Dirlewanger, E., Ascasibar, J., Cipriani, G., Ryder, C.D., Testolin, R., Abbott, A., King, G.J. and Iezzoni, A.F. 2003. P.Arus A set of simple-sequence repeat (SSR) markers covering the Prunus genome. *TAG Theoretical and Applied Genetics* 106:5.

Barranco, D., and Rallo, L., 2000. Olive cultivars in Spain. *Hortech.*10(1):107–110.

Brookes, A.J., Gyllensten, U., Howell, W.M. and Jobs, M. 1999. Dynamic allele-specific hybridization. A new method for scoring single nucleotide polymorphisms. *Nature Biotechnol.*17:87-88.

Canozer, O., 1970. *Yerli Zeytin Cesitlerinin Özelliklerinin Tesbitine Ait Arastırma ve Koleksiyon Zenginligi Tesisi (Sonuc Raporu).* Zeytincilik Arastırma Enstitusu Yayınları.

Canozer, O., 1991. *T.C. Tarım ve Koyisleri Bakanlıgı, Standard Zeytin Cesitleri Katalogu.* Ankara. Yayın Dairesi Baskanlığı.

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

Chakravarti, A. 1998. It's raining SNPs hallelujah? *Nat. genet.* 19(3):216-217.

Chiapparino, E., Donini, P. and Lee, D. 2004. Genotyping single nucleotide polymorphisms in barley by tetraprimer. ARMS-PCR. *Genome* 47:414–420.

Cirik, M.N. 1988. Farklı Iki Ekolojide Bazı Zeytin Çesitlerinin Çiçek Tomurcugu Gelisimi, Somak ve Çiçek Morfolojileri Üzerine Arastırmalar. Tarım ve Köyisleri Bakanlıgı Zeytincilik Arastırma Enstitüsü Yayınları, İzmir, Turkiye.

Collins, F.S., Guyer, M.S. and Chakravarti, A. 1997. Variations on a theme. Cataloging human DNA sequence variation. *Science* 278(5343):1580-1581.

Cuevas, J., Diaz-Hermoso, A.J., Galian, D., Hueso, J.J., Pinillos, V., Sola, D. and Polito, V.S. 2001. Response to cross pollination and choice of pollinisers for the olive cultivars (Olea europaeaL.) 'Manzanilla de Sevilla', 'Hojiblanca' and 'Picual'., *Olivae* 85: 26-32.

De la Rosa, R., James, C.M. and Tobutt, K.R. 2002. Isolation and characterization of polymorphic SSRs in olive (Olea europaea L.) and their transferability to other genera in the Oleaceae. *Mol. Ecol. Notes*.2:265–267.

Doyle, J.J and Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Fabbri, A., Hormaza, J.I. and Polito, V.S. 1995. Random amplified polymorphic DNA analysis of olive (Olea europaea L.) cultivars. 1. *Amer. Soc. Hort. Sci*. 120:(3):538-542.

Ferriol, M., Nuez, F. and Pico, B. 2003. Genetic diversity of some accessions of Cucurbita maxima from Spain using RAPD and SRAP markers. *Genetic Resources and Crop Evolution* 50:227–238.

Gupta, P.K., Mir, R.R. and Rustgi, S. 2008. Array-based high-throughput DNA markers for crop improvement. *Heredity* 101(1):5-18.

Grigg, D. 2001. Olive oil, the Mediterranean and the world. *GeoJournal*.53:163-172.

http://faostat.fao.org/site/339/default.aspx

Kwok, P.Y. and Chen, X. 1998. Detection of single nucleotide variations. *Genet Eng*. 20: 125-134.

Kwok, P. Y. 2001. Methods for Genoyping Single Nucleotide Polymophisms. *Annu.Rev. Genomics Hum. Genet*.2:235-58.

Kwok, P.Y. and Chen, X. 2003. Detection of Single Nucleotide Polymorphisms. *Caister Academic Press* 5: 43-60.

Li, G. and Quiros, C.F., 2001. Related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica *Theor Appl Genet*.103:455–461.

Mukhlesur, R. L., Li, G. and Mc Vetty, B. E. P., 2007. Development of SRAP, SNP and Multiplexed SCAR molecular markers for the major seed coat color gene in Brassica rapa. *Theor Appl Genet*.115:1101–1107 DOI 10.1007/s00122-007-0636-8123.

Mooney, S. 2005. Bio informatics approaches and researces for single nucleotide polymorphism functional analysis. *Briefings in bioinformatics*.6(1):44-56.

Panagiotakos, D.B., Pitsavos, Ch., Chrysohoou, Ch., Stefanadis, Ch. and Toutouzas, P. 2002. The role of traditional mediterranean type of diet and lifestyle, in the development of acute coronary syndromes. Preliminary results from CARDIO 2000 study. *Cent Eur J Public Health*. 10 (1-2):11-5.

Rafalski, A., 2002. Applications of single nucleotide polymorphism in crop genetics. *Curr. Opin. Plant Biol*.5:94-100.

Reale, S., Doveri, S., Diaz, A., Angiolillo, A., Lucentini, L., Pilla, F., Martin, A., Donini, P. and Lee, D. 2006. SNP-based markers for discriminating Olive (Olea europaea L.) cultivars. *Genome*49:1193-1205.

Sato, K. and Takeda, K. 2009. An application of high-throughput SNP genotyping for barley genome mapping and characterization of recombinant chromosome substitution lines *Theor Appl Genet*.119:613–619 DOI 10.1007/s00122-009-1071-9.

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

Somers, J., Kirkpatrick, R., Moniwa, M. and Walsh, A. 2003. Mining single-nucleotide polymorphisms from hexaploid wheat ESTs Daryl. *Genome* 49:431–437 doi: 10.1139/G03-027.

Taamalli, W., Geuna, F., Bassi D., Daoud, D. and Zarrouk, M. 2008. Agronomic and molecular analyses for the characterisation of accessions in Tunisian olive germplasm collections. *Electron. J. Biotechnol* 19: 467-481.

Tenaillon, M.I., Sawkins, M.C., Anderson, L.K., Stack, S.M., Doebley, J. and Gaut, B.S. 2002. Patterns of diversity and recombination along chromosome 1 of maize (Zea mays ssp. mays L.). *Genetics* 162(3):1401-13.

Testolin, R. and Lain, O. 2005. DNA Extraction from Olive Oil and PCR Amplification of Microsatellite Markers. *Journal of Food Science. 70:1.*

Tuck, K.L., Hayball, P.J. and Stupans, I. 2002. Structural characterisation of the metabolites of hydroxytyrosol, the principal phenolic component in olive oil, in rats. *Journal of Agricultural and Food Chemistry* 50:2404-2409.

Vinha A.F., Ferreres F., Silva B.M., Valentiao P., Goncalves A., Pereira J.A., Oliveira M.B., Seabra R.M., Andrad P.B.,2005: Phenolic profiles of Portuguese olive fruits (Olea europaea L.):Influences of cultivar and geographical origin. Food Chemistry 89: 561–568.

Visioli F., Poli A. and Galli C. 2002. Antioxidant and other biological activities of phenols from olives and olive oil. *Medicinal. Research Reviews* 22 : 65-75.

Weising K., Winter P., Huttel B., Kahl G. 1998. Microsatellite markers for molecular breeding. *J Crop Prod*. 11:113-43.

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

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