GENETIC DIFFERENTIATION OF LIQUIDAMBAR ORIENTALIS MILL. VARIETIES WITH RESPECT TO *matK* REGION OF CHLOROPLAST GENOME

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ASLI ÖZDİLEK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
BIOLOGY

AUGUST 2007

Approval of the Thesis

GENETIC DIFFERENTIATION OF *LIQUIDAMBAR ORIENTALIS* MILL. VARIETIES WITH RESPECT TO *matK* REGION OF CHLOROPLAST GENOME

submitted by **ASLI ÖZDİLEK** in partial fulfillment of the requirements for the degree of **Master of Science in Biology Department, Middle East Technical University** by,

Prof. Dr. Canan ÖZGEN _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Zeki Kaya _____
Head of the Department, **Biology**

Prof. Dr. Zeki Kaya _____
Supervisor, **Biology Dept., METU**

Examining Committee Members

Prof. Dr. İnci Togan _____
Biology Dept., METU

Prof. Dr. Zeki Kaya _____
Biology Dept., METU

Prof. Dr.Hayri Duman _____
Biology Dept., Gazi University

Assoc. Prof. Dr. Sertaç Önde _____
Biology Dept., METU

Assoc. Prof. Dr. İrfan Kandemir _____
Biology Dept., Zonguldak Karaelmas University

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Aslı ÖZDİLEK

Signature :

ABSTRACT

GENETIC DIFFERENTIATION OF LIQUIDAMBAR ORIENTALIS MILL. VARIETIES WITH RESPECT TO *matK* REGION OF CHLOROPLAST GENOME

ÖZDİLEK, Aslı

M.S., Department of Biology

Supervisor: Prof. Dr. Zeki Kaya

August 2007, 87 pages

Liquidambar L. genus is represented with mainly 4 species in the world and one of these species, Turkish sweet gum (*Liquidambar orientalis* Mill.) which is a relict-endemic species is naturally found in only southwestern Turkey, mainly in Muğla Province. The limited distribution of species with two disputed varieties (var. *integriloba* Fiori and var. *orientalis*) and increased anthropogenic threats to its genetic resources signify the importance of studying genetic diversity in the species to have better conservation and management programs. For this purpose, 18 different populations were sampled throughout the species range and *matK* region of chloroplast DNA (cpDNA) was sequenced to assess the genetic structure of the species. Turkish *Liquidambar orientalis* populations were evaluated at two categories: variety level and geographic level. Also, two sectors of *matK* region were examined to assess which part of the region was more variable. All molecular analysis was conducted in this study by using MEGA version 3.1 and Arlequin 2.000 softwares.

Moleculer diversity analysis indicated that the population located in Fethiye-Günlükbaşı district has the highest number of polymorphic sites. This population is also genetically the most distant from the others (average genetic distance 0.0038). Among the studied varieties, the average genetic distance within var. *integriloba* (0.0016) which also includes population Fethiye-Günlükbaşı was the greatest. Among the geographic regions, Muğla-1 including Fethiye-Köyceğiz-Aydın district as well as population Fethiye-Günlükbaşı showed the highest average genetic distances within the region with a value of 0.0015. According to the molecular variance results, among varieties and among geographic regions, there was no significant differentiation, but great amount of total variation was found (~86%) within Turkish sweet gum populations. With respect to the F_{st} values among varieties, the highest genetic differentiation was observed between var. *orientalis* and unknown group (0.040). Furthermore, based on the results of phylogenetic analysis, Turkish populations of *L. orientalis* have genetically closer to USA relative (*L. styraciflua* L.) than Chinese relatives (*L. acalycina* H.T Chang and *L. formosana* Hance).

In conclusion, 10 Turkish sweet gum populations were found to be important for conservation issues. Furthermore, eight of these located in Muğla province and sixth of them belong to var. *integriloba*. Especially Fethiye-Günlükbaşı, Marmaris-Çetibeli and Muğla-Kıyra populations should be included in either *insitu* or *exsitu* or in both conservation programs in the future.

Keywords: *Liquidambar orientalis*, *matK* gene, cpDNA

ÖZ

KLOROPLAST GENOMUNDAKİ *matK* GEN BÖLGESİNE GÖRE LIQUIDAMBAR ORIENTALIS MILL. VARYETELERİNİN GENETİK FARKLILAŞMASI

ÖZDİLEK Aslı
Y. Lisans, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. Zeki Kaya

Ağustos 2007, 87 sayfa

Dünya üzerinde 4 ana türle temsil edilen *Liquidambar* L. cinsinin bir türü de güneybatı anadoluda özellikle Muğla ili dolaylarında doğal olarak yayılış gösteren Türk sığlası yada Anadolu sığlası (*Liquidambar orientalis* Mill.)'dır. Türk sığlası Türkiye için relikt-endemik bir türdür. İki varyetesiyle bilinen türün sınırlı bir bölgede yayılış göstermesi (var. *integriloba* Fiori and var. *orientalis*) gen kaynaklarına olan tehditleri arttırmaktadır. Bu da türün genetik çeşitliliğinin çalışılmasıyla iyi bir koruma stratejisinin geliştirilmesinin ve gen kaynaklarının idaresinin önemini ortaya koymaktadır. Bu nedenle, türün 18 farklı toplumu örneklendi ve kloroplast DNA (cpDNA)' sında bulunan *matK* bölgesinin dizi analizi yapılarak türün bu bölge için genetik yapılaşması belirlendi. Türk sığla toplumlarının *matK* DNA dizi verileri iki aşamada değerlendirildi: varyeteler seviyesi ve coğrafik bölgeler seviyesi. Ayrıca, *matK* gen bölgesi iki kısımda incelenerek hangi kısmın daha fazla çeşitlilik gösterdiği belirlendi. Tüm veri setleri MEGA 3.1 ve Arlequin 2.000 bilgisayar programlarıyla değerlendirildi.

Moleküler çeşitlilik analizi, Fethiye-Günlükbaşı bölgesinde bulunan toplumun fazla miktarda polimorfik bölgeye sahip olduğunu göstermektedir. Bu toplum aynı zamanda genetik açıdan diğer toplumlardan en farklı olan toplumdur (ortalama genetik mesafe 0.0038). Çalışılan tüm varyeteler arasında, aynı zamanda Fethiye-Günlükbaşı toplumunu da içine alan varyete *integriloba*, varyete içi ortalama genetik mesafe olarak en yüksek değere sahiptir (0.0016). Coğrafik bölgeler arasında, Fethiye-Köyceğiz-Aydın bölgesini içeren Muğla-1, Fethiye-Günlükbaşı toplumunda olduğu gibi 0.0015 değeriyle en yüksek coğrafik bölge içi ortalama genetik mesafe değerini göstermektedir. Moleküler varyans sonuçlarına göre, varyeteler ve coğrafik bölgeler arasında önemli bir farklılık olmadığı, fakat Türk sığlası toplumları içinde yüksek oranda varyasyon bulunduğu saptanmıştır (~86%). Varyeteler arasındaki F_{st} değerlerine göre en yüksek genetik farklılaşmanın varyete *integriloba* ile bilinmeyen grup arasında olduğu belirlendi. Ayrıca, bu çalışmada elde edilen sonuçlar, Türk sığla toplumlarının (*L. orientalis*) genetik olarak Amerika'daki akrabasına (*L. styraciflua*), Çin'deki akrabalarına (*L. acalycina* H.T Chang and *L. formosana* Hance) oranla daha yakın olduğunu göstermektedir.

Yapılan tüm analiz sonuçlarına göre, 10 Türk sığlası toplumu, koruma kapsamına alınması bakımından önemli bulunmuştur. Ayrıca, çalışılan 10 toplumdan 6'sının varyete *integriloba*'ya ait olduğu ve bu 10 toplumun 8'inin Muğla bölgesinde bulunduğu saptanmıştır. Özellikle, Fethiye-Günlükbaşı, Marmaris-Çetibeli ve Muğla-Kıyra toplumları gelecekte, *insitu*, *exsitu* veya her iki koruma programı kapsamında dikkate alınması gerekmektedir.

Anahtar Kelimeler: *Liquidambar orientalis*, *matK* gene, cpDNA

to my unique family...

ACKNOWLEDGEMENTS

I am greatly indebted to my supervisor, Prof. Dr. Zeki Kaya for his guidance, supervision and endless patience throughout the study.

I would like to express my thanks to all jury members for their helpful comments and criticisms on the manuscript.

I wish to express my deep gratitude to Dr. Yasemin İçgen, Dr. Burcu Çengel, Dr. Gaye Kandemir and Assoc. Prof. Dr. İrfan Kandemir for their continuous help, advice, collaboration and encouragement in every step of this study.

I wish to thank Ercan Velioğlu, Süleyman Işık Derilgen, Murat Nur, Belkıs Korkmaz, Osman Eşe and Rumi Sabuncu for providing leaf sources.

I would like to thank all my colleagues from Department of Biology, Plant Genetics and Tissue Culture Laboratory for their support and friendship. I want to thank also my friends Ferhunde Aysin and Nihal Şimşek Özek.

Special thanks to my mother Hamide Özdilek, my father Sururi Özdilek, my lovely sister Arzu Özdilek and my lovely little brother Kerem Özdilek.

This thesies was supported by the research fund: TUBITAK TOVAG 104O154 and BAP-08-11-DPT2002K12510.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMOVA	Analysis of Molecular variance
β-ME	Beta Mercaptoethanol
cpDNA	Chloroplast DNA
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Tri-Phosphate
EDTA	Ethylene Diamine Tetra Aceticacid
ETOH	Ethanol
ITS	Internal Transcribed Spacer Region
IUCN	The World Conservation Union
MEGA	Molecular Evolutionary Genetic Analysis
MST	Minimum Spanning Tree
mtDNA	Mitochondrial DNA
matK	maturase K
NCBI	National Center for Biotechnology Information
NJ	Neighbour-joining
nrDNA	Nuclear ribosomal DNA
ORF	Open Reading Frame
OTUs	Operational Taxonomical Units
PCR	Polymerase Chain Reaction
rbcL	RuBisCo Large subunit
SDS	Sodium Dodecyl Sulphate
TAE	Tris-Acetate EDTA
TE	Tris EDTA Buffer
t-RNA	Transfer Ribonucleic Acid
trnK	Lysine tRNA,

CHAPTER I

INTRODUCTION

1.1. *Liquidambar* L.

Sweet gum (*Liquidambar* L.) is a genus with four main species (*L. formosana*, *L. acalycina* spread in central & southern China, *L. styraciflua* found in eastern north America (Bogle, 1986) and *L. orientalis* naturally found in southwest of Turkey) (Figure 1.1). It is a flowering plant in the subfamily Bucklandioidae, family Altingiaceae, though formerly the species often treated in the Hamamelidaceae (Örtel, 1988). *Liquidambar* is the unique genus in the Hamamelidaceae that has a disjoined distribution with species occurring in western Asia, eastern Asia and North America (Li *et al*, 1997). All members of *Liquidambar* species are large, deciduous trees (Figure 1.2), generally 25-40 m tall, with palmately lobed leaves (Figure 1.3) arranged spirally on the stems. The flowers are small, produced in a dense globular inflorescence 1-2 cm diameter (Figure 1.4), pendulous on a 3-7 cm stem. The fruit is a woody multiple capsule 2-4 cm diameter (popularly called a 'gumball'), containing numerous seeds (Figure 1.5).

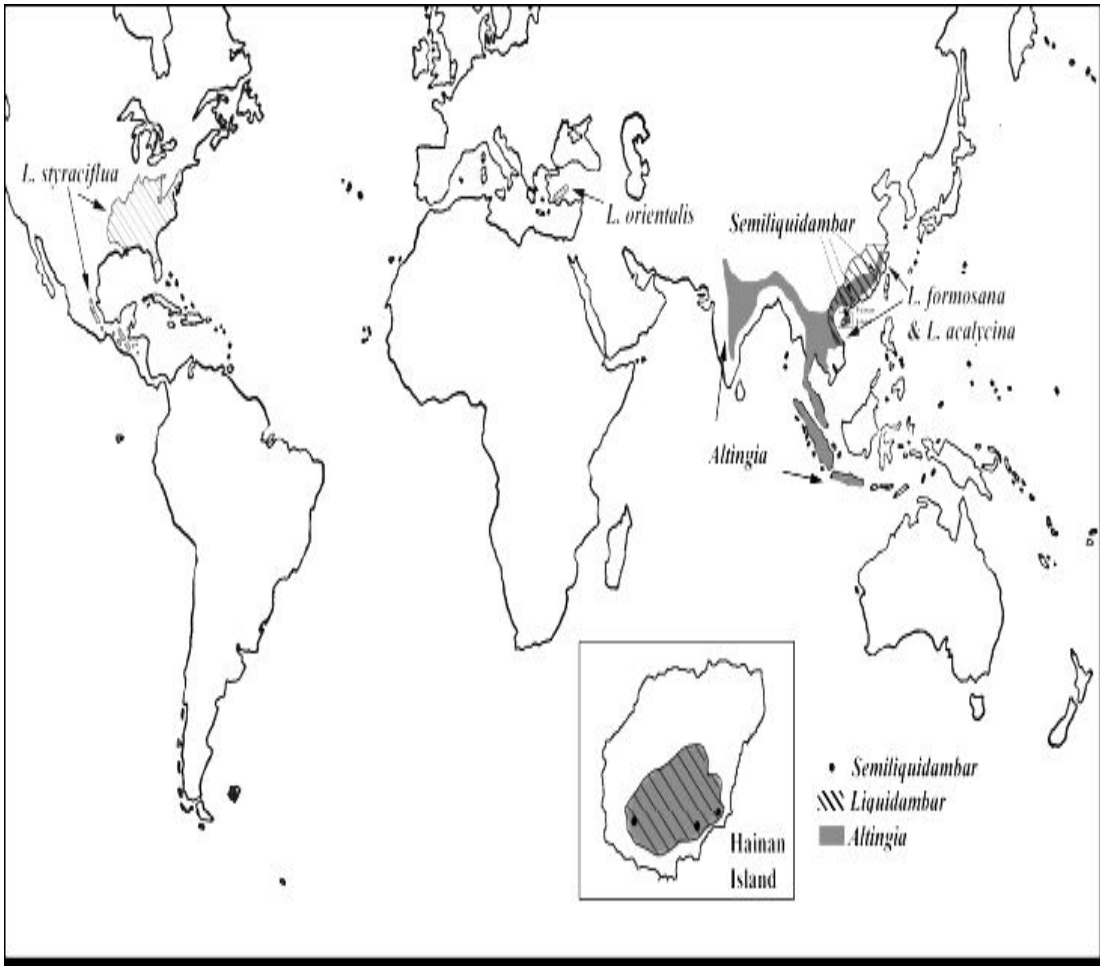


Figure 1.1 The map indicates the distribution of Altingiaceae family including *Liquidambar* species (Ickert-Bond *et al*, 2005)

1.1.1 *Liquidambar orientalis*

Turkish sweet gum is a deciduous tree growing up to 20m at a slow rate. It is hardy to zone but is frost tender and flower in May, its seeds ripen from October to November. Form of the *L. orientalis* leaves designates differences within the species (Efe, 1987). The flowers are unisexual and monoecious (individual flowers are either male or female), but both sexes can be found on the same plant. They are in bloom from March to April (Alan and Kaya, 2003) and pollinated by bees.

The species grows on slopes and dry soil, but prefers rich, deep and moist soils such as bogs, river banks and coastal areas. It can grow in semi-shade (light woodland) or no shade.

Turkish sweet gum proliferates by sprouting suckers, and in reasonable conditions, natural regeneration is also possible (Alan and Kaya, 2003).

Turkish sweet gum has a balsam obtained from the wood and inner bark. It is used for both as food additives (e.g. as a chewing gum and stabilizer for cakes) and medicinal purposes (e.g. as an irritant, expectorant, skin salve, astringent face lotion, antibacterial, anti-inflammatory, antiseptic, pectoral and stimulant). It is also taken internally for the treatment of strokes, infantile convulsions, coma, heart disease, pruritis and treatment of cancer.

Based on the result of previous studies (Acatay, 1963; Önal and Özer, 1985; Efe, 1987), Turkish sweet gum has three reported varieties in Turkey. These are;

1. *L. orientalis* var. *orientalis*
2. *L. orientalis* var. *integribola*
3. *L. orientalis* var. *suber*

However, there is no genetic or well-founded systematic data that would prove these proposed varieties to be valid. The results of the studies by Peşmen (1972) and Efe (1987) are contradicting. For example, the finding of a revision study for 'Flora of Turkey' on two varieties (var. *orientalis* and var. *integriloba* Fiori) classified, based on presence and absence of lobes on leaves, were not observed by the field studies of Efe (1987).

Figure 1.2 The tallest Turkish sweet gum tree in Turkish sweet gum (*Liquidambar orientalis* Mill.) Forest located Nature Conservation Area in Isparta (Fakir and Doğanoğlu, 2003)

Figure 1.3 The palmately five lobed leaves of *L. orientalis* from Muğla Province, Turkey (Photo: Murat Alan)

Figure 1.4 A close up of the flowers of *L. orientalis* (Photo: Mehmet ali Başaran)

Figure 1.5 *L. orientalis'* fruit ('gumball') with seeds (Photo: Mehmet ali Başaran)

1.1.2 Distribution of the species

Turkish sweet gum has a limited distribution within an altitude range between 0 and 1100m in Turkey. It has naturally distributed in Anatolia among the borders of Çine Stream in Aydın from north, seashore of Eşen Stream (Kocaçay) in Muğla to south, in Silifke from east and in Bodrum to west (Dirik, 1986). Muğla province is the main distribution area of the species (Figure 1.6).

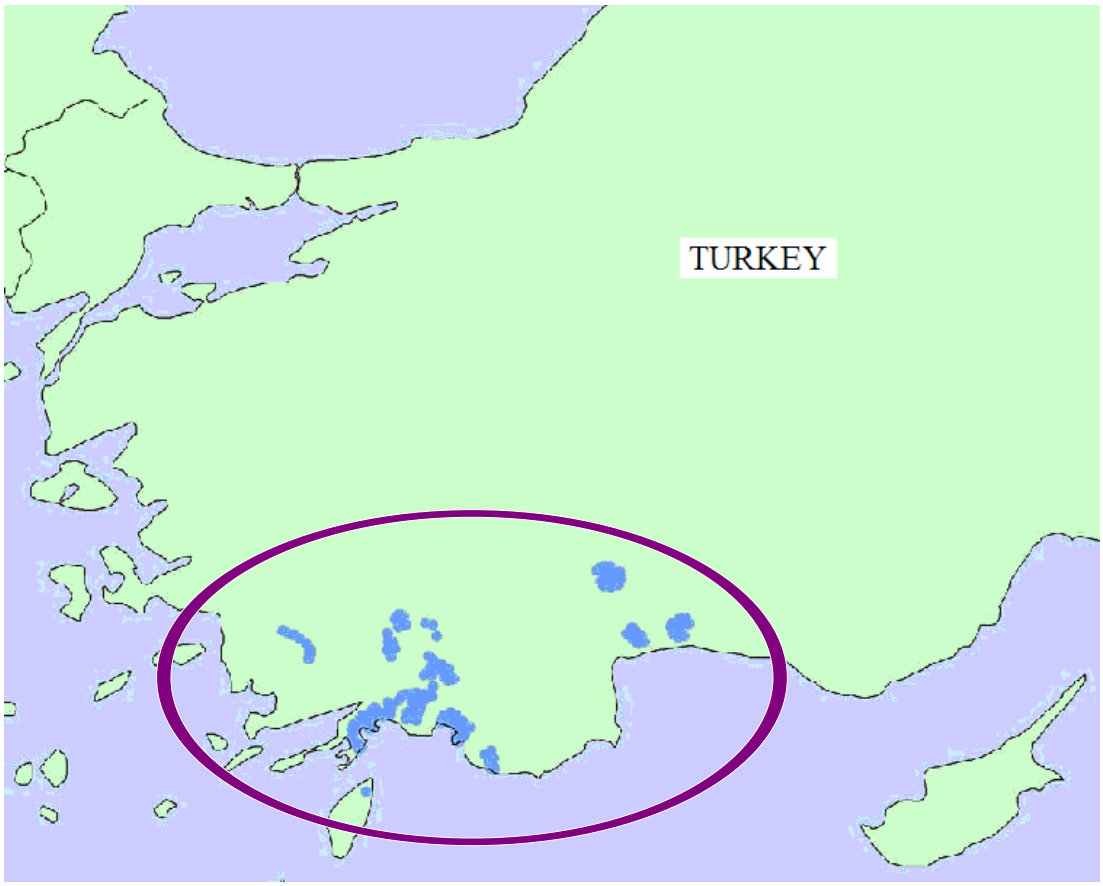


Figure 1.6 Natural distribution of the Turkish sweet gum (Alan and Kaya, 2003)

1.1.3 Importance of the species

Turkish sweet gum is a relict-endemic species in Turkey, but due to its presence in Rhodes Island, some considered that Turkish sweet gum is not an endemic species. On the other hand, it was suggested that this species may be taken from Turkey to the Cyprus and Rhodes Islands and cultivated in early days (Huş, 1949; Akman, 1995).

It is an economically important species because of its natural balsam producing ability. Sweet gum oil (Styrax Liquidus) is used in chemistry, pharmacology, and

cosmetic industry. It has brownish-yellowish color and specific aroma when it is fresh, including cinnamic acid (acid of cinnamon), sytracin, sytrol, sytron, storesinol and styrogenin. Sweet gum oil has a fixative function in perfumes and used in soap production. All oil produced is exported, providing an income for local people. The species also has some value as an ornamental, due to its attractive form and colour (Alan and Kaya, 2003).

1.1.4 Current status of Turkish sweet gum forests and conservation program

Turkish sweet gum species is threatened today. There is a sharp decrease in natural distribution of the species from 6312 ha in 1949 (Huş, 1949) to 1337 ha in 1987 (İktüeren and Acar, 1987). Since it has a highly restricted distribution, species is considered in vulnerable category by 'The World Conservation Union' (IUCN).

Turkish sweet gum forests have been badly damaged because of opening of irrigation canal, pasturage and the continuous export of balsam (poor oil production methods), especially in the period 1968-1979. Generally, the stems of trees are purposely wounded for the production of the balsam. In addition, the land covered by the species has been under continuous pressure from the local population. Trees are cut down and forests are cleared in order to gain arable land. For all these reasons, the occurence of Turkish sweet gum has been greatly reduced. At present, about 1200 hectares of natural Turkish sweet gum forests remain (Efe, 1987).

Although there is very little information available on the genetics of this species, its ecological and biological characteristics provide some indication about the patterns of genetic diversity. Trees growing from sea level to 400 m, are known as "plain sweet gum", while trees at higher altitudes are "mountain sweet gum". Trees growing at the higher altitudes form small groups and tolerate frost better.

Today, because of forest fires (Turkish sweet gum trees include oil in their stems and this causes fires easily), poor oil production and continuous pressure from the local population, this valuable species faces serious problems which may lead to vanishing, although, there are two gene conservation forests located in Isparta and Muğla Provinces, one seed orchard in Fethiye-Göcek, and finally two seed stands in Fethiye and Marmaris districts (Forest Trees and Seeds Breeding Research Directorate, http://www.ortohum.gov.tr, last visited August 2007).

1.2 DNA sequence studies in plant systematics

Since, nucleic acid sequencing is a powerful technique, the data generated by DNA studies have made it become one of the most utilized of the molecular approaches for inferring phylogenetic history. DNA sequence data are one of the most informative tool for molecular systematics. Thus, comparative analysis of DNA sequences is becoming increasingly important in plant systematics because the characters (nucleotides) are the basic units of information encoded in all organisms. Thus, for most studies, systematically informative variation is essentially inexhaustible. Furthermore, different genes or parts of the genome might evolve at different rates. Therefore, questions at different taxonomic levels can be addressed using different genes or different regions of a gene as well as different genomes (organelle or nuclear) (Liang, 1997).

1.2.1 Chloroplast DNA

Chloroplast is an organelle found in eukaryotic cells, but it has a genetic system of prokaryotic origin. Plants have a chloroplast genome (cpDNA) in addition to the nuclear (nDNA) and mitochondrial (mtDNA) genomes. The nuclear genome is used in systematic botany less frequently, because it has a complex and repetitive

characteristic. Because of its rapid changes in its structure, size, configuration, and gene order, the mitochondrial genome is used at the species level.

cpDNA;

1) is a relatively abundant component of plant total DNA, thus facilitating extraction and analysis;

2) contains primarily single copy genes;

3) has a conservative rate of 2 nucleotide substitution; and extensive background for molecular information on the chloroplast genome is available. Therefore, most phylogenetic reconstructions in plant systematics conducted so far is focused on molecular data generated from the cpDNA genes (Liang, 1997).

1.2.2 The maturase Kinase (*matK)* gene

The chloroplast gene maturase Kinase (*matK)* which is an open reading frame (ORF), located within the intron of *trnK* (lysine tRNA) gene, encodes a maturase, protein, used in RNA splicing (Neuhaus and Link, 1987; Wolfe *et al*, 1992; Mort *et al*, 2001)

The $tRNA^{Lys}$(UUU) gene (*trnK*) contains a group II intron and this group II intron encodes the *matK* (Hausner *et al*, 2006)

Group II introns are self-splicing RNAs and mobile elements which are found in eubacteria, archea and the organelles of fungi, plants, and algae (Bonen and Vogel, 2001; Lambowitz and Zimmerly, 2004; Hausner *et al*, 2006) Because of its encoding function, the *trnK* intron differs from typical group II introns (Hausner *et al*, 2006).

The *matK* ORF has been used as an indicator to construct plant phylogenies because the ORF evolves rapidly, yet is ubiquitous in plants in evolutionary studies (Hilu and Liang, 1997; Kelchner, 2002; Hausner *et al*, 2006).

1.2.3 Significance of the *matK* gene

There are various studies where *matK* gene sequence is used in phylogenetic analysis so far. These studies include family, genera and species levels. In the study concerning the family Saxifragaceae, it has been denoted that the gene *matK* evolves approximately three times faster than *rbcL* (RuBisCo Large subunit) (Johnson and Soltis, 1994, 1995; Johnson *et al*., 1996), the most common cpDNA gene used in phylogenetic analysis (Chase *et al*., 1993). In addition, the *matK* gene sequences have been used in Polemoniaceae (Steele and Vilgalys, 1994), Orchidaceae tribe Vandeae (Jarrell and Clegg, 1995), Myrtaceae (Gadek *et al*, 1996), Poaceae (Liang and Hilu 1996), Apiaceae (Plunkett *et al*. 1996), and flowering plants in general (Hilu and Liang, 1997). The *matK* was shown to have higher variation than any other studied chloroplast genes. However, the variation was slightly higher at the 5' region than at the 3' region, approximate even distribution was observed throughout the entire gene. In addition, the high proportion of transversion (a change from purine to a pyrimidine, or vice versa, is a transversion) in the *matK* gene might provide high phylogenetic information. These factors underscore the usefulness of the *matK* gene in systematic studies and suggest that comparative sequencing of *mat*K may be appropriate for phylogenetic reconstruction at subfamily, family, genera and species levels (Tanaka *et al*, 1997). Furthermore, *matK* has evolutionary patterns and pace that separate it from most genes used in angiosperm phylogeny rearrangement (Olmstead and Palmer, 1994; Hilu and Liang, 1997; Hilu *et al*, 2003).

CHAPTER II

JUSTIFICATION

1. Turkish sweet gum is an economically important species and has a potential of desired hereditary features. There is very little genetic knowledge or ongoing research on this species.

2. According to fossil evidence, although the species spread into the northern part of Anatolia in the past, its natural distribution is now limited to a small area in southwestern Turkey.

3. It is a relict-endemic species with restricted natural distribution. Natural distribution area is reduced from day to day due to anthrophogenic factors.

Besides above reasons, Turkish sweet gum constitutes a crucially important forest ecosystem and it is necessary to develop an effective conservation strategy. The genetic structure of Turkish sweet gum populations urgently needs to be investigated for conservation purposes. Some practices: seed stands, nature conservation areas and clonal seed orchards were the conservation measures to be taken so far. For species with limited genetic information, it is often assumed that genetic variation follows geographic and ecological variation. (Alan and Kaya, 2003). This study can help to eloborate such assumptions. For this reason, determination of genetical variation of Turkish sweet gum at the species and population levels are of great importance.

CHAPTER III

OBJECTIVES

The main objectives of this study are;

i. To obtain genetic data which could help to solve taxonomic problems of Turkish sweet gum at the variety, species and genus levels by means of *matK* gene region,
ii. To determine genetic diversity patterns among natural populations of the species,
iii. To develop *insitu* conservation strategies of genetic resources of the species with the aid of current data.
iv. To explore the genetic and evolutionary relationships of *L. orientalis* with other three species of *Liquidambar* which are natives of America and China.

Furthermore, to look for answers what are the rate and pattern of nucleotide variations in the *matK* gene in Turkish sweet gum and to imply evolutionary consequences of these variations.

CHAPTER IV

MATERIALS AND METHODS

4.1 Plant material

In this study, 18 different populations of Turkish sweet gum from different part of the southwest of Turkey were collected in cooperation with the Forest Trees and Seeds Breeding Research Directorate, Ministry of Environment and Forestry. These populations represent the whole natural range of the species in Turkey (Figure 4.1). Population locations, type of population and altitude of population were provided in detail in Table 4.1. Sample size (number of trees that leaf tissues obtained) was approximately 30-35 trees for each population.

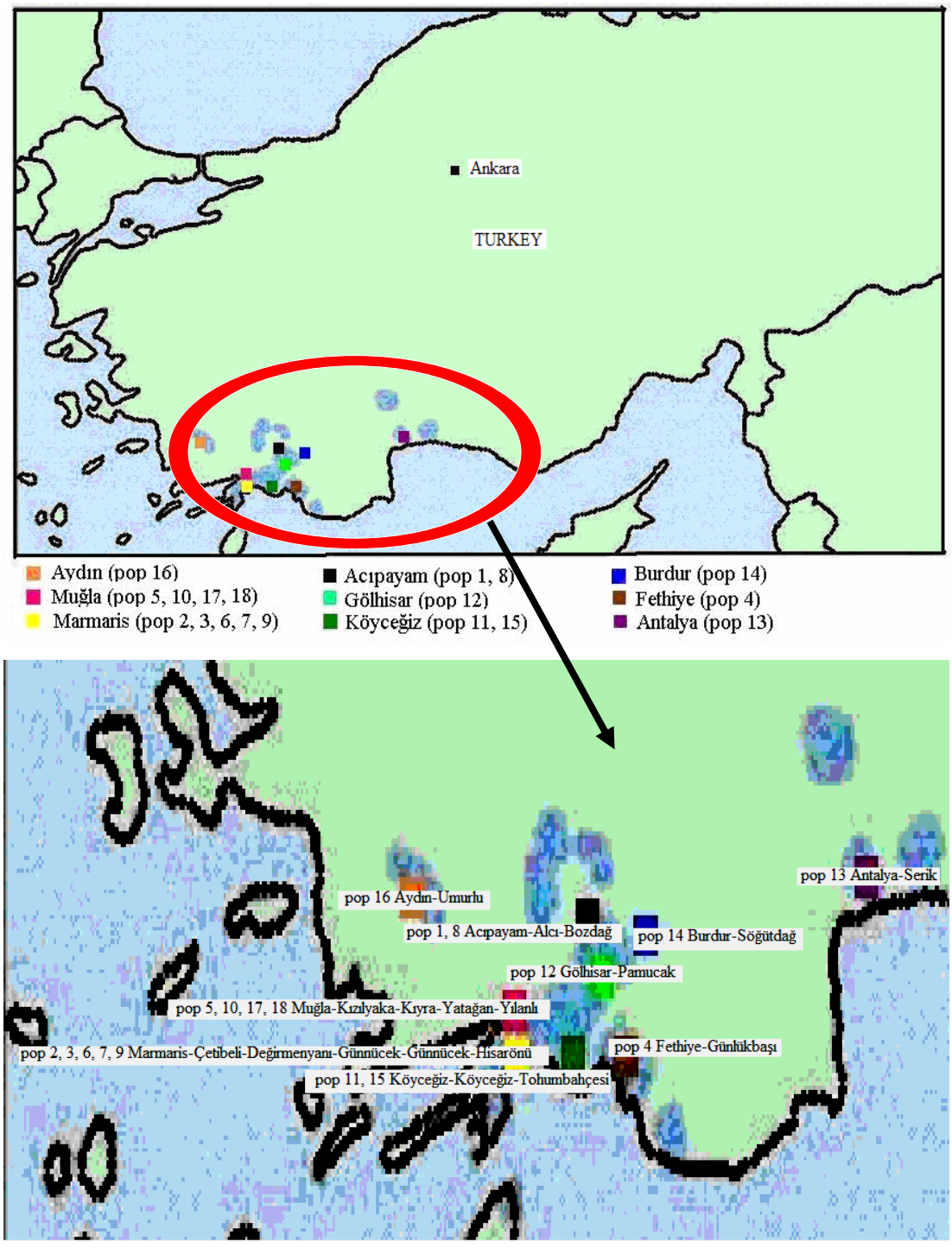


Figure 4.1 Natural range of Turkish sweet gum and location of the 18 populations sampled for the study (Map adopted from Alan and Kaya, 2003)

Table 4.1. Descriptive information on studied Turkish sweet gum populations

#	Populations Name and Locations	Abbreviation of the populations	The varieties of *L. orientalis &* (geograpgic groups)	Stand type	Altitude (m)
1	Acıpayam-Alcı	AK	unknown (1)	Pure Population	1100
2	Marmaris-Çetibeli	ÇE	var. *integriloba* (2)	Vegatation mixed with in stream	30
3	Marmaris-Değirmenyanı	DE	var. *integriloba* (2)	Pure Population	5
4	Fethiye-Günlükbaşı	FE	var. *integriloba* (2)	Pure Population	5
5	Muğla-Kızılyaka	FI	var. *integriloba* (2)	Pure Population	50
6	Marmaris-Günnücek	GC	var. *integriloba* (2)	Pure Population	5
7	Marmaris-Günnücek	GN	var. *integriloba* (2)	Pure Population	5
8	Acıpayam-Bozdağ	GÜ	unknown (1)	Pure Population	1100
9	Marmaris-Hisarönü	HÖ	var. *integriloba* (2)	Vegatation mixed with in stream	10
10	Muğla-Kıyra	KI	var. *integriloba* (2)	Scattered trees	50
11	Köyceğiz-Köyceğiz	KÖ	var. *integriloba* (2)	Pure Population	10
12	Gölhisar-Pamucak	PA	unknown (3)	Near stream	250
13	Antalya-Serik	SE	unknown (3)	Near stream	30
14	Burdur-Söğütdağ	SÖ	unknown (3)	Pure Population	550
15	Köyceğiz-Köyceğiz(TB)	TB	var. *integriloba* (2)	Clonal	10
16	Aydın-Umurlu	UM	var. *orientalis* (2)	Vegatation mixed with in stream	250
17	Muğla-Yatağan	YA	var. *orientalis* (4)	Vegatation mixed with in stream	250
18	Muğla-Yılanlı	YL	var. *orientalis* (4)	Vegatation mixed with in stream	250

(Geographic Groups→Group 1: Denizli, Group 2: Muğla-1, Group 3: Antalya and Burdur, Group 4: Muğla-2)

4.2 DNA isolation

Turkish sweet gum leaves, collected in the field by Forest Trees and Seeds Breeding Research Directorate, Ministry of Environment and Forestry in 2002, were stored at -80°C until DNA extraction. Leaf tissues were obtained from 30 trees in each population and best DNA yielding trees for each population were determined. Total cellular DNA was isolated using modified 2XCTAB (Cetyl Trimethyl Ammonium Bromide) method (Doyle and Doyle, 1987). Before selection of CTAB method, different methods were put into practice. These methods and their outcomes were given in Table 4.2. Figure 4.2 shows the DNA yields of these methods. Two different detergent were applied as CTAB and SDS (Sodium dodecyl sulfate), CTAB (Lane 1-13 and 18) was better than SDS (Lane 14, 15, 16 and 17). Selected leaf weights were 15mg, 25mg, 50mg and 75mg from which 25mg leaf weight was chosen since whole DNA yields with 25mg were satisfactory. Cell debris precipitation was carried out. For protein separation, Choloroform:octanol (24:1 v/v) solution was used and DNA precipitation was carried out with ice-cold isopropanol.

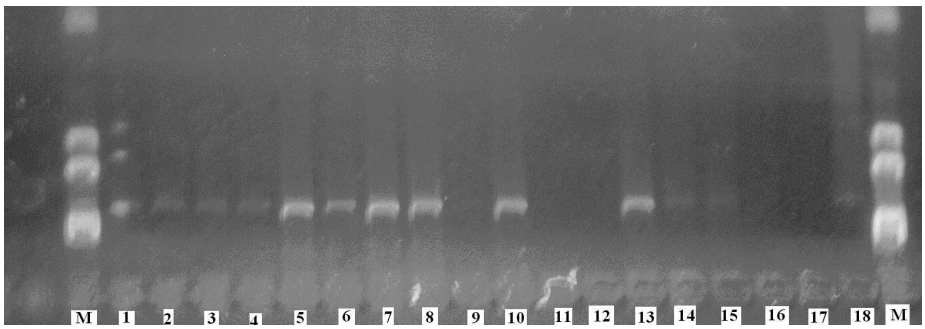


Figure 4.2 DNA yields of the 18 different methods. M is the lamda *Hind*III DNA size marker in the gel photograph. Numbers stand for DNA extraction methods. Please see Table 4.2 for the descriptions of numerical codes

Table 4.2. The protocols applied in DNA isolation studies

Lane	Method	Leaf Weight	Cell debris precipitation	Protein separation	DNA precipitation	DNA Yield
1	CTAB	15 mg	+	C: IAA*	isopropanol	good
2	CTAB	15 mg	-	C: IAA*	isopropanol	good
3	CTAB	15 mg	+	C: O**	isopropanol	good
4	CTAB	15 mg	-	C: O**	isopropanol	good
5	CTAB	25 mg	+	C: IAA*	isopropanol	very good
6	CTAB	25 mg	-	C: IAA*	isopropanol	very good
7	CTAB	25 mg	+	C: O**	isopropanol	very good
8	CTAB	25 mg	-	C: O**	isopropanol	very good
9	CTAB	50 mg	+	C: IAA*	isopropanol	pure
10	CTAB	50 mg	-	C: IAA*	isopropanol	very good
11	CTAB	50 mg	+	C: O**	isopropanol	pure
12	CTAB	50 mg	-	C: O**	isopropanol	smear
13	CTAB	50 mg	+	C: O**	NaAc: EtOH	very good
14	SDS	25 mg	+	C: O**	NaAc: EtOH	good+smear
15	SDS	50 mg	+	C: O**	NaAc: EtOH	good+smear
16	SDS	75 mg	+	C: O**	NaAc: EtOH	pure
17	SDS	50 mg	+	C: O**	NaAc: EtOH	pure
18	CTAB	-	-	C: IAA*	isopropanol	fair

Note: *Chloroform Octanol. **Chloroform-isoamyl alcohol

DNA extraction from frozen leaf tissues using CTAB (Cetyl Trimethyl Ammonium Bromide, Chemical composition is in Appendix A) method was carried out for all 18 populations. First, 25mg leaves were crushed with by the help of liquid nitrogen and put it in the ependorf tubes. Then, the method indicated in Table 4.3 was performed and repeated for each of 12 leaf samples (12 trees). A total of 216 trees (leaf tissues of 12 trees x 18 populations = 216) were used for DNA extraction.

Table 4.3. CTAB protocol

1. 900 µl CTAB [1 M pH 8.0 Tris HCl, 50 ml+0.25 M EDTA (Ethylene Diamine Tetra Acetic Acid) pH 8.0, 40 mL+41g NaCl+1 mL βME (Beta Mercaptoethanol) complete to 500 mL], vortex
2. Incubate for about 1 hour at 65°C in water bath.
3. Spin 14000 rpm, 10 min.
4. Transfer the supernatant to clean microfuge tubes.
5. Add 500 µl Chloroform: Octanol (24:1 v/v). Invert the tubes.
6. Spin 14000 rpm, 15 min.
7. Transfer the supernatant to clean microfuge tubes.
8. Add 500 µl cold isopropanol.
9. Place the tubes –80°C at least for 1 hour.
10. Spin 14000 rpm, 10 min.
11. Remove the supernatant, wash pellet 500 µl, 70 % EtOH (Ethanol) twice.
12. Dry the tubes.
13. Dissolve pellet with 25 µl of PCR grade H_2O.
14. DNA samples were stored at -20°C

4.3 DNA quantification

Concentration DNA of 12 individuals of each 18 populations were determined with Hoefer DyNA QuantTM 200 Fluorometer using the fluorometric assay of Cesarone *et al.* (1979). Instrument was zeroed using 2 mL of assay solution (Appendix A). Then, the instrument was calibrated to 100ng/μL with DNA standard solution (Appendix A). 2 μL of DNA sample was placed with 2 mL of assay solution into the cuvette then the DNA measurment was carried out. After DNA quantification, it was determined that which DNA samples could be used for Polymerase Chain Reaction (PCR) experiments. The average, minimum and maximum DNA yields were given in Table 4.4.

Table 4.4. Average, minimum and maximum DNA concentrations of the studied 18 populations

			DNA Concentration ng/mL	
L. orientalis var. integriloba	**Sample size**	**Mean±SD***	**Minimum**	**Maximum**
Marmaris-Çetibeli (ÇE)	12	3.30±2.3	0	6
Marmaris-Değirmenyanı (DE)	12	3.50±2.8	0	9
Fethiye-Günlükbaşı (FE)	12	9.50±5.0	4	20
Muğla-Kızılyaka (FI)	12	7.50±6.5	1	25
Marmaris-Günnücek (GC)	12	8.75±4.4	2	14
Marmaris-Günnücek (GN)	12	7.60±4.8	1	15
Marmaris-Hisarönü (HÖ)	12	4.25±2.5	1	8
Muğla-Kıyra (KI)	12	3.17±3.2	1	10
Köyceğiz- Köyceğiz (KÖ)	12	7.75±5.6	1	21
Mean for var. *integrilba*	**120**	**5.97±2.47**	**3.17**	**9.50**
L. orientalis var. orientalis	**Sample size**	**Mean±SD***	**Minimum**	**Maximum**
Aydın-Umurlu (UM)	12	2.25±1.9	0	6
Muğla-Yatağan (YA)	12	6.25±4.3	2	17
Muğla-Yılanlı (YL)	12	5.75±5.0	0	15
Mean for var. *orientalis*	**36**	**4.75±2.17**	**2.25**	**6.25**
Unknown	**Sample size**	**Mean±SD***	**Minimum**	**Maximum**
Gölhisar-Pamucak (PA)	12	7.91±5.3	2	19
Antalya-Serik (SE)	12	4.50±2.8	0	11
Burdur-Söğütdağ (SÖ)	12	4.92±4.3	0	14
Acıpayam-Bozdağ (GÜ)	12	9.20±6.9	0	21
Acıpayam-Alcı (AK)	12	12.00±4.8	3	23
Mean for 'unknown'	**60**	**7.71±3.11**	**4.50**	**12.00**

*SD=Standard deviation

4.4 Primer designs for *matK* region

The *matK* gene is 1512bp in length in sweet gum and is embedded within a 2.5 kb group II intron (*trnK* introns, shaded areas in Figure 4.3) that interrupts the two *trn*K exons (white part in Figure 4.3) (Sugiura, 1992).

In this study, 9 primers whose nucleotide compositions ranging from 18 to 21 nucleotides were selected based on a previous study on *Liquidambar* species (Li *et al*., 1997). Additionally, 7 primers with 20 nucleotides each were designed according to the GeneBank data on *Liquidambar orientalis*. For sequencing studies, it was decided to choose 4 of 16 primers. These 4 primers and their sequences were provided in Table 4.5. The sequences of remaining primers can be found in Appendix B.

Table 4.5. The list of the primers used for the sequencing

Primer	Primer sequence (5'-3')	Length (base pairs)	Region to be amplified
*mat*KF1	ACT GTA TCG CAC TAT GTA TCA	21	5' site of *mat*K (yellow + partial green part in Figure 4.1)
*mat*KR3	GAT CCG CTG TGA TAA TGA GA	20	*mat*K (green + partial yellow part in Figure 4.1)
*mat*KF5	TGG AGY CCT TCT TGA GCG*	18	*mat*K (green + partial blue part in Figure 4.1)
*mat*KR1	GAA CTA GTC GGA TGG AGT AG	20	3' site of *mat*K (blue + partial green part in Figure 4.1)

*This primer was synthesised with equal parts of 'C' and 'T' at base position 6.

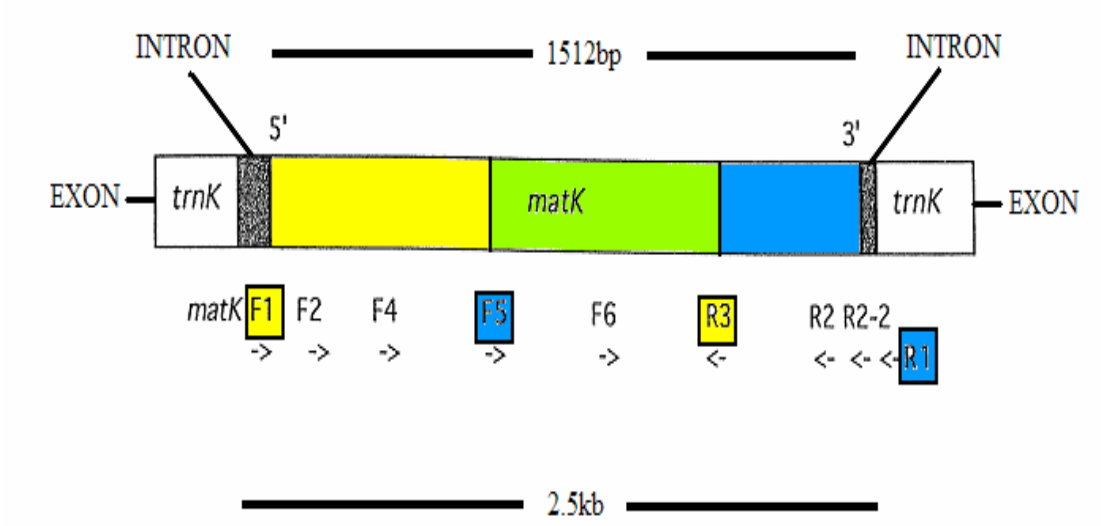


Figure 4.3 Conjectural relative positions of *matK* primers (shaded areas are introns). Green part is amplified by both *matKF1*, *matKR3 and matKF5, matKR1* (Li *et al*., 1997)

4.5 Optimization of PCR (Polymerase Chain Reaction) conditions for Turkish sweet gum

Although there are many studies related with *matK* regions of different species as well as one study related with the *Liquidambar* species, nevertheless, new PCR conditions needed to be tested to optimize the condition for the current study.

4.6 Optimization of reaction conditions

For 25µL of reaction volume, $MgCl_2$, dNTP (deoxyribonucleotide triphosphate) mixture, primers, and template DNA were selected and tested. The combinations of these parameters in different concentrations were provided in Table 4.6.

Table 4.6. The reaction mixtures tested for the optimization of PCR conditions

#	dH_2O (µL)	10X Buffer	$MgCl_2$ (25mM)	dNTP (mM)	Primer Pairs (100µM)	Taq Pol. (5u/µL)	DNA (3ng/µL)	Total (µL)
1	17.9	2.5µL	2.5µL	0.25µL	0.25µL+ 0.25µL	0.1µL	1.25µL	25
2	17.8	2.5µL	2.5µL	0.25µL	0.3µL+ 0.3µL	0.1µL	1.25µL	25
3	17.7	2.5µL	2.5µL	0.25µL	0.35µL+ 0.35µL	0.1µL	1.25µL	25
4	17.6	2.5µL	2.5µL	0.25µL	0.4µL+ 0.4µL	0.1µL	1.25µL	25
5	9.025	1.25µL	0.75µL	0.125µL	0.125µL+ 0.125µL	0.1µL	1µL	12.5
6	**9**	**1.25µL**	**1µL**	**0.125µL**	**0.125µL+ 0.125µL**	**0.1µL**	**1µL**	**12.5**
7	8.075	1.25µL	1.25µL	0.125µL	0.125µL+ 0.125µL	0.1µL	1µL	12.5
8	8.05	1.25µL	1.25µL	0.15µL	0.125µL+ 0.125µL	0.1µL	1µL	12.5

PCR reaction mixtures were tested with two different amounts as 12.5µL and 25µL. The sixth reaction mixture gave a better band among all tested ones. For the sequence analysis 55µL mixture was needed and sixth reaction mixture was modified to 55 µL. Among the total PCR reaction mixture, 5 µL of 10X Buffer ($MgCl_2$ free), 4µL of 25mM $MgCl_2$, 0.5µL of 10mM dNTPs, 2µL of ~3ng/µL of DNA, 0.5µL of 100µM each of the primers and 0.4µL of 5u/µl *Taq* polymerase and 42.1µL of dH_20 (Sterile water) was found to be the best and used for the 55µL of PCR mixture for sequence analysis (Table 4.7). About 5µL of the PCR mixture were run in to agorose gel to visualize the band quality. After detecting the good band, remainig 50µL mixture were stocked for sequencing analysis. The reaction mixtures were prepared in thin-walled 0.2 mL Eppendorf tubes and run on a thermocycler (Eppendorf-Mastercycler, Eppendorf, Canada, and Techne-genius Thermocycler, Techne, USA).

Table 4.7. The composition of optimized PCR reaction mixture

Compenent	Quantity used (µl)	Final concentration
10x Buffer	5	1x
dNTPs (10mM)	0.5	0.1mM
MgCl2 (25mM)	4	2.3mM
Primer (100µM)	0.5	1µM
Tag DNA polymerase (5u/ µl)	0.4	2.5unit
DNA (3ng/ µl)	2	6ng
dH_2O (Sterile Water)	42.1	
Total reaction mixture	55	

4.7 PCR reaction cycles and visualization of PCR product

After testing various PCR cycles, the PCR steps described in Table 4.7 were selected for the amplification of the *matK* region.

Table 4.8. The PCR cycles optimized for amplification of the *matK* region

Temperature (◦C)	Time	Cycle #	Description
94◦C	5 min.	1	Initial denaturation
94 ◦C	30 sec.	30	Denaturation
56◦C	30 sec.		Annealing
72◦C	45 sec.		Extension
72◦C	5 min.	1	Final Extension

PCR products were visualised in 1.7% agarose gels. Gels were run in 1XTAE (0.4M Tris Acetate) buffer at 90-99 volts for 1 hour. After electrophoresis, DNA bands were stained with 5µg/ml ethidium bromide and visualized under UV light. The gels were also photographed and digitialized by using a gel imaging system (Vilbor Lourmat, France).

4.8 Data collection

After amplification of the *matK* region, PCR products were stored at -20°C until sequence analysis. A PCR purification process should be performed before the sequence analysis. Thus, both purification and sequencing studies were done by Refgen Biotechnology, METU Teknokent, Ankara. In sequence analysis, ABI 310 Genetic Analyser User's Manuel was followed and sequencing was performed using the Big Dye Cycle Sequencing Kit (applied biosystems) with a ABI 310 Genetic Analyser (PE applied Biosystem) automatic sequencer. The *matK* gene region was amplified as two parts with the help of 4 primers. These parts were aligned visually before the analysis. For viewing the chromotogram data, Finch TV Version 1.4.0 developed by the Geopiza Research Team was utilized (Patterson *et al.*, 2004-2006).

4.9 Analysis of sequence data of the *matK* region

The aligned and proof read sequence data were grouped into three main data sets (Table 4.9) and three sub-data sets (Table 4.10) for the analysis. Three main data sets were created; the 5' region of the *matK* gene (*matK* F1-R3), 3' region of the *matK* gene (*matK* F5-R1) and the whole sequence of the *matK* gene that contained all individuals which were sequenced. Three sub-data sets includes one representative individual from each population (representative MEGA data can be seen in Appendix D) and other three data sets were generated containing 2, 3 and 4 individuals from each populations (Appendix C). These data sets were also rearranged according to their respective varieties of Turkish sweet gum and

geographic locations of studied populations. The geographic locations and the varieties that populations belong were provided in the (Table 4.1).

Table 4.9. The codes of individuals (genotypes) for each of 18 populations with *matK* sequence of Turkish sweet gum (main data sets)

	POPULATIONS	***matK* F1-R3 region**	***matK* F5-R1 region**	**Entire *matK* region**
1	Acıpayam-Alcı	3, 10*	3, 8, 9, 10*	3, 10*
2	Marmaris-Çetibeli	8, 9, 12	8, 9, 12	8, 9, 12
3	Marmaris-Değirmenyanı	1, 4, 6, 8	1, 4, 6, 8	1, 4, 6, 8
4	Fethiye-Günlükbaşı	1, 2, 4	1, 2, 4, 5	1, 2, 4
5	Muğla-Kızılyaka	7, 8	1, 3, 4, 8	7, 8
6	Marmaris-Günnücek	1, 7	1, 7	1, 7
7	Marmaris-Günnücek	1, 5, 6, 7	1, 2, 5, 15	1, 5
8	Acıpayam-Bozdağ	1, 3, 4	1, 2, 3, 4	1, 3, 4
9	Marmaris-Hisarönü	1, 4, 5, 9	1, 4, 5, 9	1, 4, 5, 9
10	Muğla-Kıyra	1, 5, 6, 13	1, 5, 6, 13	1, 5, 6, 13
11	Köyceğiz-Köyceğiz	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4
12	Gölhisar-Pamucak	1, 4	1, 4	1, 4
13	Antalya-Serik	1, 4, 12	1, 4, 12	1, 4, 12
14	Burdur-Söğütdağ	8, 11	8, 11	8, 11
15	Köyceğiz-Köyceğiz	4, 6, 7, 9	4, 6, 7, 9	4, 6, 7, 9
16	Aydın-Umurlu	22, 33	22, 33	22, 33
17	Muğla-Yatağan	3, 6, 7, 12	3, 6, 7, 12	3, 6, 7, 12
18	Muğla-Yılanlı	1, 2, 3, 5	1, 3, 5, 7	1, 3, 5

*These numbers stand for genotype codes within each population.

Table 4.10. The codes of genotypes with available sequences for each of 18 populations of Turkish sweet gum (sub-data sets)

	POPULATIONS	*matK* F1-R3 region	*matK* F5-R1 region	Entire *matK* region
1	Acıpayam-Alcı	3*	3*	10*
2	Marmaris-Çetibeli	8	9	12
3	Marmaris-Değirmenyanı	6	6	6
4	Fethiye-Günlükbaşı	1	1	1
5	Muğla-Kızılyaka	8	8	7
6	Marmaris-Günnücek	7	1	7
7	Marmaris-Günnücek	1	1	5
8	Acıpayam-Bozdağ	1	1	3
9	Marmaris-Hisarönü	5	5	5
10	Muğla-Kıyra	5	1	6
11	Köyceğiz-Köyceğiz	1	3	4
12	Gölhisar-Pamucak	1	1	1
13	Antalya-Serik	1	1	4
14	Burdur-Söğütdağ	8	8	11
15	Köyceğiz-Köyceğiz	7	7	7
16	Aydın-Umurlu	22	22	22
17	Muğla-Yatağan	3	3	3
18	Muğla-Yılanlı	3	1	5

*These numbers stands for genotype codes for each population

The DNA sequences were aligned vissually. Indels, insertion/deletion points in a sequence (in sequence alignments these are often referred to as "gaps") were not included in the analysis. The data sets of DNA sequences were collected and organized in MEGA format so that it could be analyzed with MEGA (Molecular Evolutionary Genetics Analysis) 3.1 software (Kumar *et al.*, 2004) and it could be used for construction of input data for the analysis by Arlequin software (version 2.000 for population genetics data analysis) (Schneider *et al*., 2000). The sequence statistics, containing nucleotide frequencies, transition/transversion (tr/tv) ratio and variability in different regions of the sequences were calculated by MEGA program (Kumar *et al*., 2004).

4.10 Molecular diversity and phylogenetic analysis based on sequence data of *matK* region

From the NCBI (National Center for Biotechnology Information) site, 10 sequences of the entire *matK* gene were selected and included to current data sets to compare with *matK* sequences of the 18 populations of the *L. orientalis*. These 10 accessions were:

1. *Liquidambar orientalis*-Genbank accession number is **AF015651** (Li *et al.*, 1997).
2. *Liquidambar orientalis*-Genbank accession number is **AF304519** (Shi *et al.*, 2001).
3. *Liquidambar orientalis*-Genbank accession number is **AF133220** (Shi *et al.*, 2001).
4. *Liquidambar acalycina*-Genbank accession number is **AF133222** (Shi *et al.*, 2001).
5. *Liquidambar acalycina*-Genbank accession number is **AF015649** (Li *et al.*, 1997).
6. *Liquidambar formosana*-Genbank accession number is **AF133221** (Shi *et al.*, 2001).
7. *Liquidambar formosana*-Genbank accession number is **AF015650** (Li *et al.*, 1997).
8. *Liquidambar styraciflua*-Genbank accession number is **AF133219** (Shi *et al.*, 2001).
9. *Liquidambar styraciflua*-Genbank accession number is **AF133218** (Shi *et al.*, 2001).
10. *Liquidambar styraciflua*-Genbank accession number is **AF015652** (Shi *et al.*, 2001).

For the sequence analysis, the Arlequin software (Schneider *et al*., 2000) and MEGA 3.1 software (Kumar *et al.*, 2004) statistics programs were used, and the following parameters were estimated: The component of molecular variance by molecular diversity indices, Analysis of Molecular Variance Approach Analysis (AMOVA), pairwise comparison of F_{st} between populations, pairwise differences according to p-distance method, the average distances between populations, bootstrap test of phylogeny, minimum spanning tree were carried out with the data including *L. orientalis* data from this study and 10 accessions from other species obtained from NCBI database. Finally, construction of phylogenetic trees for *L. orientalis* populations from Turkey alone and *Liquidambar* genus was carried out by using neighbour-joining method (Saitou and Nei, 1987).

4.10.1 Population genetic structure inferred by Analysis of Molecular Variance (AMOVA)

The differentiation between varieties of Turkish sweet gum, geographic regions of Turkish sweet gum and the genetic structure of Turkish sweet gum populations were investigated by an analysis of variance framework, as initially defined by Cockerham (1969, 1973), and extended by others (e.g. Weir and Cockerham, 1984). This is the Analysis of Molecular Variance Approach and carried out by Arlequin Software (AMOVA, Excoffier *et al*., 1992). It is essentially similar to other approaches based on analysis of variance of the gene frequencies, but it takes into account the number of mutations between molecular haplotypes, which first needed to be evaluated.

Formally, in the haploid case, it is assumed that the *i*-th haplotype frequency vector from the *j*-th population in the *k*-th group (variety in our case) is linear equation of the form as follows:

$$X_{ijk} = x + a_k + b_{jk} + c_{ijk} \qquad \text{(Equation 1)}$$

The vector **x** is the unknown expectation of $\mathbf{x}_{ijk}$, averaged over the whole study. The effects are **a** for group (variety), **b** for the population within a group (variety), assumed to be additive , random, independent, and to have the associated covariance components, σ_a^2, and σ_b^2, σ_c^2 respectively. The total molecular variance (σ^2) is the sum of the covariance component due to differences among haplotypes within a population

(σ_c^2), the covariance components due to the differences among haplotypes in different populations within a group (variety), (σ_b^2), and the covariance components due to the differences among the **G** groups (variety) (σ_a^2). The same framework could be extended to additional hierarchical levels, such as to accomodate, for instance, the covariance component due to differences between haplotypes within diploid individuals.

In terms of inbreeding coefficients and coalescent times, this F_{st} can be expressed as

$$F_{ST} = \frac{f_0 - f_1}{1 - f_1} = \frac{\bar{t}_1 - \bar{t}_0}{\bar{t}_1}$$ (Equation 2)

Where f_o is the probability of identify by descent of two different genes drawn from the same population, f_1 is the probability of identity by descent of two genes drawn from two different populations, $\bar{t}_1$ is the mean coalescence time of two genes drawn from the same population. The significance of the fixation indices is tested using a non-parametric permutation approach described in Excoeffier *et al.* (1992), consisting of permuting in haplotypes, individuals or populations, among individuals, populations or groups of populations. After each permutation round, all statistics were recomputed to get their null distribution. Depending on the tested statistic and

the given hierarchical design, different types of permutations are formed. Under this procedure, the normality assumption usual in analysis of variance tests is no longer

necessary, nor is it necessary to assume equality of variance among populations or groups of populations. A large number of permutations (1000 or more) was carried out to obtain some accuracy on the final probability.

All estimations were performed using Arlequin Software (version 2000) (Schneider *et al*., 2000). The AMOVA design and expected mean squares were given in Table 4.11, Table 4.12 and Table 4.13.

Table 4.11. Expected AMOVA table for testing variety effect in Turkish sweet gum

Source of variation	Degrees of freedom	Sum of Squares	Expected Mean Squares
Among varieites	2 (G-1)	SSD(AV)	$n''\sigma_a^2+n'\sigma_b^2+\sigma_c^2$
Among populations within varieites	15 (P-G)	SSD(AP/WV)	$n\sigma_b^2+\sigma_c^2$
Within populations	35 (N-P)	SSD(WP)	σ_c^2
Total	52 (N-1)	SSD(T)	σ_T^2

SSD(T) :Total Sum of Squared Deviations
SSD(AV) :Sum of Squared Deviations Among Varieties of Populations
SSD(WP) :Sum of Squared Deviations Within Populations
SSD(AP/WV):Sum of Squared Deviations Among Populations, Within Varieties
G :Number of Varieties in the Structure
P :Total Number of Populations
N :Total Number of Sequences Involved in the Analysis

Table 4.12. Expected AMOVA table for testing the effect of geographic regions of Turkish sweet gum

Source of variation	Degrees of freedom	Sum of squares	Expected Mean Squares
Among regions	3 (G-1)	SSD(AP)	$n''\sigma_a^2+n'\sigma_b^2+\sigma_c^2$
Among populations within regions	14 (P-G)	SSD(AP/WG)	$n\sigma_b^2+\sigma_c^2$
Within populations	35 (N-P)	SSD(WP)	σ_c^2
Total	52 (N-1)	SSD(T)	σ_T^2

SSD(T) :Total Sum of Squared Deviations

SSD(AP) :Sum of Squared Deviations Among Geographic Locations of Populations

SSD(WP) :Sum of Squared Deviations Within Populations

SSD(AP/WG):Sum of Squared Deviations Among Populations, Within Geographic Region

G :Number of Geographic Regions in the Structure

P :Total Number of Populations

N :Total Number of Sequences Involved in the Analysis

Table 4.13. Expected AMOVA table for testing the effects of Turkish sweet gum populations

Source of variation	Degrees of freedom	Sum of squares	Expected Mean Squares
Among populations	17 (P-1)	SSD(AP)	$n\sigma_a^2 + \sigma_b^2$
Within populations	35 (2N-P)	SSD(WP)	σ_b^2
Total	52 (2N-1)	SSD(T)	σ_T^2

SSD(T) :Total Sum of Squared Deviations
SSD(AP) :Sum of Squared Deviations Among Populations
SSD(WP):Sum of Squared Deviations Within Populations
P :Total Number of Populations
N :Total Number of Sequences Involved in the Analysis

4.10.2 Construction of phylogenetic trees for Turkish sweet gum population

Phylogenetic trees show the evolutionary interrelationships among various species or other entities that are believed to have a common ancestor. Phylogenetic relationships of genes or organisms usually are presented in a treelike form with a root, which is called a rooted tree (A rooted phylogenetic tree is a directed tree with a unique node). It also is possible to draw a tree without a root, which is called an unrooted tree (An unrooted trees illustrate the relatedness of the leaf nodes). The branching pattern of a tree is called a topology.

All phylogenetic trees in this study were constructed using neighbour-joining (NJ) method together with bootstrap test analysis. In the case of the NJ method (Saitou and Nei, 1987), the S (smallest value of the sum of all branches) value is not computed for all or many topologies, but the examination of different topologies is embedded in the algorithm, so that only one final tree is produced.

The bootstrap test was applied in this study. The bootstrap test, in which the reliability of a given branch pattern is ascertained by examining the frequency of its occurrence in a large number of trees, each based on the resampled dataset. The bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered "correct". If the value is greater than 50, the topology is considered as informative (Nei and Kumar, 2000). Three phylogenetic trees, those including the analysis in species level, Turkish *L. orientalis* populations and a general phylogenetic tree were constructed by MEGA 3.1 (Figure 3.2, 3.3, 3.4).

4.10.3 Models for estimating genetic distances of Turkish sweet gum

The evolutionary distance between a pair of sequences usually is measured by the number of nucleotide substitutions occurring between them. Evolutionary distances are fundamental for the study of molecular evolution and are useful for phylogenetic reconstructions and the estimation of divergence times. There are some methods for distance estimation for nucleotide sequences. Further details of these methods and general guidelines for the use of these methods are given by Nei and Kumar (2000).

In addition to the distance estimates, also the standard errors of the estimates were computed using the analytical formulas and the bootstrap method. In nucleotide method, sequences were compared nucleotide-by-nucleotide. p-distance model were chosen in this study. This distance is the proportion (p) of nucleotide sites at which two sequences being compared are different. It is obtained by dividing the number of nucleotide differences by the total number of nucleotides compared. It does not make

any correction for multiple substitutions at the same site, substitution rate biases (for example, differences in the transitional (Transition: A transition occurs when a purine is substituted by a purine, or a pyrimidine by a pyrimidine) and transversional rates (Transversion: A change from a purine to a pyrimidine, or vice versa), or differences in evolutionary rates among sites (Nei and Kumar, 2000).

4.10.4 Estimation of pairwise genetic distances (F_{st}) among populations and constraction of phylogenetic trees

Estimation of pairwise genetic distances among populations, the pairwise F_{st}'s may be used as genetic distances, with the application of a slight transformation to linearize the distances with the population divergence time (Reynolds *et al.*, 1983; Slatkin, 1995). The pairwise F_{st} values were calculated and given in the form of a matrix. The null distribution of pairwise F_{st} values under the hypothesis of no difference among the populations is obtained by permuting haplotypes between populations.

4.10.5 Minumum Spanning Tree among haplotypes of Turkish sweet gum

Minimum Spanning Tree (MST) was carried out according to Kruskal (1956) and Prim, 1957) between Operational Taxonomic Units (OTUs) with Arlequin Software. Computation of the MST from the matrix of pairwise distances calculated between all pairs of haplotypes using a modification of the algorithm described in Rohlf (1973). The Minimum Spanning Network embedding all MSTs were computed (Excoeffier and Smouse, 1994).

CHAPTER V

RESULTS

5.1 Molecular diversity in the *matK* region

In the sequence analysis, the *matK* gene begins with the start codon ATG and finishes with the stop codon TGA were divided into two sectors for MEGA 3.1 Analysis by the use of *matK*F1-R3 and *matK*F5-R1 primers. With the alignment of these sectors, the sequence of entire gene region was achieved. The first 5' region was about 1123-1124 base pairs (bp), the second sector 3' region was 754-756 bp in length and the whole *matK* region ranged from 1530 to 1531bp because of the indels. Also, geographic groups were constructed. Among the 1530 bp, 35.1% GC content, 46 variable sites (V), 1450 conserved sites, 28 parsimony-informative sites and 18 singleton sites were observed (Table 5.1). There were also 1484 identical pairs (ii), 1 transitionsal pairs (si), 3 transversional pairs (sv). The *matK*F1-R3 was 1125bp long with a GC content of 33.6%, 1058 conserved sites, 39 variable sites, 16 singleton sites, 23 parsimony informative sites, 1086 identical pairs, 1 transitionsal pairs and 2 transversional pairs. On the other hand, *matK*F5-R1 region was 726bp long with a GC content of 37.3%, 717 conserved sites, 27 variable sites, 11 singleton sites, 16 parsimony informative sites, 740 identical pairs, 1 transitional pairs and 2 transvertional pairs (Table 5.1).

Table 5.1. Estimated molecular diversity parameters for *matK*F1-R3, *matK*F5-R1 and entire *matK* gene region

	***matK*F1-R3**	***matK*F5-R1**	**entire *matK***
Total Length (bp)	1125	726	1530
GC content(%)	33.6	37.3	35.1
Conserved sites	1058	717	1450
Variable sites	39	27	46
Singleton sites	16	11	18
Parsimony informative sites	23	16	28
Identical pairs	1086	740	1484
Transitional pairs	1	1	1
Transversional pairs	2	2	3
No. of sequences used	67	72	64

5.1.1 Molecular diversity within Turkish sweet gum populations

As indicated in Table 5.2, the total length of all samples was 1530, usable site of the samples ranged from 1444 to 1496, polymorphic sites varied between ‘0’ and ‘8’. Transitions changed between ‘0’ and ‘4’, also transversition altered from ‘0’ to ‘7’. Moreover, substitutions differed from ‘0’ to ‘8’, indels varied from ‘0’ to ‘13’ and nucleotide diversity (average over total site) ranged from 0.0006 to 0.0060. Among 18 populations, population 4 (Fethiye-Günlükbaşı) was the most diversed one containing 8 polymorphic site. Also, population 2 (Marmaris-Çetibeli) with 4 polymorpic site, population 10 (Muğla-Kıyra) with 7 polymorphic site and population 11 (Köyceğiz-Köyceğiz) with 3 polymorphic site followed this population. Furthermore, population 13 (Antalya-Serik) and population 16 (Aydın-Umurlu) were the least diversed and the most conserved ones having no polymorphic site.

Table 5.2. Estimated molecular diversity parameters for studied Turkish sweet gum populations

Populations	Molecular diversity parameters								
	Gene copies	Total Site	Usable Site	Polym. Site	Transitions	Transversion	Substitutions	Indels	Nucleotide diversity (average over total site)
1-Acıpayam-Alcı	2	1530	1492	1	0	1	1	3	0.0026 (±0.0029)
2-Marmaris-Çetibeli	3	1530	1492	4	1	3	4	1	0.0022 (±0.0019)
3-Marmaris-Değirmenyanı	4	1530	1493	1	1	0	1	8	0.0030 (±0.0022)
4-Fethiye-Günlükbaşı	3	1530	1494	8	1	7	8	3	0.0049 (±0.0039)
5-Muğla-Kızılkaya	2	1530	1493	1	0	1	1	3	0.0026 (±0.0029)
6-Marmaris-Günnücek	2	1530	1444	1	0	1	1	3	0.0027 (±0.0030)
7-Marmaris-Günnücek	2	1530	1491	2	0	2	2	0	0.0013 (±0.0016)
8-Acıpayam-Bozdağ	3	1530	1492	2	0	2	2	2	0.0017 (±0.0016)
9-Marmaris-Hisarönü	4	1530	1496	1	1	0	1	9	0.0034 (±0.0025)
10-Muğla-Kıyra	4	1530	1492	7	4	3	7	4	0.0037 (±0.0027)
11-Köyceğiz-Köyceğiz	4	1530	1494	3	0	3	3	4	0.0023 (±0.0017)
12-Gölhisar-Pamucak	2	1530	1492	2	1	1	2	7	0.0060 (±0.0063)
13-Antalya-Serik	3	1530	1495	0	0	0	0	5	0.0022 (±0.0019)
14-Burdur-Söğütdağ	2	1530	1492	2	0	2	2	2	0.0026 (±0.0029)
15-Köyceğiz-Köyceğiz	4	1530	1496	1	0	1	1	13	0.0046 (±0.0033)
16-Aydın-Umurlu	2	1530	1492	0	0	0	0	1	0.0006 (±0.0009)
17-Muğla-Yatağan	4	1530	1494	1	0	1	1	7	0.0026 (±0.0020)
18-Muğla-Yılanlı	3	1530	1494	1	0	1	1	3	0.0017 (±0.0016)

5.2 Molecular variances among Turkish sweet gum populations

AMOVA analysis for 18 oriental sweet populations was performed. About 14.38% total molecular variance was among populations while 85.62% total variance was within populations (Table 5.3).

Table 5.3. AMOVA results with respect to 18 Turkish sweet gum populations

Source of variation	d.f	Sum of squares	Variance components	Percentage of variation
Among populations	17	56.596	0.37481 Va	14.38
Within populations	35	78.083	2.23095 Vb	85.62
Total	52	134.679	2.60576	

5.2.1 Genetic distances within Turkish sweet gum populations

The genetic distances among genotypes within Turkish sweet gum populations were computed. The average distance values ranged from '0' for population 13 (Antalya-Serik) and population 16 (Aydın-Umurlu) to '0.0038' for population 4 (Fethiye-Günlükbaşı). Population 2 (0.0019), population 4 (0.0038), population 7 (0.0014), population 10 (0.0012), population 11 (0.0011) and population 12 (0.0014) were with the highest average genetic distance among genotypes within population while other 12 Turkish sweet gum populations were the lowest average genetic distance ones [varied '0' (population 13, 16) and '0.0009' population 8 (Table 5.4)].

Table 5.4. Average genetic distances within populations of Turkish sweet gum

Population number and name	Genetic distance within Turkish populations of Turkish sweet gum (±standard error)
POP1-Acıpayam-Alcı	0.0007 (±0.0007)
POP2-Marmaris-Çetibeli	0.0019 (±0.0009)
POP3-Marmaris-Değirmenyanı	0.0004 (±0.0003)
POP4-Fethiye-Günlükbaşı	0.0038 (±0.0013)
POP5-Muğla-Kızılkaya	0.0007 (±0.0007)
POP6-Marmaris-Günnücek	0.0007 (±0.0007)
POP7-Marmaris-Günnücek	0.0014(±0.0010)
POP8-Acıpayam-Bozdağ	0.0009 (±0.0007)
POP9-Marmaris-Hisarönü	0.0005 (±0.0005)
POP10-Muğla-Kıyra	0.0012 (±0.0007)
POP11-Köyceğiz-Köyceğiz	0.0011 (±0.0006)
POP12-Gölhisar-Pamucak	0.0014 (±0.0010)
POP13-Antalya-Serik	0.0000 (±0.0000)
POP14-Burdur-Söğütdağ	0.0014 (±0.0009)
POP15-Köyceğiz-Köyceğiz	0.0004 (±0.0003)
POP16-Aydın-Umurlu	0.0000 (±0.0000)
POP17-Muğla-Yatağan	0.0004 (±0.0003)
POP18-Muğla-Yılanlı	0.0005 (±0.0005)

5.3 Molecular variances among Turkish sweet gum varieties

AMOVA analysis among varieties (two known and one unknown varieties of Turkish sweet gum) was carried out. One of the groups was composed of 10 populations from var. *integriloba*; the second group was formed with 5 populations

from unknown and the third group was 3 populations from var. *orientalis*. There was no variation among varieties of Turkish sweet gum, but the portion of total molecular variance due to populations within varieties was 15.28%. However, the great portion of total molecular variance (86.10%) was due to individuals within populations (Table 5.5).

Table 5.5. AMOVA results with respect to varieties of Turkish sweet gum

Source of variation	d.f	Sum of squares	Variance components	Percentage of variation
Among varieties	2	5.776	0.0000 Va	0.00
Among populations within varieties	15	50.819	0.39587 Vb	15.28
Within populations	35	78.083	2.23095 Vc	86.10
Total	52	134.679	2.59118	

5.3.1 Genetic distances among Turkish sweet gum populations as varieties

Genetic distances were computed among varieties of Turkish sweet gum populations. Among two varieties and one unknown group, the var. *integriloba* was the most divergent variety (0.0016) and var. *orientalis* was the least divergent one (0.0006). Also, average genetic distance between varieties ranged from 0.0007 to 0.0012. However, these values were too low to consider variety-differentiation and it could be interpreted as there was no variation among varieties of Turkish sweet gum (Table 5.6).

Table 5.6. Average genetic distances computed among populations of varieties of Turkish sweet gum

Average genetic distance among populations with varieties		**Average distance between varieties**		
		unknown	***integriloba***	***orientalis***
unknown	0.0009 (±0.0004)*			
integriloba	0.0016 (±0.0004)*	0.0012 (±0.0004)*		
orientalis	0.0006 (±0.0003)*	0.0007 (±0.0003)*	0.0011 (±0.0003)*	

*Standard errors of the estimates

5.4 Molecular variances among geographic locations of Turkish sweet gum

Turkish sweet gum populations were also evaluated according to the geographic locations of the populations using AMOVA analysis. 18 populations were divided into four geographic regions by considering geographical approximities, natural barriers and watersheds and their distances from the sea. These regions are:

Region 1: Denizli

Population 1 (ACIPAYAM-ALCI)

Population 8 (ACIPAYAM-BOZDAĞ)

Region 2: Muğla-1 (Fethiye-Köyceğiz-Aydın)

Population 2 (MARMARİS-ÇETİBELİ)

Population 3 (MARMARİS-DEĞİRMENYANI)

Population 4 (FETHİYE-GÜNLÜKBAŞI)

Population 5 (MUĞLA-KIZILYAKA)

Population 6 (MARMARİS-GÜNNÜCEK)

Population 7 (MARMARİS-GÜNNÜCEK)

Population 9 (MARMARİS-HİSARÖNÜ)

Population 10 (MUĞLA-KIYRA)

Population 11 (KÖYCEĞİZ-KÖYCEĞİZ)

Population 15 (KÖYCEĞİZ-KÖYCEĞİZ-Seed Orchard)

Population 16 (AYDIN-UMURLU)

Region 3: Antalya-Burdur

Population 12 (GÖLHİSAR-PAMUCAK)

Population 13 (ANTALYA-SERİK)

Population 14 (BURDUR-SÖĞÜTDAĞ)

Region 4: Muğla-2

Population 17 (MUĞLA-YATAĞAN)

Population 18 (MUĞLA-YILANLI)

The percentage of total variation due to geographic regions was found to be zero. Thus, there was no difference between geographic regions of Turkish sweet gum populations. The differences among populations within geographic regions made up 16.39% of the total variation while the great portion of total molecular variance (86.74%) was due to individuals within populations (Table 5.7).

Table 5.7. AMOVA results with respect to geographic regions of Turkish sweet gum

Source of variation	d.f	Sum of squares	Variance components	Percentage of variation
Among regions	3	8.063	0.0000 Va	0.00
Among populations within regions	14	48.533	0.42145 Vb	16.39
Within populations	35	78.083	2.23095 Vc	86.74
Total	52	134.679	2.57192	

5.4.1 Average genetic distances among geographic locations of Turkish sweet gum populations

Genetic distances were also computed among geographic locations of Turkish sweet gum populations. Among four regions, Muğla-1 (0.0015) was the most divergent region including Fethiye-Köyceğiz-Aydın populations, Denizli (0.0010), Antalya-Burdur (0.0008) and Muğla-2 (0.0004) with Yatağan-Yılanlı populations followed it.

Average genetic distance among populations with geographic locations ranged from 0.0006 to 0.0013. The most genetically distant locations were Muğla-1 and Denizli l.

While the least genetic distance was observed between Muğla-2 and Antalya-Burdur (Table 5.8).

Table 5.8. Average genetic distances computed according to the geographic locations of Turkish sweet gum

Average genetic distance for geographic regions		**Average genetic distance between geographic regions**			
		Denizli	**Muğla-1**	**Ant.-Bur.**	**Muğla-2**
Denizli	0.0010 (±0.0006)*				
Muğla-1	0.0015 (±0.0004)*	0.0013 (±0.0004)*			
Ant.-Bur.	0.0008 (±0.0004)*	0.0009 (±0.0004)*	0.0011 (±0.0003)*		
Muğla-2	0.0004 (±0.0003)*	0.0007 (±0.0004)*	0.0010 (±0.0003)*	0.0006 (±0.0003)*	

*Standard errors of the estimates

5.5 Genetic differences of among *Liquidambar* species as well as among Turkish populations of *L. orientalis* based on F_{st} values

In this part, 18 populations and 10 individuals of outgroups were grouped into seven taxonomic groups. 18 populations were grouped as var. *integriloba*, var. *orientalis* and unknown; outgroups were grouped as *L. orientalis-outgroup*, *L. acalycina-outgroup*, *L. formosana-outgroup* and *L. styraciflua-outgroup*. Group *integriloba* included 10 populations, group *orientalis* contained 3 populations and unknown implied 5 populations. *L. orientalis-outgroup* included 3 individuals, *L. acalycina-*

outgroup contained 2 individuals, *L. formosana-outgroup* implied 2 samples and *L. styraciflua-outgroup* comprised 3 samples.

Pairwise F_{st} values among 28 sequences representing *Liquidambar* species were estimated and given in Table 5.9 A and B. The values ranged between 0.000 and 0.744 in Table 5.9 A and 0.563 and 0.823 in Table 5.9 B. If F_{st} is equal to zero, compared populations do not have any difference. F_{st} value *L. orientalis* var. *integriloba* and var. *orientalis* was 0.018; among *L. orientalis* var. *integriloba* and unknown was 0.016 and between *L. orientalis* var. *orientalis* and unknown was 0.040. F_{st} values of Turkish sweet gum populations and *L. orientalis-outgroup* varied between 0.242 and 0.331. Turkish populations of Turkish sweet gum and *L. styraciflua-outgroup* F_{st} values were moderately high and ranged from 0.534 to 0.664. As expected, F_{st} values among Turkish populations of Turkish sweet gum and *L. acalycina-outgroup* (ranged from 0.537 to 0.681) as well as Turkish populations of Turkish sweet gum and *L. formosana*-outgroup (from 0.584 and 0.738) were high, indicating strong differentiation among these species. *L. formosana*-outgroup was the most distant from Turkish populations of *L. orientalis*, while *L. acalycina-outgroup* and *L. styraciflua-outgroup* were closer to Turkish populations of Turkish sweet gum (0.534-0681) than *L. formosana*_outgroup. In Table 5.9 B, the lowest F_{st} values was observed between Turkish populations of *L. orientalis* and *L. styraciflua* and the highest F_{st} values was observed between *L. styraciflua* and *L. formosana*.

Table 5.9.

A) Pairwise comparison of F_{st} values among Turkish sweet gum varieties and *Liquidambar* species

Taxonomic Units	1 LOI	2 LOO	3 LOU	4 LO_ OUT	5 LA_ OUT	6 LF_ OUT	7 LS_ OUT
***L. orientalis* var. *integriloba* (LOI)**	--						
***L. orientalis* var. *orientalis* (LOO)**	0.018	--					
***L. orientalis* var. *unknown* (LOU)**	0.016	0.040	--				
***L. orientalis-outgroup* (LO_OUT)**	0.242	0.314	0.331	--			
***L. acalycina-outgroup* (LA_OUT)**	0.537	0.608	0.681	0.591	--		
***L. formosana*_outgroup (LF_OUT)**	0.584	0.659	0.738	0.677	--	--	
***L. styraciflua-outgroup* (LS_OUT)**	0.534	0.601	0.664	0.213	0.744	0.823	--

B) Pairwise F_{st} values among *Liquidambar* species

Taxonomic Units	1	2	3	4
L. orientalis	--			
L. styraciflua	0.563	--		
L. acalycina	0.570	0.744	--	
L. formosana	0.613	0.823	--	--

5.6 Phylogenetic trees

Phylogenetic trees were formed in three levels: variety level, geographic region level and regardless of variety and geographic region level (only 18 populations and some outgroups). For each level, one representative individual of each 18 populations were utilized as shown in Figure 5.1, Figure 5.2 and Figure 5.3. Other trees which contain two individuals of each population were provided in Appendix E. Five groups were observed in variety level (Figure 5.1), but specific differentiation was not seen as indicated two varieties: var. *integriloba*, var. *orientalis* and also unknown. Although the constructed tree for Turkish sweet gum varieties did not reveal any clear pattern, it appears that those populations labeled as '*integriloba*' and '*orientalis*' varieties were somewhat in the same clusters though this tree is just an informative one.

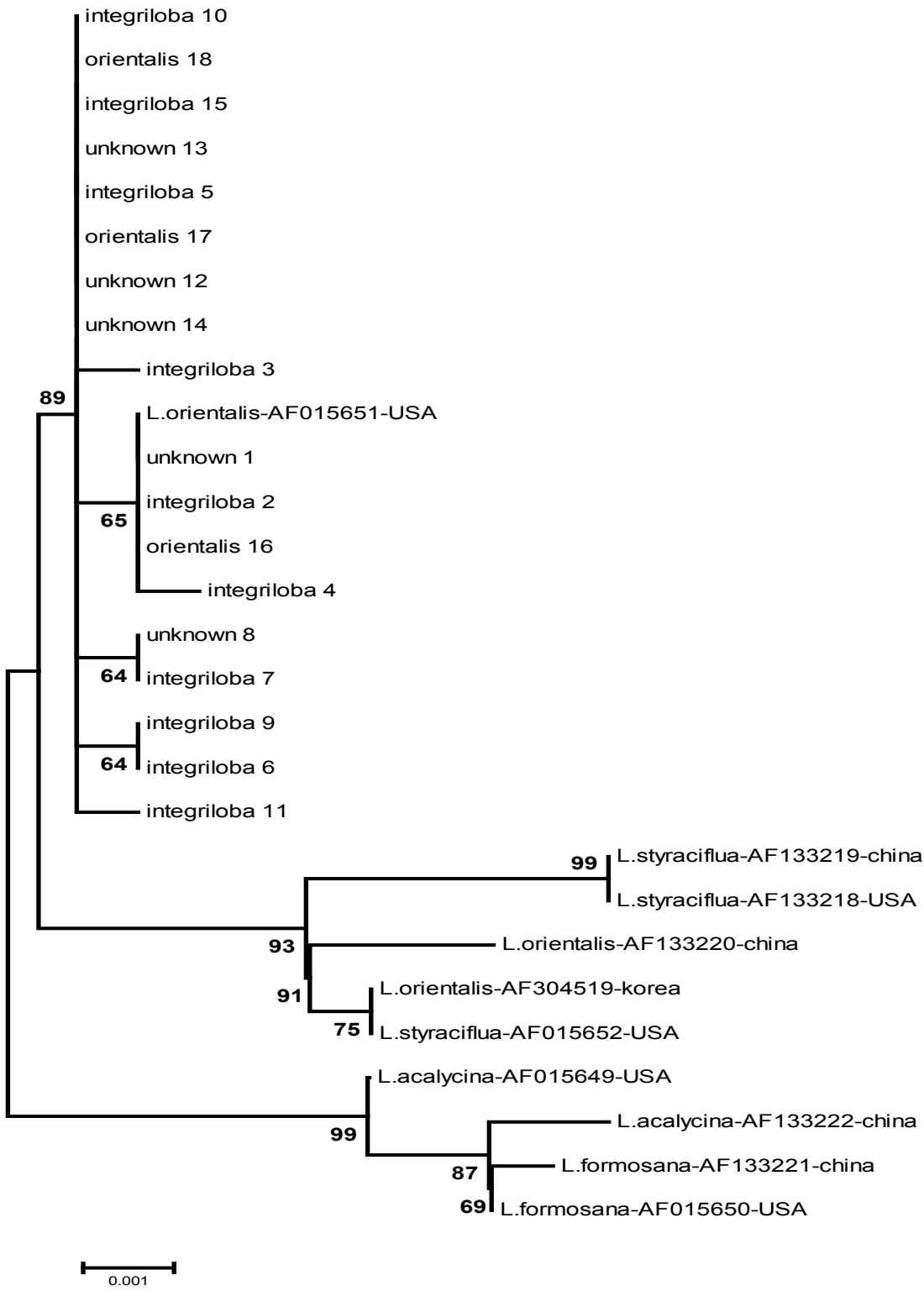


Figure 5.1 The phylogenetic tree regarding varieties of Turkish sweet gum derived from neighbour-joining methods using p-distance

The second tree was constructed with respect to the geographic locations of the Turkish sweet gum populations. As indicated before, 18 populations of Turkish sweet gum were evaluated in four geographic regions. These geographic regions were Muğla-1, Denizli, Antalya-Burdur and Muğla-2. Populations of Muğla-1 region were ended up in the cluster along the populations of Denizli region, the populations of Muğla-2 region were closer to Antalya-Burdur populations (Figure 5.2).

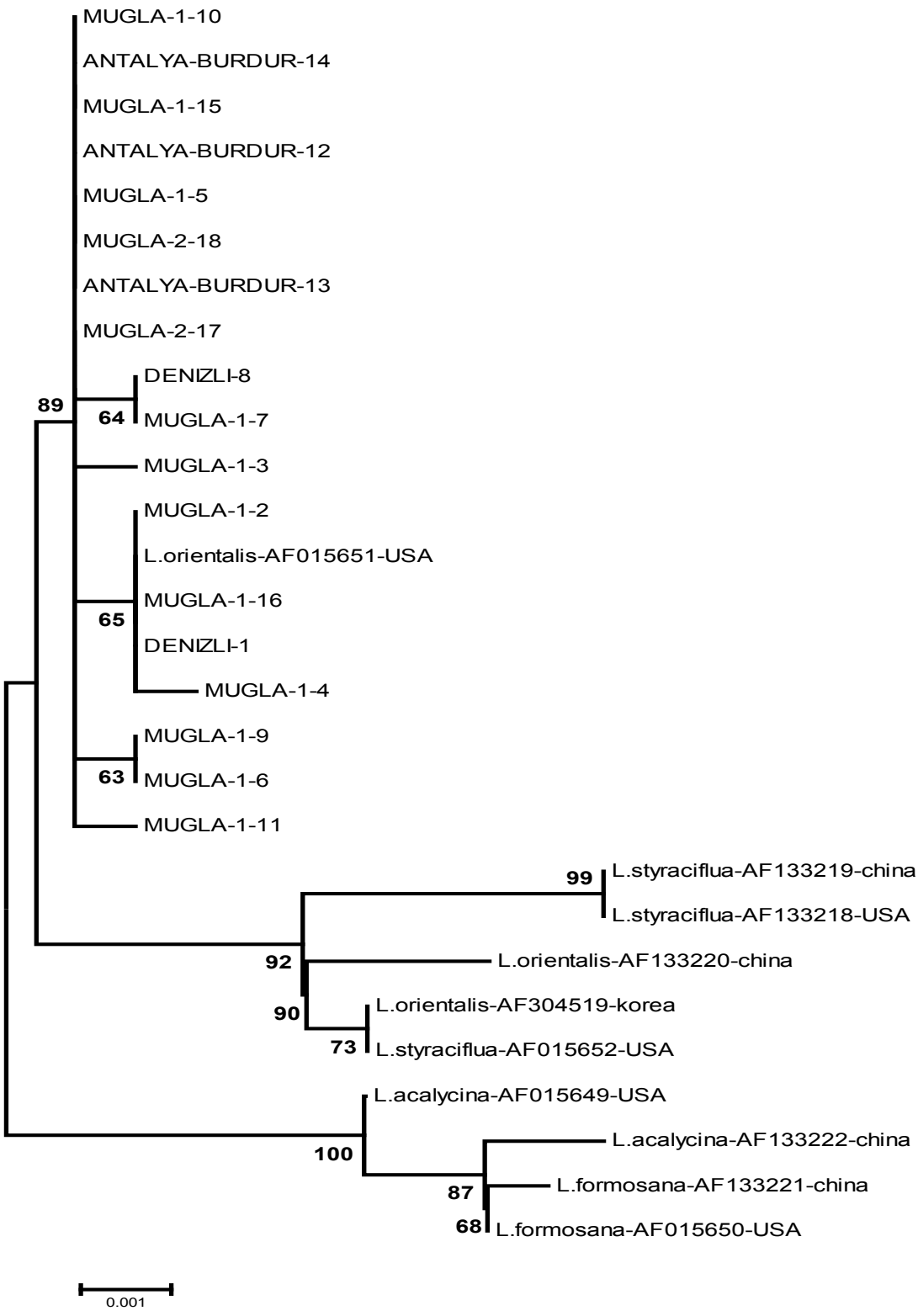


Figure 5.2 The phylogenetic tree regarding the geographic locations of populations derived from the neighbour-joining methods using p-distance

The last phylogenetic tree was formed with respect to 18 Turkish sweet gum populations regardless of variety levels and geographic locations. There were 4 clusters; one of the cluster included 8 populations originated mainly from Muğla-Köyceğiz-Yatağan-Yılanlı-Kızılyaka-Kayra, Antalya-Serik and Burdur-Söğütdağ-Gölhisar. Another major cluster group was with the populations from Muğla-Fethiye-Marmaris, Aydın and Denizli. The remaining 2 other clusters had only a few populations. One of these cluster consisted of the populations from Marmaris and the other did not have any geographic pattern (Figure 5.3).

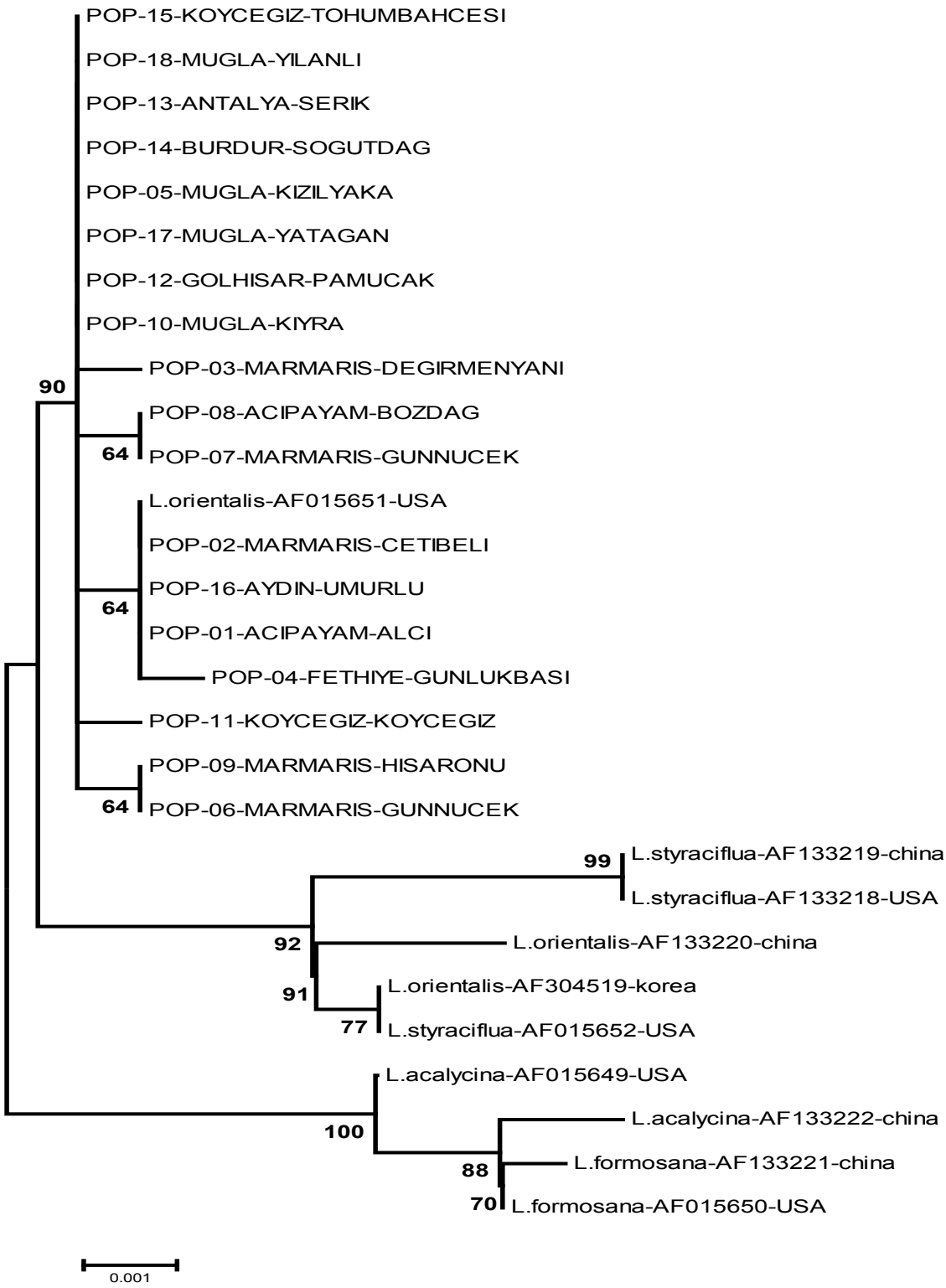


Figure 5.3 The phylogenetic tree regarding 18 Turkish sweet gum populations derived from the neighbour-joining methods using p-distance

5.7 The minimum spaning tree of *L. orientalis* varieties

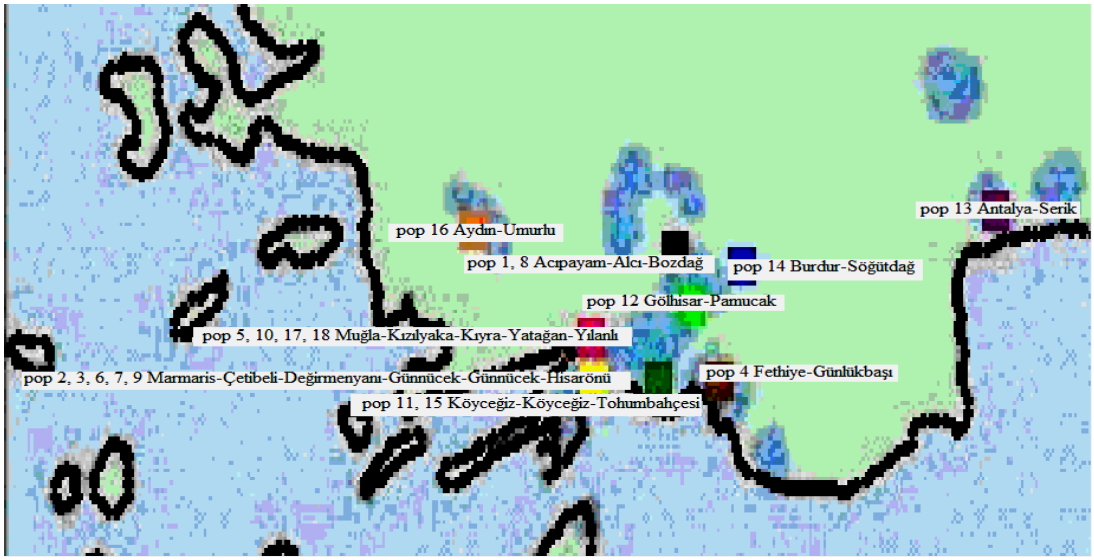
The minimum spaning tree constructed for 28 sequences (18 sequences from Turkish Turkish sweet gum populations of this study and 10 sequences from *Liquidambar* species) were shown in Figure 5.4. The results indicated 3 evolutionary groups:

Group 1 (*L. orientalis* group) involved *L. orientalis*_AF015651 (19) from America and 18 Turkish *L. orientalis* populations (1-18). The most differentiated populations were population 3 (Marmaris-Değirmenyanı), 11 (Köyceğiz-Köyceğiz), 17 (Muğla-Yatağan) and 18 (Muğla-Yılanlı);

Group 2 (*L. acalycina* and *L.formosana Hance* group) consisted of *L. acalycina*_AF133222_China (22), *L. acalycina*_AF015649_USA (23), *L. formosana*_AF133221_China (24) and *L. formosana*_AF015650_USA (25);

Group 3 (*L. styraciflua* group, but some *L. orientalis* samples were clustered with this species) composed of *L. orientalis* (20) from Korea, *L. orientalis* (21) from China, *L. styraciflua*_China (26), _USA (27), _USA (28). The most different group was *L. styraciflua*_China (26) and _USA (27).

A)



B)

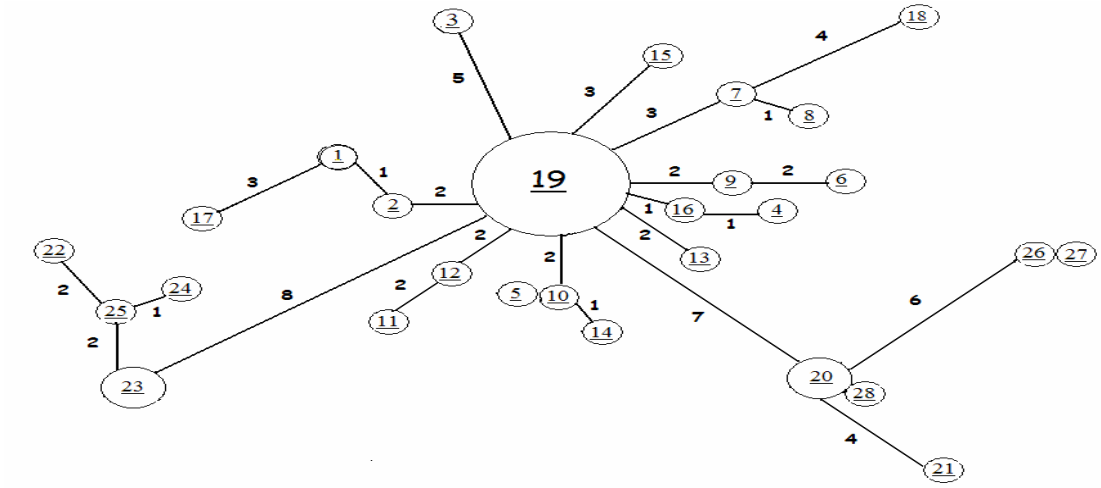


Figure 5.4 A) Map showing the location of populations. Numbers in population names correspond to population-codes in Figure 5.4 B.

B) Minimum spaning tree of 28 operational taxonomical units (OTUs) of *L. orientalis* varieties and 10 outgroups. Number on the branches indicate the base differences between OTUs

Sample 1-18→*L. orientalis* populations, Sample 19→*L. orientalis*_AF015651, Sample 20→*L. orientalis*_AF304519, Sample 21→*L. orientalis*_AF133220, Sample 22→*L. acalycina*_AF133222, Sample 23→*L. acalycina*_AF015649, Sample 24→*L. formosana*_AF133221, Sample 25→*L. formosana*_AF015650, Sample 26→*L. styraciflua*_AF133219, Sample 27→ *L. styraciflua*_AF133218, Sample 28→ *L. styraciflua*_AF015652

CHAPTER VI

DISCUSSION

6.1 Molecular diversity in the *matK* region of Turkish sweet gum populations

In this study, because of indels (insertion and deletion of bases), cpDNA *matK* region of Turkish sweet gum populations was obtained to be 1530bp in length. This is within the range of figures reported by previous studies. For example, *matK* region was approximately 1.5kb, in a study carried out by Hilu and Liang (1997) on *Liquidambar* species. They reported that *matK* region was 1512bp. Young and Pamphilis (2000) also reported the length of *matK* gene of photosynthetic and nonphotosynthetic Orobanchaceae and their relatives as about 1530bp. Because the *matK* is shown higher variation than any other studied chloroplast genes, it could be possible that *matK* region is different in length (because of indels). Furthermore, *matK* gene was also evaluated in two parts. First region (*matK*F1-R3) which is the 5' region of the gene which was found to be 1125bp and second region (*matK*F5-R1) which is the 3' region of the gene, was found to be 726bp.

The entire *matK* gene included 35.1% GC content, 46 variable sites, 1450 conserved sites, 28 parsimony informative sites and 18 singleton sites. In a study carried out by Kusumi *et al.* (2000), they reported that *matK* gene had 33% GC content of the 23 members of families including Taxodiaceae, Cupressaceae, Taxaceae and Cephalotaxaceae. *matK*F1-R3 part of the gene contained 33.6% GC, *matK*F5-R1 part consisted of 37.3% GC content. Furthermore, while first region had 39 variable sites, 1058 conserved sites, 23 parsimony informative sites and 16 singleton sites, second region had 27 variable sites, 717 conserved sites, 16 parsimony informative sites and 11 singleton sites. As a result, 3' region was found more varible than 5' region. However, a study carried out by Hilu and Liang (1997), indicated that the 5'

region of the *matK* gene is more useful at lower taxonomic levels than 3' region. Also, several studies have performed using the *matK* gene sequence at different taxonomic levels: at family level (Johnson and Soltis, 1994, 1995; Johnson *et al.*, 1996; Steele and Vilgalys, 1994; Jarrel and Clegg, 1995; Gadek *et al.*, 1996, Liang and Hilu, 1996; Plunkett *et al.*, 1996; Hilu and Liang, 1997; Wang *et al.*, 1999; Hilu and Alice, 1999; Gadek *et al.*, 2000; Shi *et al.*, 2000; Wang *et al.*, 2000; Kusumi *et al.*, 2000; Song *et al.*, 2001; Ohsako and Ohnishi, 2001; Mort *et al.*, 2001; Cameron *et al.*, 2001; Ge *et al.*, 2002; Tanaka *et al.*, 2003; Sanders *et al.*, 2003; Neel and Cummings, 2004; Hidalgo *et al.*, 2004; Bell, 2004), at genera and species level (Tanaka *et al.*, 1997; Fukuda *et al.*, 2001; Wilson, 2004; Järvinen *et al.*, 2004; Meimberg *et al.*, 2006) and also at variety level as in the case of this study. Instead of higher taxonomic levels, *matK* region may be useful at lower taxonmic levels as such in variety level in this study.

When considering the molecular diversity within 18 Turkish sweet gum populations, population 4 (Fethiye-Günlükbaşı) was the most diversed group containing 8 polymorphic site, while population 13 (Antalya-Serik) and population 16 (Aydın-Umurlu) were the least diverse groups. Today, some Turkish sweet gum populations have been set aside as conservation programs by Forest Trees and Seeds Breeding Research Directorate. One seed orchard (2.2 hectares) in Muğla-Fethiye-Göcek, two seed stands in Muğla-Fethiye-Göcek and Muğla-Marmaris-Çetibeli (200.8 hectares) and two gene conservation forests (277 hectares) found in Isparta-Bucak-Pamucak and Muğla-Ula-Kızılyaka could be given as example of such programs. According to the results of molecular diversity parameters estimated with this study, population 4 (Fethiye-Günlükbaşı) located in Muğla province could be recommended as a additional genetic conservation sites like Muğla-Fethiye-Göcek region.

6.2 Genetic differences among Turkish sweet gum populations at population, variety and geographic location levels of the populations

Considering molecular variances among Turkish sweet gum populations, 14.38% of total variation due to population could be considered as low. Since the portion of total quiete variation due to individuals within populations of Turkish sweet gum with respect to *matK* gene was high (86%). However, for the sequences of the nuclear DNA internal transcriped spacer (ITS), it was also shown that most of the variation was found within populations when compared to populations of *Hamamelidaceae* family (Shi *et al*., 1998).

Although inter-variety differentiation was not detected, the results of genetic difference analysis revealed that the most divergent variety was found to be the var. *integriloba* (0.0016). Although the *matK* gene is one of the most variable plastid genes (Olmstead and Palmer, 1994; Soltis and Soltis 1998), Turkish sweet gum varieties were not differentiated with the polymorphism revealed by this region.

While comparing the pairwise F_{st} values among Turkish sweet gum varieties, the most differentiation was observed between var. *orientalis* and unknown variety-group (0.040). In the revision study of Peşmen (1972) for 'Flora of Turkey', two varieties (var. *orientalis* and var. *integriloba*) with respect to presence or absence of secondary lobes on leaves were described, but Davis and Hedge (1975) emphasized that this subject should be restudied. Furthermore, Efe (1987) did not observe this discrimination during any of her field studies. The results support Efe (1987) that there is no true differentiation among prescribed varieties.

According to the results conducted based on the 4 geographic location differences among Turkish sweet gum populations, there was no significant variation among geographic location where populations originated. However, it was found that the populations of Turkish sweet gum from Muğla-1 was the most divergent ones (0.0015).

6.3 Genetic differences among *Liquidambar* species including Turkish sweet gum populations

In this study, 4 members of *Liquidambar* genus were compared including Turkish sweet gum populations with respect to *matK* gene. Turkish sweet gum populations have a close relationship with *L. styraciflua* while comparing the Chinese species of *Liquidambar* (*L.formosana Hance* and *L. acalycina*). While comparing Turkish populations of Turkish sweet gum with the sequences of samples from *Liquidambar* species studied before, Turkish populations have little bit different genetic structure than *L. orientalis* sampled from China, but *L. orientalis* sampled from America showed more similarity than that of China. The Chinese sample may have exchanges with the Chinese sweet gum species.

Several studies were applied among *Liquidambar* members to identify the relationships of the species. *Liquidambar* genus is a woody taxa that includes morphologically similar individuals on different continents in the world (Hoey and Parks, 1994). Because, this genus represents with mainly four species on three continents (Western Asia, *L. orientalis*, Eastern Asia, *L.formosana Hance* and *L. acalycina*, America, *L. styraciflua*) researchers have been interested in its genetic divergence. In a isozyme divergence study done by Hoey and Parks (1991), *L. orientalis* and *L. styraciflua* appeared the most closely related intercontinental pair of species. In another study by Hoey and Parks (1994), they dealt with the three species of *Liquidambar* namely *L.formosana Hance*, *L. acalycina* and *L. styraciflua*. *L.formosana Hance* and *L. acalycina* exhibited low levels of intraspecific population divergence. According to the genetic divergence study dealing with sequence data of the cpDNA *matK* gene among four species of *Liquidambar*, two clades generated. One clade includes *L.formosana Hance* and *L. acalycina*, while the other was consist of *L. orientalis* and *L. styraciflua* (Li *et al*, 1997, Li and Donughue, 1999).

6.4 The constructed phylogenetic trees as population, variety and geographic location respects

When the constructed phylogenetic tree was examined, it was clear that Turkish sweet gum populations formed three branches with bootstrap values of 63, 65 and 90 meaning that those topologies are just phylogenetically informative. In the main branch, having a bootstrap value of 89-90, *L. orientalis* from USA was grouped together with 18 populations of Turkish sweet gum (Figure 5.1, 5.2, 5.3). This proposes that there were no or very little variation between Turkish and American sweet gum samples when *matK* region is considered. 4 clusters, but mainly 2 were observed among Turkish sweet gum populations. The first main cluster which composed of 8 Turkish sweet gum populations mainly forms populations from Muğla, Burdur and Antalya; the second one includes 4 populations from Muğla, Denizli and Aydın province. However, 2 populations were aparted from these 4 clusters; population 3 (Marmaris-Değirmenyanı) and population 11 (Köyceğiz-Köyceğiz). Also, 2 other clusters including *Liquidambar* genus members were obtained. One cluster composed of 3 *L. styraciflua* sequences and 2 *L. orientalis* which were from China and Korea. These results are also consistet with the results of genetic distances. These samples were considered as hybrids, because Turkish *L.oreintalis* populations, their DNA isolation materials collected from natural distribution of the species, did not observe close relationships with them as Turkish sweet gum from USA.

6.5 MST

The minimum spanning tree constructed by Arlequin 2.000 showed consistent results with the phylogenetic tree constructed by MEGA 3.1. In both of the trees, 3 main clusters were appeared. One cluster consisted of 18 Turkish sweet gum populations and *L. orientalis* from USA; other group formed with *L. orientalis* form China, Korea and *L. styraciflua* from USA, and the last group includes Chinese members of *Liquidambar* species (*L. acalycina* and *L.formosana Hance*).

Because of the occurrence of closely related species in many plant genera in eastern Asia and North America which are two widely distributed areas (Wen, 2001), the close relationship between *L. orientalis* and *L. styraciflua* could be explained with the help of a land bridge in the early Oligocene between North America and Europe.

In addition to the divergent populations stated in the phylogenetic trees, Population 3 and 11, Population 17 and population 18 which are the members of var. *orientalis* and the members of Muğla-2 region could be given priorities in future conservation programs of both *insitu* and *exsitu*.

CHAPTER VII

CONCLUSION

The main purpose of this study was to obtain genetic data that will help to solve taxonomic status of Turkish sweet gum (endemic) at variety, species and genus levels by means of studying *matK* gene region.

Turkish sweet gum *matK* gene was found to be 1530bp in length. The 5' and 3' regions as well as entire *matK* gene were compared and 3' region was found more variable than 5' region. Also, before this study, *matK* gene have not been used at variety level, it was used especially at species and higher taxonomic levels, but this study show that this region may not be given accurate distinction for variety level.

In respect of the results of molecular diversity analysis; the most divergent population was found to be population 4 (Fethiye-Günlükbaşı) of Turkish sweet gum. The population 2 (Marmaris-Çetibeli) and population 10 (Muğla-Kıyra) were the other populations with high diversity values. Among the varieties, the most divergent variety was var. *integriloba* followed by the unknown group and var. *orientalis*. Furthermore, the most divergent populations were from the variety *integriloba*. These populations could be members of true taxonomic entity that is var. *integriloba*.

The most separated populations, that show differences in *matK* sequences when compared to other populations, are population 2 (Marmaris-Çetibeli), population 3 (Marmaris-Değirmenyanı), population 4 (Fethiye-Günlükbaşı), population 7 (Marmaris-Günnücek), population 10 (Muğla-Kıyra) population 11 (Köyceğiz-

Köyceğiz), population 12 (Gölhisar-Pamucak), population 14 (Burdur-Söğütdağ), population 17 (Muğla-Yatağan) and population 18 (Muğla-Yılanlı). All of these populations should be considered as potential candidates for *insitu* gene conservation programs. However, population 4 (Fethiye Günlükbaşı) should be urgently included in *insitu* and *exsitu* conservation programs.

When the geographic distributions are considered, the most divergent region was found to be region 2 (Muğla-1) which includes genetically the most distant populations as expected.

When the results of phylogenetic and minimum spanning trees were consisted, *L. orientalis* from USA together with the Turkish populations of the species and *L. styraciflua* from the USA were genetically the closest neighbors. However, the Chinese representative of *Liquidambar* genus did not show close relationship between Turkish populations of Turkish sweet gum with respect to *matK* gene and they were the most distant species to *L. orientalis*.

REFERENCES

Acatay, A. 1963. Sığla ağacı (*Liquidambar orientalis* Mill.)'nın Türkiye'de yayılışı, yeni tesbit edilen varyetesi ve sığla ağaçlarına musallat olan böcekler. *İ.Ü. Orman Fakültesi Dergisi*, Seri A, Cilt 8, Sayı 2, İstanbul.

Akman, Y. 1995. Türkiye orman vejetasyonu. A.Ü. Fen Fakültesi Botanik Ana Bilim Dalı. Ankara.

Alan M. and Z. Kaya, 2003. EUFORGEN Technical Guidelines for genetic conservation and use for Turkish sweet gum (*Liquidambar orientalis*). International Plant Genetic Resources Institute, Rome, Italy. 6 pages.

Bell, C. D. 2004. Preliminary phylogeny of Valerianaceae (Dipsacales) inferred from nuclear and chloroplast DNA sequence data. *Molecular Phylogenetics and Evolution* 31: 340–350.

Bogle, A. L. 1986. The floral morphology and vascular anatomy of the Hamamelidaceae: subfamily Liquidambaroideae. *Annals of the Missouri Botanic Garden*. 73: 325-347.

Bonen, L. and J. Vogel 2001. The ins and outs of group II introns. *Trends in Genetics* . 6: 322–331.

Cameron, K.M., Chase, M.W., Anderson, W.R. and H.G. Hills 2001. Molecular systematics of Malpighiaceae: evidence from plastid *rbcL* and *matK* sequences. *American Journal of Botany* 88: 1847–1862.

Cesarone, C., Bolognesi, C. and Santi, L. 1979. Improved microfluorometric DNA determination in biological material using 33258 Hoechst. *Analytical Biochemistry*. 100: 188-197.

Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B.D., Duvall, M. R., Price, R. A., Hills, H. G., Qiu, Y. -L., Kron, K. A., Rettig, J. H., Conti, E., Palmer, J. D., Manhart, J. R., Sytsma, K. J., Michaels, H. J., Kress, W. J., Karol, K. G., Clark, W. D., Hedren, M., Gart, B. S., Jansen, R. K., Kim, K. -J., Wimpee, C. P., Smith, J. F., Furnier, G. R., Strauss, S. H., Xiang, Q. -Y., Plunkett, G. M., Soltis, P. S., Swensen, S. M., Williams, S. E., Gradek, P. A., Quinn, C. J., Cockerham, C.C. 1969. Variance of gene frequencies. *Evolution*, 23: 72-83.

Chase, M. W., D. E. Soltis, R. G. Olmstead, D. Morgan, H. D. Les, B. D. Mishler, M. R. Duvall, R. A. Price, H. G. Hills, Y.-L. Qiu, K. A. Kron, J. H. Rettig, E. Conti, J. D. Palmer, J. R. Manhart, K. J. Sytsma, H. J. Michaels, W. J. Kress, K. G. Karol, W. D. Clark, M. Hedrén, B. S. Gaut, R. K. Jansen, K.-J. Kim, C. F. Wimpee, J. F. Smith, G. R. Furnier, S. H. Strauss, Q.-Y. Xiang, G. M. Plunkett, P. S. Soltis, S. M. Swensen, S. E. Williams, P. A. Gadek, C. J. Quinn, L. E. Eguiarte, E. Golenberg, G. H. Learn Jr., S. W. Graham, S. C. H. Barrett, S. Dayanandan, and V. A. Albert. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. *Annals of the Missouri Botanical Garden* 80:526-580.

Cockerham, C.C. 1973. Analysis of gene frequencies. *Genetics*, 74: 679-700.

Davis, P. H. and Hedge, I. C. 1975: The Flora of Turkey. Past, present and future. *Condollea* 30: 331-351.

Dirik, H. 1986. Anadolu Sığlası (*Liquidambar orientalis* Mill.)'nın gençleştirilmesi üzerine çalışmalar. İ.Ü. Orman Fakültesi Yüksek Lisans Tezi, İstanbul.

Doyle J. J. and J. L. Doyle 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull*. 19: 11-15.

Efe, A. 1987. *Liquidambar orientalis* Mill. (Sığla ağacı)'ın morfolojik ve palinolojik özellikleri üzerine araştırmalar. *İ.Ü. Orman Fakültesi Dergisi* Seri A. Cilt: 37, Sayı:2, 273-286, İstanbul.

Excoeffier, L., Smouse, P.E. and Quattro. 1992. Analysis of molecular variance inferred from Metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479-491.

Excoeffier, L., Smouse, P.E. 1994. Using allele frequencies and geographic subdivision to reconstruct gene genealogies within a species. Molecular Variance Parsimony. *Genetics*, 136: 343-359.

Fakir, H. ve Ö. Doğanoğlu 2003. Isparta Sığla (*Liquidambar orientalis* Mill.) ormanı tabiatı koruma alanı bitki taksonları. *Süleyman Demirel Üniversitesi Orman Fakültesi Dergisi*. Seri: A, Sayı: 1, ISSN: 1302-7085, Sayfa: 67-86

Forest Trees and Seeds Breeding Research Directorate, http://www.ortohum.gov.tr, last visited August 2007.

Fukuda, T., J. Yokoyama, and H. Ohashi 2001. Phylogeny and biogeography of the genus *Lycium* (Solanaceae): inferences from chloroplast DNA sequences. *Molecular Phylogenetics Evolution* 19: 246–258.

Gadek, P. A., Wilson P. G. and C. J. Qinn 1996. Phylogenetic reconstruction in Myrtaceae using *mat*K, with particular reference to the position of *Psiloxylon* and *Heteropyxis*. *Australian Systematic Botany*.

Gadek, P.A., D.L. Alpers, M.M. Heslewood and C.J. Quinn 2000. Relationships within Cupressaceae sensu lato: a combined morphological and molecular approach. *American Journal of Botany* 87(7):1044-1057.

Ge, S., Li, A., Lu, B.R., Zhang, S.Z. and D.Y. Hong 2002. A phylogeny of the rice tribe oryzeae (poaceae) based on *matK* sequence data. *American Journal of Botany* 89(12): 1967–1972.

Hausner, G., Olson, R., Simon, D., Johnson, I., Sanders, E. R., Karol, K. G., McCourt, R. M. and S. Zimmerly 2006. Origin and evolution of the choloroplast *trnK* (*matK*) intron: A model for evolution of group II intron RNA structures. *Molecular Biology and Evolution* 23(2): 380-391.

Hidalgo, O., Garnatje, T., Susanna, A. and J. Mathez, 2004. Phylogeny of Valerianaceae based on *matK* and *ITS* markers, with reference to *matK* individual polymorphisms. *Annals of Botany* 93, 283–293.

Hilu, K. W., and H. Liang 1997. The *matK* gene: sequence variation and application in plant systematics. *American Journal of Botany* 84: 830-839

Hilu, K. W., and L. A. Alice. 1999. Evolutionary implications of *matK* indels in Poaceae. *American Journal of Botany* 86: 1735-1741.

Hilu, K. W., Borch T., Müller, K., Soltis, D. E., Soltis, P. S., Savolainen, V., Chase, M. W., Powell, M. P., Alice, L. A., Evans, R., Sauquet, H., Neinhuis, C., Slotta, T. A. B., Rohwer, J. G., Campbell, C. S. and L. W., Chatrou 2003. Angiosperm Phylogeny based on *matK* sequence information. *American Journal of Botany* 90(12): 1758-1776.

Hoey, M.T., Parks, C.R., 1991. Isozyme divergence between eastern Asian, North American, and Turkish species of *Liquidambar* (Hamamelidaceae). *American Journal of Botany* 78, 938–947.

Hoey, M.T., Parks, C.R., 1994. Genetic divergence in *Liquidambar styraciflua L.*, *L. formosana*, and *L. acalycina* (Hamamelidaceae). *Systematic Botany* 19, 308–316.

Huş, S. 1949. Sığla ağacının *Liquidambar orientalis* Mill. ormancılık bakımından önemi ve sığla yağının kimyasal araştırılması. Orman Genel Müdürlüğü Yayını Özel Sayı: 83. 7-61.

Ichert-Bond, S., Pigg K., and J. Wen 2005. Comparative infructescence morphology in *Liquidambar* (Altingiaceae) and its evolutionary significance. *American Journal of Botany*. 92(8): 1234–1255.

İktüeren, Ş. ve İ. Acar 1987. Sığla ağacının (*Liquidambar orientalis* Mill.) doğal yayılışı, sığla yağı üretimi ve pazarlaması. *Ormancılık Araştırma Enstitüsü Yayınları Dergi Serisi*, Cilt:33, Sayı:2, No:66.

Jarrel, D. C. and M. T. Clegg 1995. Systematic implications of the chloroplast-encoded *mat*K gene on the tribe Vandeae (Orchidaceae). *American Journal of Botany* 82: 137.

Järvinen, P., Palme, A., Morales, L.O., Lännenpää, M., Keinänen, M., Sopanen, T. and M. Lascoux 2004. Phylogenetic relationshıps of *Betula* species (Betulaceae) based on nuclear *adh* and chloroplast *matK* sequences. *American Journal of Botany* 91(11): 1834–1845. 2004.

Johnson L. A., and D. E. Soltis. 1994. *mat*K DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic Botany* 19: 143-156.

Johnson L. A., and D. E. Soltis 1995. Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *mat*K sequences. *Annals of the Missouri Botanic Garden* 82: 149-175.

Johnson L. A., J. L. Schultz, D. E. Soltis and P. S. Soltis 1996. Monophyly and generic relationships of Polemoniaceae based on *mat*K sequences. *American Journal of Botany* 83: 1207-1224.

Kelchner, S. A. 2002. Group II introns as phylogenetic tools: Structure, function, and evolutionary constraints. *American Journal of Botany* 89(10): 1651–1669.

Kruskal, J.R. 1956. On the shortest spanning subtree of a graph and the travelling salesman problem. *Proceedings of the American Mathematical Society* 7: 48-50.

Kumar, S., Tamura, K. and M. Nei 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Briefings in Bioinformatics* 5:150-163.

Kusumi, J., Y. Tsumura, H. Yoshimaru, and H. Tachida. 2000. Phylogenetic relationships in Taxodiaceae and Cupressaceae sensu stricto based on matK gene, chlL gene, trnL-trnF IGS region, and trnL intron sequences. *American Journal of Botany* 87: 1480–1488.

Lambowitz, A. M. and S. Zimmerly 2004. Mobile group II introns. *Annual Review of Genetics* 38: 1-35.

Li J.-H., Bogle A. L. and A. S. Klein 1997. Interspecific relationships and genetic divergence of the disjunct genus *Liquidambar* (Hamamelidaceae). *Rhodora* 99: 229-241.

Li, J. and M. J. Donoghue 1999. More molecular evidence for interspecific relationships of Liquidambar (Hamamelidaceae). *Rhodora* 101: 37-41.

Liang, H. and K. W. Hilu 1996. Application of the *mat*K gene sequences to grass systematics. *Canadian Journal of Botany* 74: 125-134.

Liang, H. 1997. The Phylogenetic Reconstruction of the Grass Family (Poaceae) Using *mat*K Gene Sequences. The PhD thesis of the Faculty of the Virginia Polytechnic Institute and State University. Blacksburg, Virginia.

Meimberg, H., Thalhammer, S., Brachmann, A. 2006. Comparative analysis of a translocated copy of the *trnK* intron in carnivorous family Nepenthanceae. *Molecular Phylogenetics and Evolution*, 39:478–490.

Mort, M. E., D. E. Soltis, P. S. Soltis, J. Francisco-Ortega, and A. Santos-Guerra. 2001. Phylogenetic relationships and evolution of Crassulaceae inferred from matK sequence data. *American Journal of Botany* 88:76-91.

Neel, M.C. and Cummings, M.P. 2004. Section-level relationships of North American *Agalinis* (Orobanchaceae) based on DNA sequence analysis of three chloroplast gene regions. *BMC Evolutionary Biology* 4: http://www.biomedcentral.com/1471-2148/4/15

Nei, M. and Kumar, S. 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Neuhaus H. and G. Link 1987. The chloroplast tRNA lys (UUU) gene from mustard *(Sinapis alba)* contains a class II intron potentially coding for a maturase-related polypeptide. *Current Genetics* 11: 251-257.

Ohsako, T., and Ohnishi, O. 2001. Nucleotide sequence variation of the chloroplast *trnK*/*matK* region in two wild *Fagopyrum* (Polygonaceae) species, *F. leptopodum* and *F. statice*. *Genes and Genetic Systems* 76: 39–46.

Olmstead, R. G., and J. D. Palmer, 1994. Choloroplast DNA systematics: a review of methods and data analysis. *American Journal of Botany* 81: 1205-1224.

Önal, S. ve S. Özer 1985. Ülkemizdeki sığla yağı üretimi ve değerlendirilmesindeki sorunlar. Orman Ürünleri Endüstri Kongresi (ORENKO). Trabzon.

Örtel, E. 1988. Sığla ormanlarımızın durumu. *Ormancılık Araştırma Enstitüsü Dergisi Cilt* 17, sayı 194:16-19.

Patterson, J., Chamberlain, B. and D. Thayer 2004-2006. Finch TV Version 1.4.0

Peşmen, H. 1972. The Genus *Liquidambar* L. in Davis, Flora of Turkey. Vol:4, 264-265, Edinburgh.

Plunkett, G. M., Soltis, D. E. and P. S. Soltis 1996. Evolutionary pattern in Apiaceae: inferences based on *mat*K sequence data. *Systematic Botany* 21: 477-495.

Prim, R.C. 1957. Shortest connection networks and some generalizations. *Bell System Technical Journal*.36: 1389-1401.

Reynolds, J., Weir, B.S, and Cockerham, C.C. 1983. Estimation for the coancestry coefficient: basis for a short term genetic distance. *Genetics* 105: 767-779.

Rohlf, F.J. 1973. Algorithm 76. Hierarchial clustering using the minimum spaning tree. *The computer Journal*. 16:93-95.

Saitou, N. and Nei., M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

Sanders, E.R., K.G. Karol and R.M. McCourt 2003. Occurrence of *matK* in a *trnK* group II intron in charophyte green algae and phylogeny of the Characeae. *American Journal of Botany* 90: 628-633.

Schneider, S., Roessli, D., and Excoeffier, L. 2000. ARLEQUIN ver 2.000. A software for population genetics data analysis. Department of Anthropology and Ecology. University of Geneva, Switzerland.

Shi, S., Chang H. T., Chen, Y., Qu L., Wen, J., 1998. Phylogeny of the Hamamelidaceae based on the ITS sequences of nuclear ribosomal DNA. *Biochemical Systematics and Ecology* 26, 55–69.

Shi, S., Jin, H., Zhong, Y., He, X., Huang, Y. and F. Tan 2000. Phylogenetic relationships of the Magnoliaceae inferred from cpDNA *matK* sequences. *Theoretical and Applied Genetics* 101:925-930.

Shi, S., Huang, Y., Zhong, Y., Du, Y., Zhang, Q., Chang, H. and D.E. Boufford 2001. Phylogeny of the Altingiaceae based on cpDNA *matK* PY-IGS and nrDNA ITS sequences. *Plant Systematics and Evolution* 230: 13-24.

Slatkin, M. 1995. A measure of population subdivision based on micro-satellite allele frequencies. *Genetics*: 139: 457-462.

Soltis, D.E., and Soltis, P.S. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. In: Soltis DE, Soltis PM, Doyle JJ (eds) Molecular systematics of plants. II. DNA sequencing. Kluwer Academic Publishers, Boston, pp 1–42.

Song B. H., Wang X. Q., Li F. Z., Hong D.Y. 2001. Further evidence for paraphyly of the Celtidaceae from the chloroplast gene *matK*. *Plant Systematics and Evolution*. 228: 107–115.

Steele, K. P., and Vilgalys, R. 1994. Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the Plastid gene *mat*K. *Systematic Botany*. 19:126-142.

Sugiura M., 1992. The chloroplast genome. *Plant Molecular Biology*19: 149–168

Tanaka, N., Setoguchi H. and J. Murata 1997. Phylogeny of the family Hydrocharitaceae inferred from *rbcL* and *matK* gene sequence data. *Journal of Plant Research* 110:329-337.

Tanaka, N., Kuo, J., Omori, Y., Nakaoka, M. and K. Aioi 2003. Phylogenetic relationships in the genera Zostera and Heterozostera (Zosteraceae) based on *matK* sequence data. *Journal of Plant Research.* 116, 273–279.

Wang, X.R., Tsumura, Y., Yoshimaru, H., Nagasaka, K. and A.E. Szmidt 1999. Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast rbcL, *MATK, RPL20-RPS18* spacer, and TRNV intron sequences. *American Journal of Botany* 86: 1742-1753.

Wang, X.Q., Tank, D.C. and T. Sang 2000. Phylogeny and Divergence Times in Pinaceae: Evidence from Three Genomes. *Molecular Biology* and *Evolution* 17:773-781.

Wen, J., 2001. Evolution of eastern Asian-eastern North American biogeographic disjunctions: a few additional issues. *International Journal of Plant Sciences*. 162, S117–S122.

Weir, B.S. and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38: 1358-1370.

Wilson, C.A. 2004. Phylogeny of *Iris* based on chloroplast *matK* gene and *trnK* intron sequence data. *Molecular Phylogenetics and Evolution*, (Vol. 33) (No. 2) 402-412.

Wolfe K. H., Morden C. W. And J. D. Palmer 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proceedings of the. National. Academy of. Sciences of the USA* 89: 10648-10652.

Young, N. D. and C. W. De Pamphilis 2000. Purifying selection detected in the plastid gene *matK* and flanking ribozyme regions within a group II intron on nonphotosynthetic plants. *Molecular Biology and Evolution* 17: 1933–1941.

APPENDIX A: BUFFERS, CHEMICALS and EQUIPMENTS

Buffers and solutions for DNA extraction and quantification

DNA Extraction

2X CTAB: 10 g CTAB (Cetyl Trimethyl Ammonium Bromide), (SIGMA)
50 mL (pH: 8.0) Tris HCl, (SIGMA)
40 mL (pH: 8.0) 0.5M EDTA, (FLUKA)
41 g 5M NaCl is completed with 500 mL with dH_2O

Chloroform-Octanol, (FLUKA)**:** (24:1)

β-Mercaptoethanol, (SIGMA)**:** 17.5 ml β-Mercaptoethanol is completed with 250 mL dH_2O

Isopropanol, (FLUKA)**:** Pure isopropanol, ice cold

Ethanol: 70% in dH2O

TE buffer: 10 mM Tris HCl (pH: 7)
10 mM Ethylenediaminetetraaceticacid disodium salt (EDTA)

DNA Quantification

Assay Solution (Low Range:10-500ng/mL final concentration**)**
0.1µg/mL H 33258 in 1XTNE (0.2M NaCl, 10mM Tris-Cl, 1mM pH 7.4)

H 33258 stock solution	10µL
10XTNE buffer	10mL
Distilled filtered water	90mL

DNA Standard for Low Range

1:10 dilution (100µg/mL) of 1mg/mL DNA standard stock solution. Mix:

1mg/mL DNA standard stock	100µL

10XTNE buffer	100µL
Distilled water	800µL

Hoechst dye stock solution

(10mL, 1mg/mL Hoechst H 33258)

Add 10mL distilled water to 10mg H 33258. Do not filter. Store at 4°C for up to 6 months in an amber bottle.

10XTNE buffer (1000mL, buffer stock solution)

12.11g	Tris	100mM
3.72g	EDTA $Na_2.2H_2O$	10mM
116.89g	NaCl	2M

Dissolve in 800mL distilled water. Adjust pH to 7.4 with concentrated HCl. Add distilled water to 1000mL. Filter before use (0.45µm). Store at 4°C for up to 3 months.

Buffers and solutions for Polymerase Chain Reaction (PCR)

10X PCR Buffer ($MgCl_2$ free) (BIORON)

$MgCl_2$ Stock Solution (BIORON)**:** 25mM $MgCl_2$

dNTPs (LAROVA)**:** 5mM

Taq DNA polymerase (BIORON)**:** 5U/µL

Sterile Water: dH_2O

Primer-pairs: 100µM

Electrophoresis buffers and gel systems

Agarose Gel Electrophoresis

Running Buffers: 1XTAE prepared in dH_2O

Agarose, (SIGMA)**:** 1 or 1.7 percent (w/v) Agarose gel

Ethyidium Bromide, (SIGMA)**:** 5mg/mL

Loading Buffer: 9.5mL Formamide, (SIGMA)
500µL EDTA (0.5 M)
15mg Bromophenolblue, (SIGMA)
15mg Xylene cyanol, (SIGMA)

Equipments

Autoclave: Kermanlar – ISTANBUL
Centrifuge: Sigma 113
Deepfreezer: Sanyo – Medical Freezer
Horizontal Electrophoresis System: Maxicell EC360M Elect. Unit
Thermocyclers: Eppendorf- Mastercycler, Techne-genius
Magnetic Stirrer: Labor Brand/Hotplate L-81
Ovens: Dedeoğlu
pH meter: Hanna Inst.
Power Supplies: EC135-90 E-C
Refrigerator: AEG
UV Transilluminator : Vilbor Lourmant
Vortex: Nüve NM110
Water Bath: Memmert
Micropipettes: GILSON

APPENDIX B: SEQUENCES OF THE PRIMERS

Table B.1. Sequences of the primers

Primer	Sequence (5'-3')	Length (bp)
mKF1	CCC TTC GAT ACT GGC TGA AA	20
mKR1	TCA AGA AGG GCT CCA GAA GA	20
mKF2	TAT CGA CCG ATT TGT GCG TA	20
mKR2	AGC TGG GAC GAT CAA AGA AA	20
mKR3	AGA AGA AGC TGG GAC GAT CA	20
mKR4	AGG GCT CCA GAA GAT GTT GA	20
mKR5	GCT GGG ACG ACT AAA GAA AG	20
matKF1	ACT GTA TCG CAC TAT GTA TCA	21
matKF2	GTT CAC TAA TTG TGA AAC GT	20
matKR3	GAT CCG CTG TGA TAA TGA GA	20
matKF4	ACC CCA CCC CAT CCA TCT	18
matKF5	TGG AGY CCT TCT TGA GCG	18
matKF6	TCA GTG GTA CGG AAT CAA ATG C	22
matKR1	GAA CTA GTC GGA TGG AGT AG	20
matKR2	TTC ATG ATT GGC CAG ATC A	19
matKR2 2	ACG GGG CCA TAA GAA AGT CG	20

APPENDIX C: DATA SETS

Table C.1. Individual numbers of 18 populations with 2 individuals

	POPULATIONS	*matK* F1-R3	*matK* F5-R1	Entire *matK*
1	Acıpayam-Alcı	3, 10	3, 8	3, 10
2	Marmaris-Çetibeli	8, 9	8, 9	8, 9
3	Marmaris-Değirmenyanı	1, 6	1, 6	1, 6
4	Fethiye-Günlükbaşı	1, 2	1, 2	1, 2
5	Muğla-Kızılyaka	7, 8	1, 8	7, 8
6	Marmaris-Günnücek	1, 7	1, 7	1, 7
7	Marmaris-Günnücek	1, 5	1, 5	1, 5
8	Acıpayam-Bozdağ	1, 3	1, 2	1, 3
9	Marmaris-Hisarönü	1, 5	1, 5	1, 5
10	Muğla-Kıyra	1, 5	1, 6	1, 5
11	Köyceğiz-Köyceğiz	1, 3	1, 3	1, 3
12	Gölhisar-Pamucak	1, 4	1, 4	1, 4
13	Antalya-Serik	1, 4	1, 4	1, 4
14	Burdur-Söğütdağ	8, 11	8, 11	8, 11
15	Köyceğiz-Köyceğiz	4, 7	7, 9	4, 7
16	Aydın-Umurlu	22, 33	22, 33	22, 33
17	Muğla-Yatağan	3, 7	3, 6	3, 7
18	Muğla-Yılanlı	2, 3	1, 5	1, 3

Table C.2. Individual numbers of 18 populations with 3 individuals

	POPULATIONS	*matK* F1-R3	*matK* F5-R1	Entire *matK*
1	Acıpayam-Alcı	-	3, 8, 9	-
2	Marmaris-Çetibeli	8, 9, 12	8, 9, 12	8, 9, 12
3	Marmaris-Değirmenyanı	1, 6, 8	1, 4, 6	1, 6, 8
4	Fethiye-Günlükbaşı	1, 2, 4	1, 2, 4	1, 2, 4
5	Muğla-Kızılyaka	-	1, 3, 8	-
6	Marmaris-Günnücek	-	-	-
7	Marmaris-Günnücek	1, 5, 6	1, 2, 5	-
8	Acıpayam-Bozdağ	1, 3, 4	1, 2, 3	1, 3, 4
9	Marmaris-Hisarönü	1, 4, 5	1, 4, 5	1, 4, 5
10	Muğla-Kıyra	1, 5, 6	1, 5, 6	1, 5, 6
11	Köyceğiz-Köyceğiz	1, 3, 4	1, 3, 4	1, 3, 4
12	Gölhisar-Pamucak	-	-	-
13	Antalya-Serik	1, 4, 12	1, 4, 12	1, 4, 12
14	Burdur-Söğütdağ	-	-	-
15	Köyceğiz-Köyceğiz	4, 6, 7	4, 7, 9	4, 7, 9
16	Aydın-Umurlu	-	-	-
17	Muğla-Yatağan	3, 7, 12	3, 6, 7	3, 7, 12
18	Muğla-Yılanlı	2, 3, 5	1, 5, 7	1, 3, 5

Table C.3 Individual numbers of 18 populations with 4 individuals

	POPULATIONS	*matK* F1-R3	*matK* F5-R1	Entire *matK*
1	Acıpayam-Alcı	-	3, 8, 9, 10	-
2	Marmaris-Çetibeli	-	-	-
3	Marmaris-Değirmenyanı	1, 4, 6, 8	1, 4, 6, 8	1, 4, 6, 8
4	Fethiye-Günlükbaşı	-	1, 2, 4, 5	-
5	Muğla-Kızılyaka	-	1, 3, 4, 8	-
6	Marmaris-Günnücek	-	-	-
7	Marmaris-Günnücek	1, 5, 6, 7	1, 2, 5, 15	-
8	Acıpayam-Bozdağ	-	1, 2, 3, 4	-
9	Marmaris-Hisarönü	1, 4, 5, 9	1, 4, 5, 9	1, 4, 5, 9
10	Muğla-Kıyra	1, 5, 6, 13	1, 5, 6, 13	1, 5, 6, 13
11	Köyceğiz-Köyceğiz	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4
12	Gölhisar-Pamucak	-	-	-
13	Antalya-Serik	-	-	-
14	Burdur-Söğütdağ	-	-	-
15	Köyceğiz-Köyceğiz	4, 6, 7, 9	4, 6, 7, 9	4, 6, 7, 9
16	Aydın-Umurlu	-	-	-
17	Muğla-Yatağan	3, 6, 7, 12	3, 6, 7, 12	3, 6, 7, 12
18	Muğla-Yılanlı	2, 3, 5, 1	1, 5, 7, 3	-

APPENDIX D: A PART OF THE MEGA DATA FILE

Mega sequence data for Turkish sweet gum populations

#POP-03-MAR-DEG-06-DE06-integriloba

ATGGAGGAATCTCAAGGATATTTAGAACTAGATAAATCTGGGCAACATG
ACTTCCTATATCCACTTATCTTTCAGGAGTATATTTATGTACTTGCTCATG
ATCATGGTTTAAATAGATCGATTTTGTTGGAAAATTTGGGTTCTGACAAT
AAATTCAGTTCATTAATTGTGAAACGTTTAATTACTCGAATGTATCAACA
GAACCGTTTGATTATTTCCGCTAATGATTCTAACCAAAATCCATTTTTGG
GGCACAACAAGGATTTGTATTCTCAAATGATATCAGAGGGATTTGCAGT
CATTGTGGAAATTCCATTTCCCCTACGATTAGTATCTTCCCTAGAGAGGA
AAGAAATAGTAAAATCGCATAATTTACGATCAATTCATTCAGTATTTCCT
TTTTTAGAGGACAAGTTTTTACATTTAAATTATGTGTCAGATATACTAAT
ACCCCACCCCATCCATCTGGAAATATTGGTTCAAACCCTTCGATACTGGG
TGAAAGATGCTTCTTCTTTGCATTTATTACGATTCTTTCTCTACGAGTATC
GTAATTGGAATAGTCTTATTAATCCAAAGAAATCCATTTTCGTTTTTTCA
AAAAGGAATCAAAGATTATTCTTGTTTCTATATA??????????????????GA
ATCCGTCTTCGTTTTTCTTCGTAACCAATCTTCTCATTTACGATCAACATC
TTCTGGAGCCCTTCTTGAGCGAATATATTTCTATGGAAAAATAAAACATC
TTGTAGAAGTCTTTGCTAATGATTTTCAGGCCATCCTGTGGTTGTTCAAG
GATCCTTTCGTGCATTATGTTAGGTATCAAGGAAAATCAATTCTCGCTTC
AAAAGGAACACCTCTTCTGATGAATAAATGGAAATATTACCTTGTCAAC
TTCTGGCAATGTCATTTTTACGTGTGGTCTCAACCAGTAAGGATCTATAT
AAACCAATTATCCAATCATTCCCTTTACTTTCTGGGCTATCTTTCGAGTGT
GGGATTAAATCCTTTAGTGGTACGGAATCAAATGCTAGAAAATTCGTTT
ATAATAGATAATGCTATTAAAAAGTTCGATATCATAGTTCCAATTATTCC
TCTGATTGGATCATTGGCTAAAGCGAAATTTTGTAACGTATTAGGGCATC
CTATTAGTAAGCCGGCCCGGGCCGATTCATCAGATTCTGATATTATCGAC
CGATTTGTGCGTATATGCAGAAATCTTTCTCATTATCACAGCGGATCCTC
GAAAAAAAAGAGTTTGTATCGAATAAAGTATATACTTCGACTTTCTTGTG
CTAGAACTTTGGCTCGTAAACACAAAAGTCCTGTACGTGCTTTTTTGAAA
AGATTAGGTTCGGAATTATTGGAAGAATTCCTTACGGAGGAAGAACAAG
TTCTTTCTTTGATCGTCCCAGCTTCTTCTACTTCGCGGAGGTTATATAGAG
GGCGTATTTGGTATTTGGATATTATTTGTATCAACGATCTGGCCAATCAT
GAATGA

Arlequin sequence data of some of Turkish sweet gum populations

[Profile]

Title="*matK* gene"

NbSamples=28
GenotypicData=0
DataType=DNA
LocusSeparator=NONE
MissingData='?'

[Data]
[[Samples]]

SampleName="POP01-ACI-ALC-Unknown"
SampleSize=1
SampleData= {
AK03 1 ----

ATGGAGGAATCTCAAGGATATTTAGAACTAGATAAATCTGGGCAACATG
ACTTCCTATATCCACTTATCTTTCAGGAGTATATTTATGTACTTGCTCATG
ATCATGGTTTAAATAGATCGATTTTGTTGGAAAATTTGGGTTCTGACAAT
AAATTCAGTTCATTAATTGTGAAACGTTTAATTACTCGAATGTATCAACA
GAACCGTTTGATTATTTCCGCTAATGATTCTAACCAAAATCCATTTTTGG
GGCACAACAAGGATTTGTATTCTCAAATGATATCAGAGGGATTTGCAGT
CATTGTGGAAATTCCATTTCCCCTACGATTAGTATCTTCCCTAGAGAGGA
AAGAAATAGTAAAATCGCATAATTTACGATCAATTCATTCAGTATTTCCT
TTTTTAGAGGACAAGTTTTTACATTTAAATTATGTGTCAGATATACTAAT
ACCCCACCCCATCCATCTGGAAATATTGGTTCAAACCCTTCGATACTGGG
TGAAAGATGCTTCTTCTTTGCATTTATTACGATTCTTTCTCTACGAGTATC
GTAATTGGAATAGTCTTATTAATCCAAAGAAATCCATTTTCGTTTTTTCA
AAAAGGAATCAAAGATTATTCTTGTTTCTATATA??????????????????GA
ATCCGTCTTCGTTTTTCTTCGTAACCAATCTTCTCATTTACGATCAACATC
TTCTGGAGCCCTTCTTGAGCGAATATATTTCTATGGAAAAATAAAACATC

TTGTAGAAGTCTTTGCTAATGATTTTCAGGCCATCCTGTGGTTGTTCAAG
GATCCTTTCGTGCATTATGTTAGGTATCAAGGAAAATCAATTCTCGCTTC
AAAAGGAACACCTCTTCTGATGAATAAATGGAAATATTACCTTGTCAAC
TTCTGGCAATGTCATTTTTACGTGTGGTCTCAACCAGTAAGGATCTATAT
AAACCAATTATCCAATCATTCCCTTTACTTTCTGGGCTATCTTTCGAGTGT
GGGATTAAATCCTTCAGTGGTACGGAATCAAATGCTAGAAAATTCGTTT
ATAATAGATAATGCTATTAAAAAGTTCGATATCATAGTTCCAATTATTCC
TCTGATTGGATCATTGGCTAAAGCGAAATTTTGTAACGTATTAGGGCATC
CTATTAGTAAGCCGGCCCGGGCCGATTCATCAGATTCTGATATTATCGAC
CGATTTGTGCGTATATGCAGAAATCTTTCTCATTATCACAGCGGATCCTC
GAAAAAAAAGAGTTTGTATCGAATAAAGTATATACTTCGACTTTCTTGTG
CTAGAACTTTGGCTCGTAAACACAAAAGTCCTGTACGTGCTTTTTTGAAA
AGATTAGGTTCGGAATTATTGGAAGAATTCCTTACGGAGGAAGAACAAG
TTCTTTCTTTGATCGTCCCAGCTTCTTCTACTTCGCGGAGGTTATATAGAG
GGCGTATTTGGTATTTGGATATTATTTGTATCAACGATCTGGCCAATCAT
GAATGA
}

SampleName="POP02-MAR-CET-Integriloba"

SampleSize=1

SampleData= {

CE09 1 ----
ATGGAGGAATCTCAAGGATATTTAGAACTAGATAAATCTGGGCAACATG
ACTTCCTATATCCACTTATCTTTCAGGAGTATATTTATGTACTTGCTCATG
ATCATGGTTTAAATAGATCGATTTTGTTGGAAAATTTGGGTTCTGACAAT
AAATTCAGTTCATTAATTGTGAAACGTTTAATTACTCGAATGTATCAACA
GAACCGTTTGATTATTTCCGCTAATGATTCTAACCAAAATCCATTTTTGG
GGCACAACAAGGATTTGTATTCTCAAATGATATCAGAGGGATTTGCAGT
CATTGTGGAAATTCCATTTCCCCTACGATTAGTATCTTCCCTAGAGAGGA
AAGAAATAGTAAAATCGCATAATTTACGATCAATTCATTCAGTATTTCCT
TTTTTAGAGGACAAGTTTTTACATTTAAATTATGTGTCAGATATACTAAT
ACCCCACCCCATCCATCTGGAAATATTGGTTCAAACCCTTCGATACTGGG
TGAAAGATGCTTCTTCTTTGCATTTATTACGATTCTTTCTCTACGAGTATC
GTAATTGGAATAGTCTTATTAATCCAAAGAAATCCATTTTCGTTTTTTCA
AAAAGGAATCAAAGATTATTCTTGTTTCTATATA??????????????????GA
ATCCGTCTTCGTTTTTCTTCGTAACCAATCTTCTCATTTACGATCAACATC
TTCTGGAGCCCTTCTTGAGCGAATATATTTCTATGGAAAAATAAAACATC
TTGTAGAAGTCTTTGCTAATGATTTTCAGGCCATCCTGTGGTTGTTCAAG
GATCCTTTCGTGCATTATGTTAGGTATCAAGGAAAATCAATTCTCGCTTC
AAAAGGAACACCTCTTCTGATGAATAAATGGAAATATTACCTTGTCAAC
TTCTGGCAATGTCATTTTTACGTGTGGTCTCAACCAGTAAGGATCTATAT
AAACCAATTATCCAATCATTCCCTTTACTTTCTGGGCTATCTTTCGAGTGT
GGGATTAAATCCTTCAGTGGTACGGAATCAAATGCTAGAAAATTCGTTT
ATAATAGATAATGCTATTAAAAAGTTCGATATCATAGTTCCAATTATTCC
TCTGATTGGATCATTGGCTAAAGCGAAATTTTGTAACGTATTAGGGCATC

CTATTAGTAAGCCGGCCCGGGCCGATTCATCAGATTCTGATATTATCGAC
CGATTTGTGCGTATATGCAGAAATCTTTCTCATTATCACAGCGGATCCTC
GAAAAAAAAGAGTTTGTATCGAATAAAGTATATACTTCGACTTTCTTGTG
CTAGAACTTTGGCTCGTAAACACAAAAGTCCTGTACGTGCTTTTTTGAAA
AGATTAGGTTCGGAATTATTGTGAGAATTCCTTACGGAGGAAGAACAAG
TTCTTTCTTTGATCGTCCCAGCTTCTTCTACTTCGCGGAGGTTATATAGAG
GGCGTATTTTGTATTTGGATATTATTTGTATCAACGATCTGGCCAATCAT
GAATGA
}

SampleName="POP03-MAR-DEG-Integriloba"

SampleSize=1

SampleData= {

DE06 1 ----
ATGGAGGAATCTCAAGGATATTTAGAACTAGATAAATCTGGGCAACATG
ACTTCCTATATCCACTTATCTTTCAGGAGTATATTTATGTACTTGCTCATG
ATCATGGTTTAAATAGATCGATTTTGTTGGAAAATTTGGGTTCTGACAAT
AAATTCAGTTCATTAATTGTGAAACGTTTAATTACTCGAATGTATCAACA
GAACCGTTTGATTATTTCCGCTAATGATTCTAACCAAAATCCATTTTTGG
GGCACAACAAGGATTTGTATTCTCAAATGATATCAGAGGGATTTGCAGT
CATTGTGGAAATTCCATTTCCCCTACGATTAGTATCTTCCCTAGAGAGGA
AAGAAATAGTAAAATCGCATAATTTACGATCAATTCATTCAGTATTTCCT
TTTTTAGAGGACAAGTTTTTACATTTAAATTATGTGTCAGATATACTAAT
ACCCCACCCCATCCATCTGGAAATATTGGTTCAAACCCTTCGATACTGGG
TGAAAGATGCTTCTTCTTTGCATTTATTACGATTCTTTCTCTACGAGTATC
GTAATTGGAATAGTCTTATTAATCCAAAGAAATCCATTTTCGTTTTTTCA
AAAAGGAATCAAAGATTATTCTTGTTTCTATATA??????????????????GA
ATCCGTCTTCGTTTTTCTTCGTAACCAATCTTCTCATTTACGATCAACATC
TTCTGGAGCCCTTCTTGAGCGAATATATTTCTATGGAAAAATAAAACATC
TTGTAGAAGTCTTTGCTAATGATTTTCAGGCCATCCTGTGGTTGTTCAAG
GATCCTTTCGTGCATTATGTTAGGTATCAAGGAAAATCAATTCTCGCTTC
AAAAGGAACACCTCTTCTGATGAATAAATGGAAATATTACCTTGTCAAC
TTCTGGCAATGTCATTTTTACGTGTGGTCTCAACCAGTAAGGATCTATAT
AAACCAATTATCCAATCATTCCCTTTACTTTCTGGGCTATCTTTCGAGTGT
GGGATTAAATCCTTTAGTGGTACGGAATCAAATGCTAGAAAATTCGTTT
ATAATAGATAATGCTATTAAAAAGTTCGATATCATAGTTCCAATTATTCC
TCTGATTGGATCATTGGCTAAAGCGAAATTTTGTAACGTATTAGGGCATC
CTATTAGTAAGCCGGCCCGGGCCGATTCATCAGATTCTGATATTATCGAC
CGATTTGTGCGTATATGCAGAAATCTTTCTCATTATCACAGCGGATCCTC
GAAAAAAAAGAGTTTGTATCGAATAAAGTATATACTTCGACTTTCTTGTG
CTAGAACTTTGGCTCGTAAACACAAAAGTCCTGTACGTGCTTTTTTGAAA
AGATTAGGTTCGGAATTATTGGAAGAATTCCTTACGGAGGAAGAACAAG
TTCTTTCTTTGATCGTCCCAGCTTCTTCTACTTCGCGGAGGTTATATAGAG
GGCGTATTTGGTATTTGGATATTATTTGTATCAACGATCTGGCCAATCAT
GAATGA
}

----and the other populations are included as sampled above----

```
#Definition of the group structure:
((Structure))
StructureName="18 populations and 4 outgroups"
                NbGroups=1
            #18 populations
            Group= {
                        "POP01-ACI-ALC-Unknown"
                        "POP02-MAR-CET-Integriloba"
                        "POP03-MAR-DEG-Integriloba"
                        "POP04-FET-GUL-Integriloba"
                        "POP05-MUG-KIZ-Integriloba"
                        "POP06-MAR-GUN-Integriloba"
                        "POP07-MAR-GUN-Integriloba"
                        "POP08-ACI-BOZ-Unknown"
                        "POP09-MAR-HIS-Integriloba"
                        "POP10-MUG-KIY-Integriloba"
                        "POP11-KOY-KOY-Integriloba"
                        "POP12-GOL-PAM-Unknown"
                        "POP13-ANT-SER-Unknown"
                        "POP14-BUR-SOG-Unknown"
                        "POP15-KOY-KOY-Integriloba"
                        "POP16-AYD-UMU-Orientalis"
                        "POP17-MUG-YAT-Orientalis"
                        "POP18-MUG-YIL-Orientalis"
                    "L. orientalis_AF015651"
                        "L. orientalis_AF304519"
                        "L. orientalis_AF133220"
                    "L. acalycina_AF133222"
                        "L. acalycina_AF015649"
                    "L. formosana_AF133221"
                        "L. formosana_AF015650"
                    "L. styraciflua_AF133219"
                    "L. styraciflua_AF133218"
                    "L. styraciflua_AF015652"
}
```

APPENDIX E: PHYLOGENETIC TREES

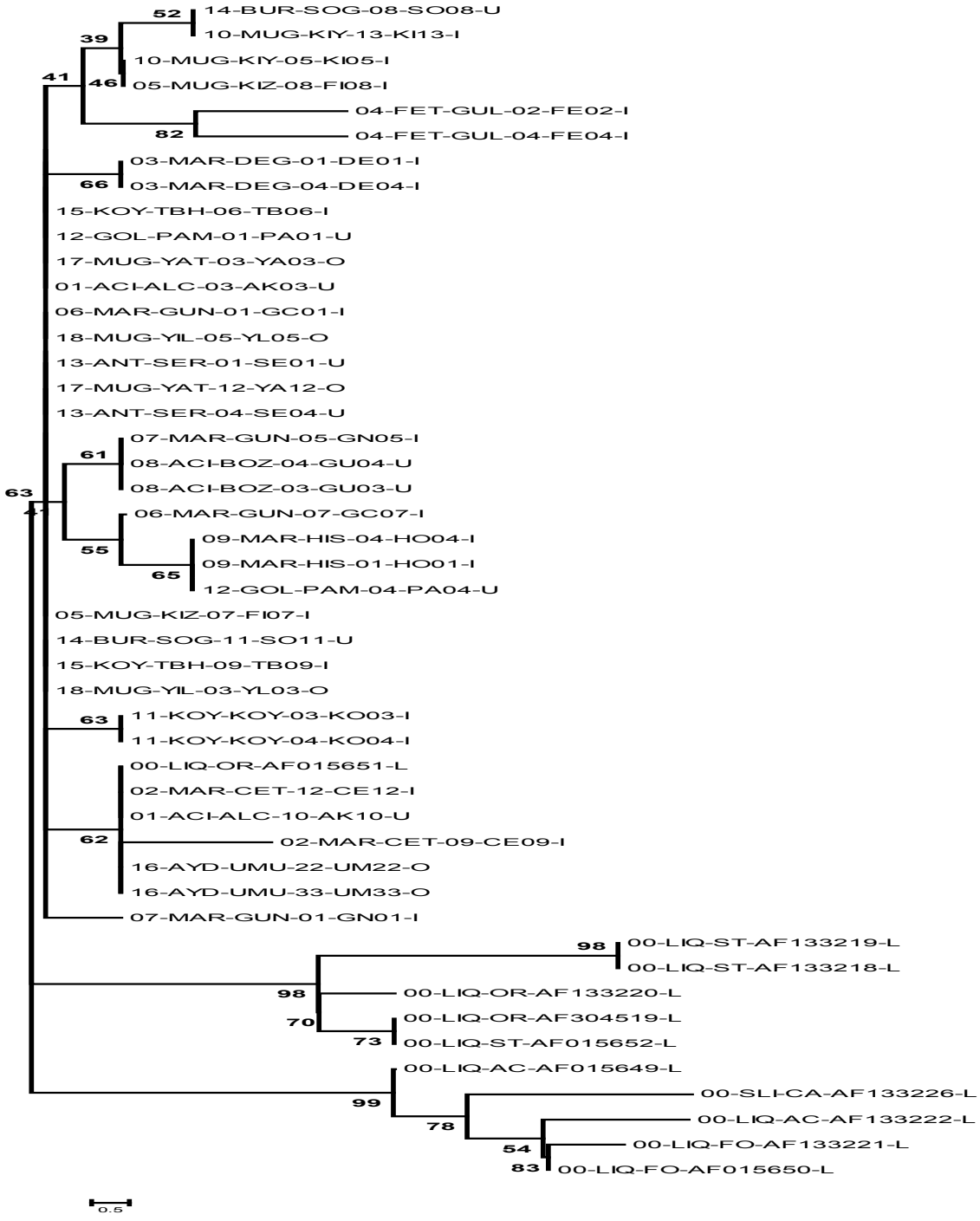


Figure E.1 Phylogenetic trees with two individuals from each 18 populations

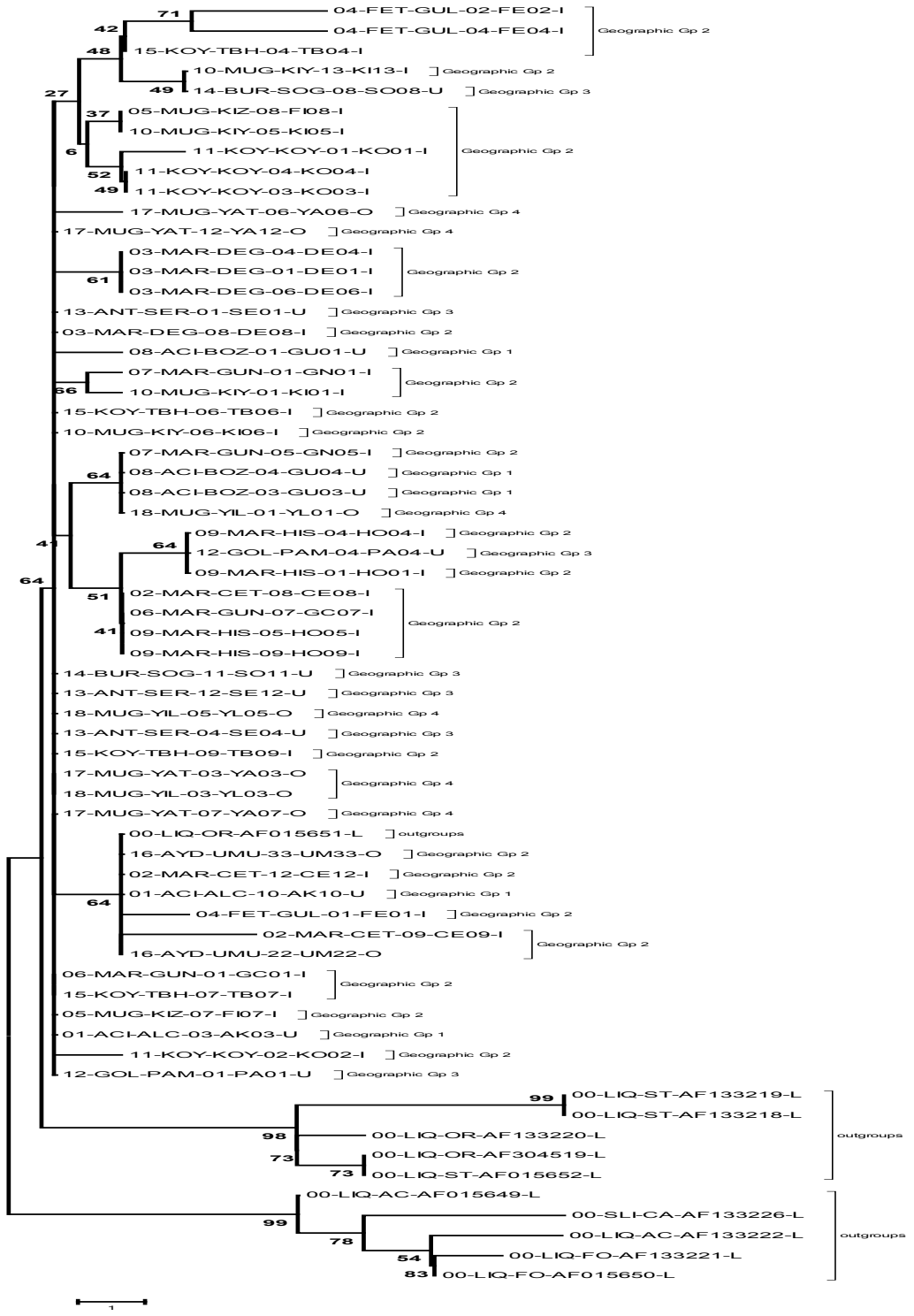


Figure E.2 Phylogenetic trees with all sequences obtained from each 18 populations