GENETIC COMPOSITION OF FOUR MARGINALLY LOCATED ANATOLIAN BLACK PINE (*Pinus nigra* subsp. *pallasiana*) POPULATIONS DETERMINED BY SSR MARKERS

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

BY

SILA AYGÜN

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

SEPTEMBER 2018

Approval of the thesis:

GENETIC COMPOSITION OF FOUR MARGINALLY LOCATED ANATOLIAN BLACK PINE (*Pinus nigra* subsp. *pallasiana*) POPULATIONS DETERMINED BY SSR MARKERS

Submitted by **SILA AYGÜN** in partial fulfillment of the requirements for the degree of **Master of Science in Department of Biological Sciences, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Orhan Adalı
Head of Department, **Department of Biological Sciences** _____

Prof. Dr. Zeki KAYA
Supervisor **Department of Biological Sciences, METU** _____

Examining Committee Members:

Prof. Dr. Musa Doğan
Department of Biological Sciences., METU _____

Prof. Dr. Zeki Kaya
Department of Biological Sciences, METU _____

Prof. Dr. C. Can Bilgin
Department of Biological Sciences,METU _____

Prof. Dr. Sümer Aras
Department of Biology, Ankara University _____

Assoc. Prof. Dr. Mehmet Somel
Department of Biological Sciences,METU _____

Date: 06/09/2018

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 : Sıla,Aygün

Signature :

ABSTRACT

GENETIC COMPOSITION OF FOUR MARGINALLY LOCATED ANATOLIAN BLACK PINE (*Pınus nigra* subsp. *pallasiana*) POPULATIONS DETERMINED BY SSR MARKERS

Aygün, Sıla

Msc.Department of Biological Sciences

Supervisor: Prof. Dr. Zeki Kaya

September 2018, 73 pages

Anatolian black pine (*Pinus nigra* Arnold subsp. *pallasiana* (Lamb.) Holmboe) is one of the most economically and ecologically important coniferous tree species of Turkey. Global climate change is obviously going to affect distribution and development of forest tree species and the marginal populations will be the most vulnerable ones. Since these populations are located in outer borders of their natural distribution area, they are expected to have original genetic makeup shaped by unsuitable living conditions. To predict the future of forest ecosystems, it is important to determine the genetic composition and adaptation processes on lifecycle of these populations.

In this thesis, genetic makeup of 4 Anatolian black pine marginal populations were studied in 3 different life stages (seed, seedling and mature stages). Totally 720 genotypes were investigated by means of 10 microsatellite DNA (SSR: single sequence repeats) markers. It was found that, heterozygosity values were low (Ho: 0.22±0.01) and level of genetic differentiation was high (F_{ST}: 0.13±0.03) among populations.

In addition, it was determined how natural selection and fitness affect these populations' genetic diversity in different life stages. Obtained data showed that genetic diversity gradually diminished after natural selection from seed to mature stages. Among populations in different life stages, seed stage of Beynam population possessed the highest diversity level (Ho: 0.26±.0.04). Hence, it is advised that when afforestation or conservation activities dealing with marginal black pine populations ,

genetic diversity patterns at the three stages of marginal populations should be considered to improve adaptation of plantations to changing adverse environmental conditions.

Keywords: Anatolian black pine, *Pinus nigra* subsp. *pallasiana*, marginal population, genetic diversity, inbreeding

ÖZ

MARJİNAL DAĞILIMLI DÖRT ANADOLU KARAÇAMI (*Pinus nigra* subsp. *pallasiana*) POPÜLASYONLARININ SSR MARKÖRLERİ YARDIMIYLA GENETİK YAPILARININ BELİRLENMESİ

Aygün,Sıla

Yüksek Lisans, Biyoloji Bölümü

Tez Yöneticisi: Prof. Dr. Zeki KAYA

Eylül 2018, 73 sayfa

Anadolu karaçamı (*Pinus nigra* Arnold subsp. *pallasiana* (Lamb.) Holmboe) ekonomik ve ekolojik açıdan ülkemizin en önemli ibreli orman ağacı türlerinden birisidir. Türkiye'de en geniş yayılışa sahip ikinci ibreli orman ağacıdır; ayrıca ağaçlandırma çalışmalarında en fazla kullanılan ikinci türdür**.** Küresel iklim değişikliğinin ağaç türlerinin dağılımını ve gelişimlerini de etkilemesi kaçınılmazdır. Bu değişimlerden en fazla etkilenecek popülasyonların başında marjinal popülasyonların olduğu düşünülmektedir. Çünkü bu popülasyonlar, türün yetişme alanının uç sınırlarında yer alır ve uygunsuz yaşam şartları nedeniyle de orijinal bir genetik yapıya sahiptir. Orman ekosistemlerinin geleceğinin anlaşılabilmesi için bu popülasyonlardaki uyum süreçlerinin ve yaşam döngüsündeki genetik yapının anlaşılması çok önemlidir.

Bu tezde, 4 Anadolu karaçamı popülasyonlarının genetik yapısı 3 farklı yaşam evresinde (tohum, fidan ve olgunluk) çalışılmıştır. Toplam 720 birey 10 SSR markörü kullanılarak incelenmiştir. Bunun sonucunda heterozigotluk değerlerinin düşük (Ho: 0.22±0.01) ve genetik farklılaşmanın yüksek olduğu(F_{ST}: 0.13±0.03) tespit edilmiştir.

Ek olarak, doğal seçilim ve uyumun bu popülasyonların farklı evrelerine nasıl etki ettiği de saptanmıştır. Elde edilen sonuçlara göre, tohumdan olgunluk çağına doğru genetik çeşitlilik değerleri doğal seçilimle birlikte gitgide azalmaktadır. Tüm popülasyonlar arasında is Beynam popülasyonunun tohum çağı en yüksek çeşitlilik derecesini göstermiştir. Dolayısı ile, marjinal karaçam ile yapılacak olan

ağaçlandırma ve gen koruma çalışmalarında yapılan ağaçlandırmaların değişen kötü çevresel koşullara uyumunu artırmak için marjinal popülasyonlarda genetik çeşitliliğin üç farklı yaşam evresinde yapılanmasının dikkate alınması önemle tavsiye edilmektedir.

Anahtar Kelimeler: Karaçam, *Pinus nigra* subsp. *pallasiana*, marjinal popülasyon, genetik çeşitlilik, kendileme

ACKNOWLEDGEMENTS

I owe my deepest gratitude to my supervisor Prof. Dr. Zeki Kaya for his guidance, supervision and endless patience throughout the study.

Additionally, I would like to thank Dr. Burcu Çengel for her help, support and guidance. Her guidance helped and motivated me all the time of the study. She is always more than an advisor. I can't never thank enough.

I would like to express my thanks to rest of the thesis committee members Prof. Dr. Musa Doğan, Prof. Dr. Can Bilgin, Prof. Dr. Sümer Aras and Assoc. Prof. Dr. Mehmet Somel for their helpful comments and criticisms on the manuscript.

This thesis was supported by the research fund: TÜBİTAK/TOVAG 116O719 Project "Genetic Structure and Relatedness of Four Marginally Located Anatolian Black Pine (*Pinus nigra* Subsp. *pallasiana*): The Potential of Genetic Fitness in the Afforestation Activities with These Populations"

I specifically thank Turan Demiraslan Scholarship Foundation under TEMA (the Turkish Foundation for Combating Erosion Reforestation and the Protection of Natural Habitats) because of the MS scholarship that I was honored.

I am also thankful to Forest Tree Seeds and Tree Breeding Research Directorate, General Directorate of Forestry (GDF) for the material and working opportunity, and the staff for their help, kindness and cosiness, specially, Dr. Burcu Çengel, Dr. Yasemin Tayanç, Dr. Selim Kaplan and Zühal Yılmaz.

I would like to thank Dr. Funda Değirmenci, Asya Çiftçi and Dr. Çiğdem Kansu for their endless help and support. Morover I am also thankful to all previous and current lab mates ; Dr. Pelin Acar, Dr. Alev Ateş, Bircan Taşkıran, Nurbahar Usta, Mert Çelik, Baki Çoban and Nehir Nebioğlu.

Last but not the least, I would like to present my special thanks to my family for supporting me and standing right beside in any circumstances throughout my life and my besty Duygu Pamukçu for finding her whenever I need.

To my dear family...

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

CLUMPP	Cluster Matching and Permutation Program
cp	Chloroplast
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribonucleic Acid
dNTP	Deoxy ribonucleotide triphosphate
EDTA	Ethylenediaminetetraaceticacid disodium salt
Fis	Inbreeding Coefficient Within Individuals
Fst	Differences Among Subpopulation (AMOVA)
GDA	Genetic Data Analysis
ha	Hectares
He	Expected Heterozygosity
Ho	Observed Heterozygosity
HWE	Hardy–Weinberg Equilibrium
m	Meters
MgCl2	Magnesium Chloride
Na	Number of Alleles
Ne	Number of Effective Alleles
NJ	Neighbor Joining
Nm	Number of Migrants
PCoA	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
SSR	Simple Sequence Repeats
TAE	Tris-Acetate-EDTA

CHAPTER 1

INTRODUCTION

1.1.Pinus

Pines are the most important genus of the 11 genera comprising Pinaceae. It is the largest and also most heteromorphic genus of conifers, having 113 species. Several authorities accepted that there are about 105-124 species, segregating some species or splitting the genus. The pines are almost entirely Northern Hemisphere taxon, growing on temperate and sub-tropical regions of the world (Farjon, 2010).

Pines are coniferous, evergreen and resinous trees or rarely shrubs reaching 15 - 45 m tall. They are long-lived trees, living from 100 up to 1000 years. The longest-lived pine is called "the Great, bristlecone pine" that is *Pinus longaeva*. One individual of this species is one of the world's oldest living organisms at around 4600 years old (http://conifersociety.org/conifers/conifer/pinus/). Pines are mostly monoecious, male and female cones present on the same tree. The seeds are mostly small and winged, that they are wind dispersed.

Pines are commercially important throughout the world, mostly on timber production since their timber is denser, resinous and durable.

There are five naturally distributed *Pinus* species in Turkey (Akkemik, 2014). These are;

- *Pinus brutia* Tenore: Turkish red pine,
- *Pinus nigra* subsp. *pallasiana* Arnold: Anatolian black pine
- *Pinus sylvestris* L.: Scots pine,
- *Pinus pinea* L.*:* Stone pine,
- *Pinus halepensis* Mid.: Aleppo pine.

Among these, *Pinus brutia* is the most widespread forest tree species, which occupies 5.6 million ha of forest area in Turkey. In addition, by reforestation volume, it is the most important forest tree species (Ormancılık istatistikleri, 2015). Highest priority was given to *Pinus brutia* in "National Tree Breeding and Seed Production

Programme", since it is a fast-growing species, resistant to drought and has high genetic diversity (Koski and Antola, 1993). Anatolian black pine, is the second widespread coniferous species with 4.2 million ha distribution. By reforestation volume, it is the second most important tree species. It has given high priority in the breeding program, due to its high reforestation volume.

1.2 Pinus nigra

Pinus nigra Arnold or black pine has some several common names associated with its distribution; Austrian pine, Calabrian pine, Corsican pine, Crimean pine and European black pine. The taxonomy of the black pine has occupied taxonomist for years, since it is a very variable taxon with a discontinuous range. Some authors accept it as a collective species due to its genetic and phenotypic variability (Fukarek, 1958; Vidakovic 1974, 1991; Farjon, 2010). Black pine is mainly native to Europe where its distribution extends from Spain through southern Europe to Anatolia (Farjon, 2010). It grows widely on eastern Spain, southern France, Italy, Austria, Macedonia, western Romania, Bulgaria and Greece on the Balkan Peninsula; east to southern Russia in the Crimes and south to Turkey; and on the islands of Cyprus, Sicily and Corsica with outliers in Algeria and Morocco (Figure 1.1).

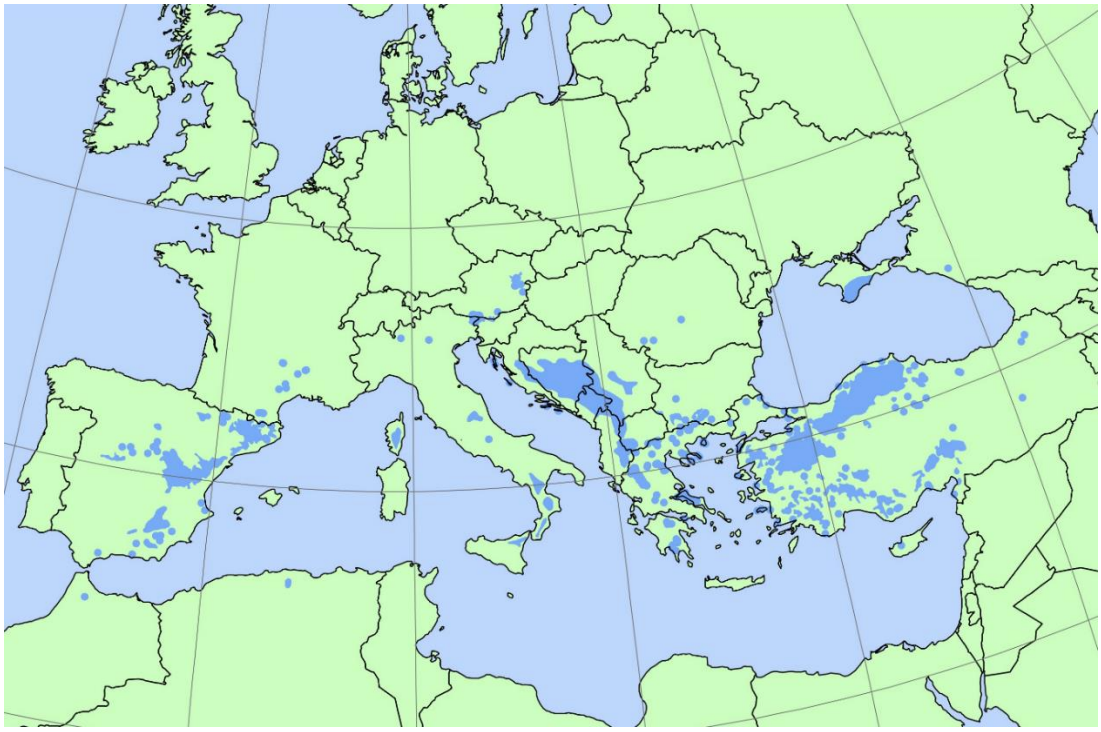


Figure 1.1 Natural distribution of *Pinus nigra* (Isajev *et al*., 2004)

Figure 1.1 Natural distribution of *Pinus nigra* (Isajev *et al*., 2004)

Black pine has a large variability based on morphological, anatomical, physiological and genetic characteristics (Vidakovic, 1974; Kaya *et al*., 1985; Matziris, 1989; Portfaix, 1989; Alptekin, 1986; Işık, 1990; Economou, 1990; Kaya and Temerit, 1994; Şimşek *et al.,* 1995; Velioğlu *et al*., 1999a).

In the latest taxonomical classifications, *Pinus nigra* is divided into 5 subspecies (Farjon, 2010; Debreczy and Racz, 2012). These are;

- **subsp. *salzmannii* (Dun) Franco:** Central and southern Spain, a few isolated populations in Pyrenees and Cevennes in France;
- **subsp. *laricio* (Poir.) Marie:** Corsica, Calabria and Sicily;
- **subsp. *nigra*:** Austria, Italy, Balkans and Greece;
- **subsp. *dalmatica* (Vis.) Franco:** Croatia and Dinaric Alps;
- **subsp. *pallasiana* (Lamb.) Holmboe:** Balkans, southern Carpathians, Crimea, Turkey, Cyprus and Syria.

1.2.1. ***Pinus nigra* subsp. *pallasiana***

Anatolian black pine (*Pinus nigra* subsp. *pallasiana)* is the subspecies of black pine naturally occurring in Turkey. It is differentiated from other subspecies by its shape of crown, needle and cone length (Yaltırık, 1993). Anatolian black pine or *Pinus nigra* subsp. *pallasiana* has also five varieties occurring in Turkey (Davis, 1965 ; Akkemik *et al*., 2010). These varieties are;

- *Pinus nigra* Arn. subsp. *pallasiana* (Lamb.) Holmboe var. *pallasiana*
- *Pinus nigra* Arn. subsp. *pallasiana* (Lamb.) Holmboe var. *şeneriana* (Saatçioğlu) Yaltırık
- *Pinus nigra* Arn. subsp. *pallasiana* (Lamb.) Holmboe var. *yaltirikiana* Alptekin
- *Pinus nigra* Arn. subsp. *pallasiana* (Lamb.) Holmboe var. *fastigiata* Businsky
- *Pinus nigra* Arn. subsp. *pallasiana* (Lamb.) Holmboe var. *columnaris-pendula* Boydak

1.2.2. Distribution

Anatolian black pine is a widespread mid-elevation species in the Taurus, western Anatolia and northern Anatolian mountains mainly west of the Anatolian Diagonal (Davis, 1965) (Figure 1.2). Anatolian black pine has 4.2 million ha distribution in Turkey where 1.5 hectares of these is fragmented and degraded. (Ormancılık İstatistikleri, 2015)

Anatolian black pine's distribution covers Mediterranean region, where is hot and drought in summer; to Black Sea region, with mild and rainy and to Central Anatolian steps. In the Black Sea region, it grows generally at 400-1400 m high, forming pure stands. After 1400 m sea level, it forms mixed stands with *Pinus sylvestris*, *Abies* spp., and *Quercus* spp. In Thrace, there are some small stands. In the western Anatolia, some of the best stands are found in Bozüyük, Keles, Dursunbey, Bigadiç, Sındırgı, Demirci, Simav, Emet and Tavşanlı. There are some pure stands at between elevations of 200 and 1400 m in the Ida Mountains. It forms some of the best stands at Muğla-Yılanlı, Köyceğiz, Fethiye, Gölhisar, Acıpayam and Denizli. In the southern Anatolia, it forms mixed stands especially with *Juniperus* spp. At 1200-1400 m. Finally, south margin of the distribution is in the Samandağ.

Anatolian black pine is adapted to several soil types such as calcareous, sandy and even pure limestone soils but requires deep soil. It is a shade-intolerant species, grows best on temperate (cool to cold) climates. Anatolian black pine is resistant to wind, drought and urban conditions.

At the climatic extreme points, *P. nigra* subsp. *pallasiana* populations have tendency to growing slowly, yet, are represented with long-lived individuals (Akkemik, 2010).

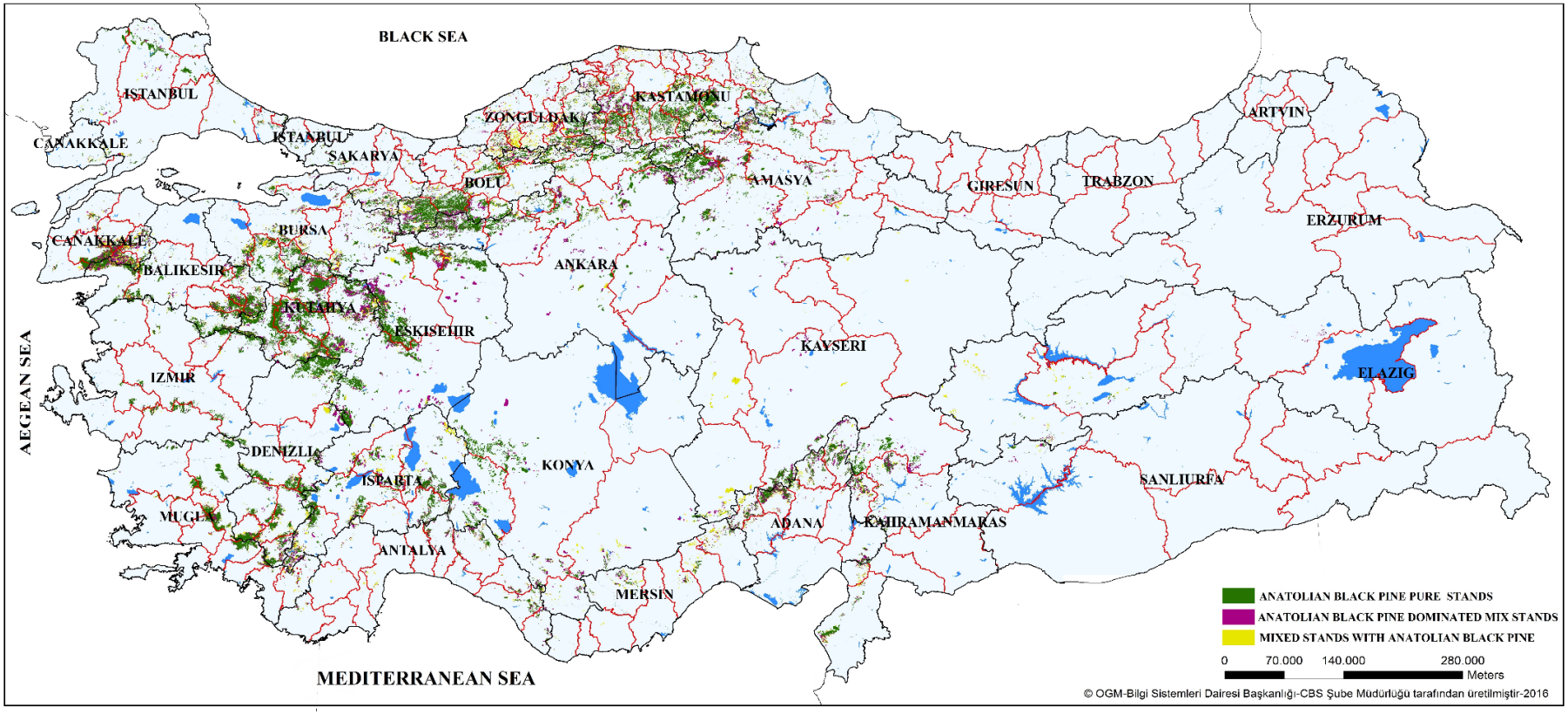


Figure1.2 Distribution of Anatolian black pine in Turkey *(*constructed by GDF, 2018*)*

1.2.3. Botany

Anatolian black pine is a high-quality, evergreen forest tree growing up to 20-40 m high. The crown of the young individuals is conical, becoming rounded and flat-topped with maturity (Figure 1.3a).

Figure 1.3 A view of an Anatolian black pine tree (Photographed by Burcu Çengel)

Figure 1.3b Anatolian black pine trunk (Photographed by Burcu Çengel)

Figure 1.3c Anatolian black pine male cones (Photographed by Burcu Çengel)

Figure 1.3 d Anatolian black pine mature female cone (Photographed by Burcu Çengel)

Figure 1.3 e Anatolian black pine one-year-old female cones (Photographed by Burcu Çengel)

The bark of Anatolian black pine is thick, scaly and furrowed on older trees (Yaltırık, 1988). Its bark is dark-grey to black in younger individuals, however it is paler colored in older ones (Figure 1.3b).

Branches are thick and dark colored; slightly ascending on young trees, upturned on older trees. One-year shoots are glabrous, light brown to orange-brown. Buds are resinous, ovate, gray-brown and pointed (Yaltırık, 1988).

Needles are stiff, 4-18 cm long, 1-2 mm in diameter, straight or curved, finely serrate and have pointed tips. Needles are green to dark green and bundled in clusters (fascicles) of two; leaf sheath persistent, 10-12 mm long.

Flowering occurs every year. However, abundant seed year is every 2-4 years. In their natural habitats, trees reach sexual maturity at about 15-20 years. Flowering period begins in May and fecundation occurs 13 months later. Pollen cones often grouped

close together on long shoots, axillary, clustered from a single bud, catkin-like, deciduous (Figure 1.3c).

Female cones are 5-12 cm long, yellow-brown, sessile and horizontally spreading (Figure 1.3d, Figure 1.3e). Cones ripen September-October and opening in the third year; contain about 30-40 seeds. Seeds are gray, 5-7 mm long with a 16-26 mm long wing. Seed scales are persistent, woody; bearing two inverted seeds.

1.2.4. Importance of *Pinus nigra* subsp. *pallasiana*

Anatolian black pine is a widespread and important timber species to Turkey. Its wood is rough and durable, so it can be used in poles, posts, mines, rail road ties, furniture, veneers and plywood, wood containers pulp and paper, sound insulation materials (Göker, 1969). Since it is a species with decorative varieties, it is also used in parks in small or larger groups

Since it is adapted to diverse soil types, topographic and climatic conditions, Anatolian black pine is the second most widely used forest tree for afforestation activities, mainly in Central Anatolian region,

1.3. Microsatellite Markers

Microsatellites, short tandem repeats (STRs) or simple sequence repeats (SSRs) are genomic repetitive regions in length with 1-6 bp represented. They occur both in prokaryotes and eukaryotes. They show high rate of presence in genome, in euchromotin in Eukaryotes and both coding and non-coding regions of prokaryotes.

Polymorphisms in DNA may be the result from insertion or deletion. If they are not fixed by DNA repairing mechanisms, they cause to differentiations called as mutations in genome. The mutations can be seen in SSR regions and generates different SSR alleles at loci. For this reason, different alleles may be seen in SSR regions, this makes SSR markers more informative than others (Vieira *et al*., 2016).

1.3.1. Microsatellite Markers in Population Genetics

Microsatellite markers are exclusive markers, have various features such as, hyper viability, wide genetic distribution, co-dominant inheritance, multi-allelic nature and chromosomal-specific location (Parida *et al*., 2009). Therefore, they can be used in genome mapping, parental tests and population genetics studies.

1.4. Population genetics studies dealing with Black pine and Anatolian black pine

Conventional population genetics studies on determination of genetic variation of forest trees were initiated with field trials or common garden studies. These tests are still widely used in tree breeding and very effective in identification of families and clones that are adapted to particular environments. There are numerous studies dealing with Anatolian black pine populations' genetic variation and/or adaptation in Turkey by means of field trials (Kaya and Temerit, 1994; Şimşek *et al*., 1995; Velioğlu *et al*.,1999a; Üçler and Gülcü, 1999).

Protein electrophoresis techniques such as isozymes followed field trial studies during the 1980s (Hamrick and Godt, 1990). The isozymes are selectively neutral markers and widely used in forest genetics for resolving many issues such as determination of genetic variation, mating patterns, pollen contamination in seed orchards, etc. There are abundant number of isozyme studies related to *Pinus nigra* (Bonnet-Masimber and Bikay-Bikay, 1978; Nicolic and Tucic, 1983; Finesci, 1984; Silin and Goncharenko, 1996).

There are several isozyme studies related to Anatolian black pine populations in Turkey. Isozyme studies were performed for the first time in Turkey, within the framework of a project called "*In Situ* conservation of plant genetic resources in Turkey" funded by Global Environment Facility (GEF). The project aimed to protect genetic resources of important forest tree species as well as wild relatives of crop species in selected pilot sites. Anatolian black pine was one of the target species in selected pilot areas of the Bolkar and Kazdağı Mountains of Turkey (Doğan *et al*., 1998).

Tolun and her co-workers (2000) determined the potential populations suitable for *in situ* conservation of genetic resources of Anatolian black pine in Bolkar Mountains. For this purpose, 14 enzyme systems were investigated by starch gel electrophoresis to determine the magnitude and pattern of genetic diversity among four populations. 24 loci were resolved for the 14 enzyme systems assayed. Only 7% of the total genetic diversity was among populations. Based on the results of this study, 2 populations were recommended for *in situ* conservation of genetic resources of the species in the Bolkar Mountains.

For the same purpose, Kazdağı populations of Anatolian black pine were investigated by starch gel electrophoresis (Cengel *et al.,* 2000). As a result, 94% of the total observed genetic variation was within populations. Based on the genetic diversity measurements and genetic distance between populations, three populations were recommended to conserve the genetic resources of the species in the Kazdağı.

In the last 30 years, along with DNA molecular marker development, molecular and population genetics studies were extremely accelerated for forest tree species. Kaya and Neale (1993) investigated the utility of RAPD markers in Anatolian black pine, for the first time. Then, several studies were carried out by RAPD markers for several forest tree species (Icgen *et al*., 2005; Velioğlu *et al.,* 2004; Velioğlu *et al.,* 2008).

Çengel and her colleagues (2012), was aim to test how efficiently plantations and seed orchards captured genetic diversity from natural populations of Anatolian black pine. For this purpose, Anatolian black pine seed sources were chosen from 3 different categories comprising seed stands, seed orchards and plantations from 4 different breeding zones of the species. Seed sources were screened with 11 RAPD markers. Estimated genetic diversity parameters were found to be generally high in all Anatolian black pine seed sources and the majority of genetic diversity is contained within seed sources (94%). No significant difference in genetic diversity parameters among seed source categories was found. In general, genetic diversity in seed stands has been transferred successfully into seed orchards and plantations. However, the use of seeds from seed orchards can increase the amount of genetic diversity in plantations further.

Naydenov and his co-workers (2006) studied genetic diversity of *P. nigra* populations from Bulgaria by chloroplast DNA microsatellites and terpene analysis. Nine populations were screened by 3 cpDNA microsatellite markers and significant genetic differentiation was observed between studied populations.

Afzal-Rafii and Dodd (2007), tried to test the hypothesis that fragmented populations in western Europe survived *in situ* during the last glacial rather than having been re-colonized in the postglacial period. For this purpose, genetic variation was assessed using 10 chloroplast DNA microsatellites. Among 311 individuals analyzed, 235 haplotypes were detected revealing high levels of chloroplast haplotype diversity in most populations. Obtained data suggested that chloroplast DNA is structured in black pine and disjoint populations in Western Europe are likely to have been present during the Last Glacial Maximum.

Jaramillo-Correa and his co-workers (2010) investigated the genetic structure of five conifers distributed on both sides of the Strait of Gibraltar(SG) using mitochondrial and chloroplast DNA markers. Across taxa, there was a significant overlapping between the SG and the genetic breaks detected, especially for the four Tertiary species surveyed including *Pinus nigra*. However, both markers' data revealed that *P. nigra* populations in Spain are significantly differentiated from Moroccan populations.

Giovanelli and his colleagues (2017), interested to resolve the taxonomy of the European black pine, developing specific markers for *P. nigra*. Nine nSSRs were specifically developed for the *P. nigra* starting from six transcriptomes sequenced using Next Generation Sequencing (NGS) techniques, while five previously developed for other pines were successfully transferred *to P. nigra*. Thus, they showed 14 nSSRs to be highly informative for population genetic studies.

1.5. Marginal Population Concept

By definition, marginal populations are the populations with original genetic structure that occurs on the edge of natural distribution of the species. They can be called as peripheral as well. They generally have small population size and isolated from the central populations. These populations have lower fitness compared with central populations. This argument is grounded by two characteristics of the populations: (i) The centrally distributed populations are larger than peripheral ones and (ii) As population gets smaller, the tendency of inbreeding in population gets higher. Increased inbreeding causes to collapse of gene pool and extinction of population.

The tendency of the low genetic diversity presence in marginal populations based on the three main assumptions: 1) Small populations have lower genetic diversity due to genetic drift (Lewontin, 1974); 2) The selective forces that are effective on the edge of distribution may erode the genetic diversity (Mayr, 1970; Lewontin, 1974); 3) Because of high isolation level, the emigration rate is become barely lower than it is in central populations. The lower level of immigration, may the population composed of individuals with lower fitness and accelerate the extinction rate.

1.6. Impact of Global Climate Change on Marginal Populations

Global climate change challenges the persistence and sustainability of forest ecosystems all over the world. Climatic changes have direct impacts on tree survival, growth and reproduction and possibly will cause shifts in species ranges. In addition

to climatic changes, human population growth, industrialization, urbanization and construction (airports, dams, highways, and hotels), pollution, over-exploitation of forestland and wrong political practices/policies also have pressure on forest ecosystems. Conserving forest genetic resources is both crucial for forest ecosystems and people's welfare. Marginal or peripheral forest tree populations may contain unusual adaptive traits that make them able to survive in marginal environments. Such traits may be particularly important for adapting forests to climate change. These unique genetic resources could be used to help forests adapt to the challenges of global climatic changes. Therefore, they should have the highest conservation priority. The FAO Global Action Plan for the conservation, sustainable use and development of forest genetic resources defines its Strategic Priority 7 as "support assessment, management and conservation of marginal and/or range limits forest tree population"

Turkey, like other countries in the Mediterranean basin is predicted to be seriously affected by climate change (Ministry of Forestry and Water Affairs, 2013). Turkey's forests especially Mediterranean region are subject to highly variable climatic and topographic conditions characterized by the presence of forest tree populations with high levels of marginality and differentiation (Ducci *et al*., 2014).

There are numerous studies on climate scenarios, adaptation capacity and the impacts of global climate change on forest populations in Europe and the Mediterranean (Echeverria *et al*., 2007; Aitken *et al*., 2008; Linare and Tiscar, 2010; Lindner *et al*., 2010; Ruiz-Labourdette *et al.,* 2012).

In order to understand the future of forest ecosystems, it is very important to recognize the adaptive processes and genetic structure in marginal populations. For several forest trees, such studies have been conducted (Ortego *et al*., 2010; Buschbom *et al*., 2011; Chybicki *et al.,* 2012).

Kaya and Temerit (1994) performed a study about marginal populations of Anatolian black pine in Central Anatolian region. They have established a field trial with the seed material sampled from seven marginal populations. After two years of observations, they concluded that genetic diversity is maintained within the populations rather than between populations.

Pandey and Rajora (2012) performed a research on *Thuja occidentalis*'s central and marginal populations' genetic diversity. They screened 13 populations with 6 SSR

markers and concluded that there was no differentiation between studied central and marginal populations.

CHAPTER 2

JUSTIFICATION AND OBJECTIVE OF THE STUDY

Anatolian Black pine (*Pinus nigra* Arnold subsp. *pallasiana*) is one of the most important conifer species of Turkey in terms of economic and ecological aspects. This species is the second most used species in afforestation activities. It covers nearly 16% of forest area of Turkey, yet it is fragmented. It is distributed as continuous core populations as well as marginal populations, which are residuals of natural old populations, specially located around Central Anatolian steps. There is no molecular study performed to reveal the genetic structure of these fragmented populations.

Global climate change is obviously going to affect distribution and development of forest tree species. The most vulnerable ones that will be affected from climatic alterations are marginal populations. Since these populations are located in outer borders of core distribution area of the species, they are expected to have original genetic makeup caused by unsuitable living conditions. To predict the future of forest ecosystems, it is important to determine the genetic composition and adaptation processes on lifecycle of these populations.

In this study, genetic structure (genetic diversity and relatedness) of marginal populations of Anatolian black pine aimed to be explored reveal in three life stages (mature, seed, seedling) by means of SSR markers.

Specifically, the study focused on;

- To determine genetic diversity parameters (polymorphism, heterozygosity and allele number) and genetic structure of studied marginal Anatolian black pine populations in three life stages (mature, seed and seedling)
- To help for development of new strategies in nursery and afforestation activities with information generated from the current study.

CHAPTER 3

MATERIAL AND METHODS

3.1. Plant Material

In this study, plant material was used from two projects called "Quantitative Variation in Marginally Located Anatolian Black Pine (*Pinus nigra* subsp. *pallasiana*) Populations in Central Anatolia." and "Genetic Diversity of Marginally Located Anatolian Black Pine (*Pinus nigra* subsp. *pallasiana*) Populations Central Anatolia and Development of Conservation Strategies" performed by Dr. Burcu Çengel and her co-workers. These projects were performed and funded Forest Tree Seeds and Tree Breeding Research Directorate, General Directorate of Forestry (GDF). From these projects' study materials, four populations were chosen, which represents northern and southern borders of the Central Anatolia (Figure 3.1 and Table 3.1.) Three of these were considered as marginal populations. Specifically, Beyşehir population is chosen from the optimum distribution area to compare with other marginal populations.

As a part of the projects that mentioned above, needles and cones were collected from selected 20 trees for each population. While selecting the trees, following criteria has been applied;

Trees had to be separated by at least 100 meters within a population,

- The elevation range of selected trees had to be within 300 meters,

- Trees had to be at similar ages.

In addition, cones were collected from upper 1/3 of the trees to avoid inbreeding

Unfortunately, these criteria could not be implemented for each population since marginal populations are generally had limited number of individuals.. Needle collection was performed from the healthy and 1-year old seedlings. 40-50 cones per tree was collected from the 1/3 upper of tree crown. All of the collected needles and cones were labelled with their locations and transferred to laboratory. In GDF labs,

needles were kept at -20°C till DNA extraction. Cones were opened in stoves to obtained seeds. Than seeds were stored at +4ºC, until its use for DNA extraction.

For the seedling stage experiments, needles were collected from the healthy 1- year old seedlings from the nursery. Seeds from 4 populations were sown in İlyakut Nursery of the General Directorate of Forestry. In the first year, seeds germinated and seedlings reached about 10 cm height. Their needles were collected with respect to populations and family numbers.

Remaining seeds were grouped with respect to corresponding mature trees in each population and germinated in laboratory so that their cotyledons(to increase tissue amount) could be used for DNA extraction..

The number of the tree genotypes sampled and used for each life stages in the study as follows:

1. Mature stage: 20 trees X 4 populations = 80 genotypes
2. Seed Stage : 20 trees X 4 populations X 4 seeds =320 genotypes
3. Seedling Stage: 20 trees X 4 populations X 4 seedlings =320 genotypes

In total, 720 different genotypes were sampled and studied.

Table 3.1 General information about studied populations

No	Status	Population Name	Forestry District	Latitude	Longitude	Altitude (m)	Area (ha)
1	GCF94*	Hasayaz	Ankara - Elmadağ	40°14’15”	33°22’06”	1400	580
2	GCF202	Beyşehir	Beyşehir - Beyşehir	37°59’20”	31°44’00”	1500	169.5
3	GCF219	Beynam	Ankara - Bala	39°41’10”	32°55’08”	1300	110.7
4	GCF204	Kadınhanı	Ilgın - Kadınhanı	38°07’24”	32°16’01”	1550	258

***GCF94: Gene Conservation Forest and its National Code**

Figure 3.1 Locations of studied populations

In this study, genetic structure of four marginal Anatolian black pine populations in 3 different life stages, totally 720 genotypes were studied. All populations and life stages are coded given below, after this point the codes of the populations were used.

Table 3.2 Codes given for the populations and life stages

Population	**Codes for Populations**	**Population codes for mature stage**	**Population codes for seed stage**	**Population codes for seedling stage**
Hasayaz	HAS	MatureHAS	SeedHAS	SeedlingHAS
Beyşehir	BEY	MatureBEY	SeedBEY	SeedlingBEY
Beynam	BYN	MatureBYN	SeedBYN	SeedlingBYN
Kadınhanı	KDN	MatureKDN	SeedKDN	SeedlingKDN

3.2.DNA Isolation

DNA isolation is performed by the Doyle and Doyle's CTAB protocol (1990) with small modifications.

Fresh needle tissues of mature and seedling stages were pulverized with liquid nitrogen in mortar by pestle. Powdered samples were stocked in -80°C until DNA isolation. For DNA extraction from seed tissue, seeds were germinated on Jacopsen Germination Table (Cophenhagen Table) for a week, in order to increase the tissue amount. Tissues were ground with glass rod with 200µl CTAB (Cetyl Trimethyl

Ammonium Bromide) solution, then step 1&2 were skipped and continued with step 3.

Doyle and Doyle's CTAB protocol (1990) used in this study:

1. 100 mg tissue were put into each 2 ml microcentrifuge tube and 500 µl CTAB solution (0.1 M Tris HCl pH: 8.0, 0.1 M EDTA, 0.25M NaCl) was added.
2. Tissues were ground with homogenizer (Omni International TH) for 1 minute.
3. 500 µl CTAB solution was added and tubes were transferred to water bath (65°C) for at least 1 hour. As the duration of incubation gets longer, isolated DNA becomes purer since it helps solubilization of lipids and proteins from DNA.
4. Following the incubation, tubes were centrifuged at 14 000 x rpm for 20 minutes at 4 °C.
5. Supernatant was transferred to clean microfuge tubes and mixed with 500 µl chloroform-octanol (24:1) solution to denature proteins and inactivate DNAse.
6. Centrifugation was repeated at14000 x rpm for another 20 minutes; resulting 2 phases; an aqueous phase, (containing DNA) and chloroform phase, (containing damaged proteins, lipids and many secondary compounds).
7. The aqueous phase was transferred to clean tubes and 750 µl cold isopropanol was added to precipitate the DNA.
8. The tubes were placed at -80ºC for overnight.
9. They were centrifuged at 14000 x rpm for 20 minutes and supernatant was discarded and the pellet was washed with 70% ice cold ethanol solution for removing chemical remains.
10. The pellets were left to dry at room temperature and re-suspended in 100 µl Tris buffer (0.1 M; pH: 8) to dissolve DNA.
11. DNA samples were stored at -20 ºC, until they were used.

3.3 DNA Quantification

DNA quantification of all individuals was performed with BioDrop µLITE-Nanodrop Spectrophotometer. The range of DNA concentration were 80-300 ng for mature, 200-3500 ng for seeds and seedlings. After quantification, all DNA samples were diluted to 20 ng/µl.

3.4. SSR Primers

The 20 SSR primers were used to determine the genetic differentiation among populations. SSR primers were selected from the previously developed for several Pinus species such as *Pinus sylvestris* (Sebastiani *et al*., 2011), *Pinus taeda* (Gonzales-Martinez *et al*., 2004) and *Pinus nigra* (Giovannelli *et al.,* 2017. Among these primers, 10 primers with the highest polymorphism were chosen to be used for Anatolian black pine samples. In order to detect on genetic analyzer system, primers' forward sequences were labeled with fluorescent dye. (Table 3.3)

Table 3.3 Detailed information on used microsatellite markers

Name of primer	SSR Locus	Primer Motif	Primer sequence (5'-3')	Expected allele size (bp)	Dye	Reference
CK 3	PtTX4001	$(CA)_{15}$	F:CTATTTGAGTTAAGAAGGGAGTC R:CTGTGGGTAGCATCATC	200	Red	Gonzalez-Martinez *et al.*, 2004
CK 8	psy118	$(GCA)_7$	F:ACTACCTGGCATTCGTCCTG R:GGATCTGGTCCATTTCGTGT	297-306	Blue	
CK 9	psy119	$(GCT)_7$	F:GGCTGTAATTGGCACAGGTT R:CGAGGTGGTACACAGCAACA	315-324	Blue	
CK 10	psy125	$(GCA)_5$	F:CAGCACGCGTTCTTTGTATC R:ACCGTTGCTCGTTGTCTTCT	214-244	Red	Sebastiani *et al.*, 2011
CK 11	psy136	$(GTC)_7$	F:TATCATCGAGAGCCCCAAAA R:GAAAGGCGAAAGCAAAAGTG	245-257	Green	
CK 13	psy144	$(CGG)_5$	F:TCCAAGTTCGGTTCCTTGTC R:GACACGATGGATTCCCTGAT	166-175	Blue	
CK 15	pn6360	$(TC)_{20}$	F:CAACTTCCTTCACCTGGCAC R:GACCCTTCTCAGCATCAACAC	324	Blue	
CK 16	pn7754	$(TA)_{12}$	F:TCAAGCTAATGCTGGGAACTC R:GTTGGTCGTGGTGACAAAGG	114	Blue	Giovannelli *et al.*, 2017
CK 18	pn6175	$(AT)_{12}$	F:ATTTCCCGCCTACCATTACC R:GTTACCTGCAATTCGTGTGG	201	Green	
CK 19	pn1403	$(GA)_{13}$	F:AGGCAATGATCATGTGGGTC R:GGATAGCCTGCAACTCCAATG	307	Black	

3.5 Polymerase Chain Reaction Amplification of Microsatellite Loci

The total volume of the PCR amplifications is 20 µl. The major component of the PCR mix is 5x HOT FIREPol® Blend Master Mix Ready to Load(Solis BioDyne, Tartu, Estonia), that contains, HOT FIREPol® DNA polymerase, proofreading enzyme,5x blend Master Mix Buffer,15mM MgCl2, 2mM dNTPs of each, BSA, blue dye, yellow dye, compound that increases sample density for direct loading. The other ingredients are; forward primer, reverse primer, template DNA and dH_2O.

Labelled primers were prepared as 10 pmol/µl solutions. For CK 16 and CK 18 primers "touch-down" PCR conditions were used. The optimized PCR mixtures is given in Table 3.4 and the conditions which optimized in Polymerase Chain Reaction is given in Table 3.5. The amplification of SSR loci were carried out with thermal cycler (Eppendorf-Mastercycler, Eppendorf). All template DNA were amplified using labeled forward primer and unlabeled reverse primer.

Table 3.4 Optimized PCR mixtures for studied primers

Primer name	**dH_2O**	**Master Mix(5x)**	**TaqDNA polimerase (5 u/µl)**	**Primer (10 mM)**	**DNA (20 ng/µl)**	**Total volume**
CK 16, 18	9.55 µl	4 µl	0.25 µl	0.6 µl+0.6 µl	5 µl	20 µl
CK3,8,9,10, 11,13,15,19	10.8 µl	4 µl	-	0.6 µl+ 0.6 µl	4 µl	20 µl

Table 3.5 PCR amplification conditions for studied SSR primers

Primer name	Step	Temprature	Duration	Cycle number	Definition
CK3	1	94ºC	10 min.	1	Denaturation
CK8	2	94ºC	30 sn.	35	Denaturation
CK9		58 ºC	40 sn.		Annealing
CK10		72ºC	50 sn.		Extension
	3	72ºC	10 min.	1	Final Extension
CK11 CK13	1	94 ºC	10min.	1	Denaturation
	2	94 ºC	30 sn.	35	Denaturation
		56 ºC	40 sn.		Annealing
		72 ºC	50 sn.		Extension
	3	72 ºC	10 min.	1	Final Extension
CK15 CK19	1	95ºC	15 min.	1	Denaturation
	2	95ºC	30 sn.	30	Denaturation
		58 ºC	1 min.		Annealing
		72ºC	50 sn.		Extension
	3	72ºC	10 min.	1	Final Extension
CK16 CK18	1	95ºC	15 min.	1	Denaturation
	2	94ºC	30sn.	10	Denaturation
		60oC (1oC drops in each cycle)	30 sn.		Annealing
		72ºC	40 sn.		Extension
	3	94ºC	30 sn	30	Denaturation
	4	50ºC	30 sn.		Annealing
	5	72ºC	40 sn		Extension
	6	72ºC	10 min	1	Final Extension

3.6. Agarose Gel Electrophoresis

In order to screen the PCR products, 2.5% agarose gel is used for electrophoresis. 5x HOT FIREPol® Blend Master Mix (5X) which includes reagents required for PCR and additional dyes for direct loading to agarose gel (Bioshop® Agarose, Routine

Use) to be observed during electrophoresis. Agarose (Bioshop® Agarose, Routine Use) is mixed with 1x TAE (Tris Acetate EDTA) buffer solution (AppliChem 50X) and boiled via microwave to get gels. In order to detect the PCR products on gel, each sample (5 µl) was mixed with 2 µl dye (Bioshop Bio-View White) before loading to gel. Electrophoresis is processed with 1X TAE running buffer under 100 Volts for 50-60 minutes. To determine the magnitude of bands, a DNA standard (Thermo GeneRuler 50bp) was used. Gels were visualized under UV light by using a gel documentation system (Vilber Lourmat, Infinity).Further information about used chemicals and equipments was given in Appendix A.

3.7. SSR Fragment Analysis

According to results of gel electrophoresis, templete DNAs from sampled trees, seeds and seedlngs which result in clear bands were collected as quart. Each of them labelled with different fluorescent dye and was sent to be analyzed. Fragment analysis by Capillary Electrophoresis was carried out at the BM Laboraratory Facilities, Ankara. Procedure for fragment analysis was provided by Applied Biosystems 3730 XL DNA Analyzer (Applied Biosystems, Foster city, CA, USA). This system uses a standard marker (The GeneScan ROX labelled 400H). Further information about assay procedure was given in Appendix B.

The result of this analysis was electrophenogram, which was visualized with Peak Scanner V.1 (Applied Biosystems). An example of electrophenogram was given in Figure 3.2. In electrophenogram, since each primer was labeled different color of fluorescent dye, each locus had fragment peaks with different colors (Table 3.3). In the software, each peak was checked for allele size and allelic data were recorded for all locus of each individual to get microsatellite genotypes. A part of Excel matrix showing genotypes allele sizes was given in Appendix C.

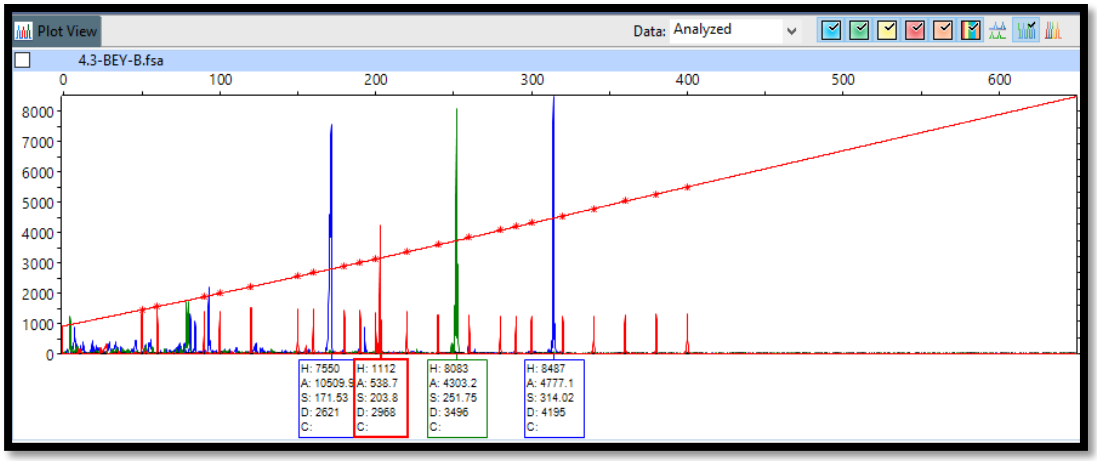


Figure 3.2 A microsatellite electrophenogram of 4 different SSR loci used in the study

(y- axis: Relative Flourescence Units, x- axis: Size of fragment).

3.8 Genetic Diversity Analysis

In order to determine the population diversity parameters, all data were resulted of locus and genotypes within populations were collected and arranged to be analyzed with various population genetics software. Examples of the data formats were given in Appendix D. After that, the data were processed by the packaged software of POPGENE V1.31 (Yeh and Boyle, 1997) and GenAlEx 6.5 (Peakall and Smouse, 2006, 2012).With these analysis, number of alleles (Na), mean effective number of alleles (Ne), observed heterozygosity (Ho) and expected heterozygosity (He), Fixation index (F), private alleles and F statistics (F_{IS} and F_{ST}), were estimated for loci and populations.

The number of alleles represents observed alleles that were detected in the studied populations. The effective number of alleles is used in calculating expected heterozygosity in studied populations. The effective number of alleles is used to designate the number of alleles that are equally relevant to the expected heterozygosity as in the studied population. Private alleles are special to a population and present at any frequencies. The observed heterozygosity is the number of heterozygous individuals per loci. The expected heterozygosity is computed with individual allele frequencies. F-statistics (F_{IS} and F_{ST}) are basically explain the genetic situation of populations. (Nei, 1977) F_{IS} is also known as inbreeding coefficient. F_{ST} is used for determining the genetic divergence among

subpopulations. The further information on calculation of the statistics that used in thesis is given in Appendix E.

In order to evaluate the amount of genetic variation within life stages, among populations within life stages, and among populations of *P. nigra,* Analyses of Molecular Variance (AMOVA) was performed. Besides, it was used to estimate of fixation indices (Fct, Fsc and Fst) based on the Fst values (number of different alleles) with 1,000 permutations using ARLEQUIN. (Excoffier and Lischer, 2010)

3.9. Genetic Structure Analysis

Genetic structure analysis were performed by STRUCTURE v2.3.4 (Pritchard *et al*., 2000) which uses Bayesian clustering approach to determine actual population number (K) via clustering of individual genotyping data. Despite the initial information of the clusters coming by geographic location, this analysis detects clusters apart from the prior information. Admixture model was used with a burn-in of 50,000 and 250,000 Markov chain Monte Carlo iterations, possible cluster numbers (K) tested from K=1 to K=4 and 10 replications were performed for each K.

First, genetic grouping of considered populations were determined. The actual K was identified with the web-based program called Structure-Harvester. Using Evanno Method (Evanno *et al.,* 2005) the highest value of Delta K with lowest value of Standard deviation was estimated. Multiple runs for the true K value were averaged with the CLUMPP software (Jakobsson and Rosenberg, 2007) to identify the best alignment to the replicate results of the cluster analysis. After determining K, cluster pattern figure is revealed by another web-based program that is the PopHelper (Francis, 2017).

In order to show genetic differentiation of the populations, a Principle Coordinate Analysis (PCoA) method based on pairwise Fst matrix was carried out by GenAlEx software. (Peakall and Smouse, 2012). The genetic relationship among the populations were revealed as phenogram with the GDA software based on co-ancestry identity by Neibor-Joining method (Nei, 1978).

CHAPTER 4

RESULTS

4.1 Genetic Diversity

4.1.1. Genetic Diversity Parameters for all Loci

In this study, 10 SSR primers were used to understand the genetic diversity pattern of marginal *Pinus nigra* subsp. *pallasiana* populations. The obtained genetic diversity parameters were given at Table 4.1. All studied primers were found to be polymorphic. Observed allele numbers for each locus were varied between 2.58 (CK9) and 11.33 (CK15). When all loci were considered together, expected heterozygosity (He=0.44) was found to be higher than observed heterozygosity (Ho=0.22). Both expected and observed heterozygosity values were detected in the lowest value for the locus CK9 (He=0.10, Ho=0.04), while it was the highest in the CK3 locus (He=0.66; Ho=0.51). The most diverse locus with highest He (He=0.84), effective and observed allele numbers (Ne=7.09; Na=11.33) was the CK15 locus. When F_{ST} values of all loci were considered, the CK8 locus has the highest F_{ST} value and makes the highest contribution to genetic differentiation of populations. Genetic differentiations among all seed, seedling and mature Anatolian black pine populations were found to be high ($F_{ST=}$ 0.13). F_{IS} values known as inbreeding coefficient for each locus were found to be positive and high (except for CK 16) indicating occurrence of considerable inbreeding in the populations (Table 4.1)

Table 4.1 Genetic diversity parameters of the studied 10 SSR loci (standard errors of estimation)

Locus	N	Na	Ne	Nm	Ho±SE	He±SE	F_{IS}	F_{ST}
CK8	59.17	4.00	1.54	0.47	0.08±0.04	0.23±0.07	0.56	0.35
CK10	58.33	3.83	1.79	0.99	0.18±0.04	0.33±0.07	0.31	0.20
CK3	57.58	5.17	3.15	1.63	0.51±0.05	0.66±0.03	0.22	0.13
CK9	59.08	2.58	1.13	4.37	0.04±0.02	0.10±0.03	0.56	0.05
CK11	58.08	3.25	1.56	3.29	0.24±0.03	0.33±0.04	0.23	0.07
CK13	58.33	3.25	1.31	2.22	0.14±0.03	0.19±0.04	0.17	0.10
CK15	58.75	11.33	7.09	4.39	0.24±0.03	0.85±0.02	0.72	0.05
CK16	59.50	4.75	1.54	3.41	0.29±0.04	0.32±0.04	0.08	0.07
CK18	59.25	8.00	4.00	1.38	0.19±0.03	0.68±0.05	0.73	0.15
CK19	58.83	9.25	5.13	2.17	0.26±0.04	0.77±0.04	0.63	0.10
Mean	58.69	5.56	2.83	2.43	0.22±0.02	0.44±0.03	0.46	0.13

4.1.2. Genetic diversity parameters for all populations

Genetic diversity parameters calculated for all populations are polymorphic regarded to studied loci (Table 4.2). Observed allele number (Na) is ranged from 4.00 in matureKDN to 7.7 in seedlingBEY with the mean of 5.56 while effective allele number (Ne) is ranged from 1.95 seedlingKDN to 3.82 seedBYN. The highest expected heterozygosity is calculated as 0.64 in seedBYN population and the lowest value of 0.38 in matureKDN population. Similarly, the lowest observed heterozygosity (Ho=0.16) was found in matureBEY and in matureKDN populations and the highest one (Ho=0.28) was found in seedlingBEY population. The SeedBYN population had the highest number of private allele (8), whereas some of the populations did not contain private alleles at all. Allelic patterns of populations were shown in Figure 4.1. It is clearly observed that the highest values of allele numbers were seen in the seedBYN and highest value of expected heterozygosity was found in the matureBYN.

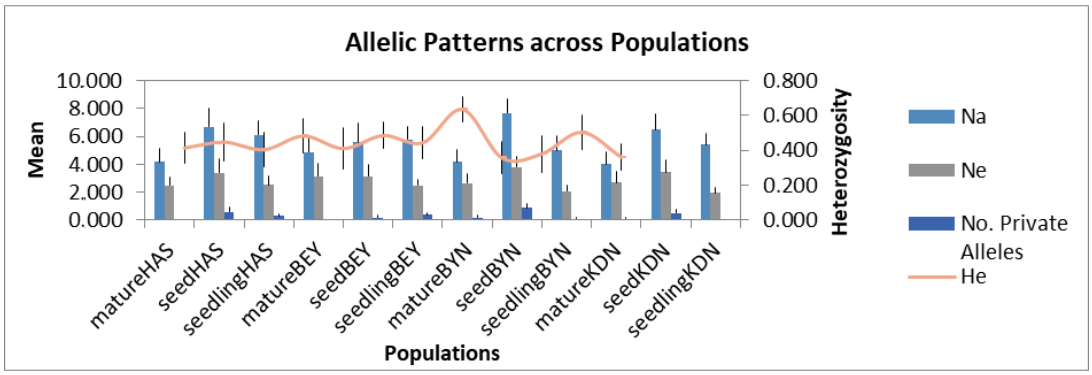


Figure 4.1 Allelic pattern across populations

Table 4.2.a Genetic diversity parameters for three stages of studied populations (SE: standard errors of estimation)

Population	**N**	**Na**	**Ne±SE**	**Ho±SE**	**He**	**P.allele**	**F_{IS}**
matureHAS	18.40	4.20	2.48	0.20±0.06	0.41±0.09	0.00	0.56
seedHAS	79.00	6.70	3.40	0.25±0.06	0.45±0.10	6.00	0.35
seedlingHAS	79.00	6.10	2.52	0.17±0.04	0.40±0.10	3.00	0.47
matureBEY	19.30	4.90	3.17	0.17±0.04	0.48±0.10	0.00	0.56
seedBEY	78.90	5.60	3.15	0.23±0.06	0.41±0.11	4.00	0.31
seedling BEY	77.90	5.80	2.48	0.28±0.06	0.49±0.08	2.00	0.36
matureBYN	19.60	4.20	2.67	0.27±0.06	0.44±0.09	2.00	0.22
seedBYN	76.80	7.70	3.82	0.26±0.04	0.64±0.07	8.00	0.60
seedling BYN	79.30	5.00	2.09	0.21±0.06	0.36±009	1.00	0.23
matureKDN	17.90	4.00	2.72	0.16±0.05	0.38±0.10	1.00	0.41
seedKDN	79.40	6.70	3.43	0.22±0.06	0.51±0.10	5.00	0.57
seedling KDN	78.80	5.40	1.95	0.19±0.05	0.36±0.07	0.00	0.39
Mean	58.69	5.56	2.83	0.22±0.02	0.44±0.03	2.67	0.42

Table 4.2.b: Analyses of molecular variance (AMOVA) as weighted averages over loci in the populations

Source of variation	Sum of squares	Variance components	Percentage of total variation	Fixation Indices
Among stages	124.288	0.08	3.03	FCT: 0.03*
Among populations	293.958	0.26	9.94	FSC : 0.10*
within stages				
Within populations	3168.320	0.26	87.03	FST : 0.13*
Total	3586.566	2.60	100	-

Fst=differences among subpopulation, Fsc=differences among population within groups, and Fct=difference among groups for the total population. *Significant at $p < 0.05$.

The analysis of 10 loci in four populations at three stages were resulted in positive F_{IS} value, which is a clear indication of heterozygosity deficiency compared with Hardy-Weinberg expectations. Positive F_{IS} values are also indicator of inbreeding in the studied populations. Three percentage of total variation was found to be among stages, while 10% of variation was shared among populations within stages. When AMOVA was carried out, the result revealed that this groups were significantly differed from each other. (Table 4.2.b)

4.2. Genetic diversity pattern in different life stages

In order to reveal the genetic structure of the four populations, genetic diversity parameters were calculated for each life stage, separately.

4.2.1. Mature stage

For this phase of the experiment, 80 trees (20 individuals from each of the four populations) were analyzed. The results of the analysis were given at Table 4.3.

Table 4.3 Genetic diversity parameters estimated for populations at the mature stage

Population	**N**	**Na**	**Ne**	**Ho±SE**	**He±SE**	**F_{IS}**	**F_{ST}**
MatureHAS	18.40	4.20	2.48	0.20±0.06	0.41±0.09	0.56	
MatureBEY	19.30	4.90	3.17	0.17±0.04	0.48±0.10	0.56	
MatureBYN	19.60	4.20	2.67	0.27±0.06	0.44±0.09	0.22	
MatureKDN	17.90	4.00	2.72	0.16±0.05	0.38±0.10	0.41	
Mean	18.80	4.33	2.76	0.20±0.03	0.43±0.05	0.43	0.06

For each mature population, observed allele number and effective allele numbers for populations were in close range With respect to genetic diversity, the most diverse population appeared to be matureBEY with Na=4.90; Ne= 3.17 and He=0.48. As seen in Table 4.3, expected heterozygosity values were higher than observed heterozygosity values for mature populations. Among mature populations, mean F_{IS} value was 0.43; which indicates considerable inbreeding in the mature Anatolian black pine populations. The genetic differentiation among mature populations was found to be at moderate level (F_{ST} = 0.06).

4.2.2 Seed Stage

To assess the pattern of genetic diversity at seed stage of populations, a total of 320 seeds were analyzed from four populations (20 trees X 4 seeds per tree X 4populations)). Results of these analysis were given in Table 4.4

Table 4.4 Genetic diversity parameters estimated for the seed stage of populations (SE : standard error of estimation)

Population	**N**	**Na**	**Ne**	**Ho±SE**	**He±SE**	**F_{IS}**	**F_{ST}**
SeedHAS	79.00	6.70	3.40	0.25±0.06	0.45±0.11	0.35	
SeedBEY	78.90	5.60	3.15	0.23±0.06	0.41±0.12	0.31	
SeedBYN	76.80	7.50	3.80	0.26±0.04	0.64±0.07	0.59	
SeedKDN	79.40	6.50	3.43	0.22±0.06	0.51±0.10	0.57	
Mean	78.53	6.58	3.45	0.24±0.03	0.50±0.05	0.46	0.11

In the seed stage of the populations, effective allele numbers for each population are close to each other. Most genetically diverse population was seedBYN with Na=7.5,

Ne= 3.80, Ho=0.26, and He=0.59). It is appeared that mean observed heterozygosity was considerably smaller than expected heterozygosity in all populations. The expected heterozygosity value was ranged from 0.41 in seedBEY to 0.64 in seedBYN. When the populations were considered separately, the lowest Ho value was found in seedKDN. All estimated F_{IS} values were positive indicating inbreeding occurrence. Mean F_{ST} value was calculated as 0.11 which indicates high rate of genetic differentiation among populations at the seed stage.

4.2.3. Seedling Stage

In the seedling stage of populations, 320 seedlings from four populations (20 trees X 4seedlings per tree X 4 populations) were screened with SSR markers. The results of genetic diversity analysis at the seedling stage of populations were given in Table 4.5.

Table 4.5 Genetic diversity parameters estimated for populations at the seedling stage (SE : standard error of estimation)

Population	**N**	**Na**	**Ne**	**Ho±SE**	**He±SE**	**F_{IS}**	**F_{ST}**
seedlingHAS	79.00	6.10	2.52	0.18±0.04	0.41±0.10	0.48	
seedlingBEY	77.90	5.80	2.48	0.28±0.06	0.49±0.08	0.36	
seedlingBYN	79.30	5.00	2.08	0.21±0.06	0.36±0.09	0.23	
seedlingKDN	78.80	5.40	1.95	0.20±0.05	0.36±0.08	0.39	
Mean	78.75	5.57	2.26	0.21±0.03	0.40±0.04	0.36	0.09

The seedling stage of the populations, although observed allele numbers are high and similar to the seed stage of the populations, effective allele numbers were found to be lower than those in mature and seed stage of the populations. In terms of allelic diversity, seedlingHAS is the most diverse population with Na=6.10 and Ne= 2.52. When seedling stage of the populations were evaluated together, it is clear that mean observed heterozygosity (Ho=0.21) is lower than expected heterozygosity (He=0.40). Similar to mature and seed stage of the populations, seedling stage of the populations had positive F_{IS} values indicating excess of homozygote individuals and inbreeding in these populations. F_{ST} value was found to be 0.09 shows increased level of genetic differentiation of seedling populations.

4.3. Population Genetic Structure

4.3.1. Principal Coordinate Analysis

Principal coordinate analysis of three different life stages was conducted using F_{ST} values of the populations. Separations or aggregations were determined only by their genetic relatedness. So that, this analysis was independent from the geographic locations of the populations.

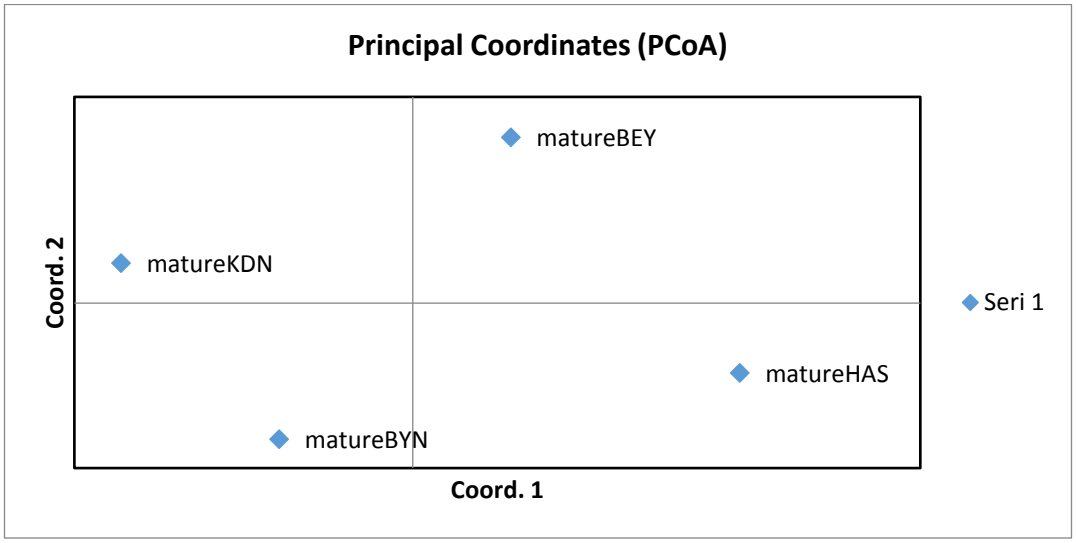


Figure 4.2 Principal coordinates analysis (PCoA) for mature stage of the populations

It is clear that all of four mature populations are separated well based on their genetic structures. When considering principle coordinate 1, matureKDN and matureBYN were well separated from matureBEY and matureHAS, while the coordinates were useful for separation of matureKDN and matureBEY from matureBYN and matureHAS populations (Figure4.2).

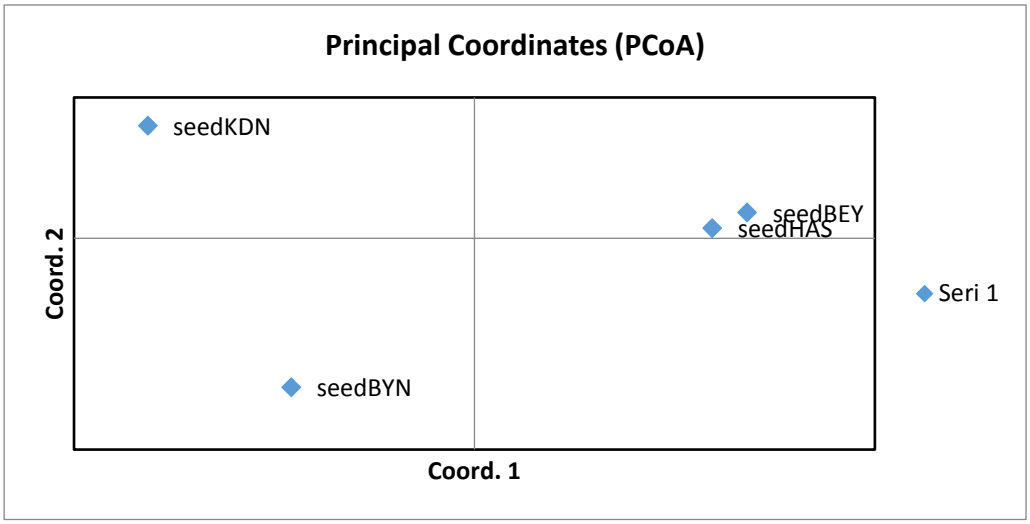


Figure 4.3 Principal coordinates analysis (PCoA) for seed stage of the populations

When all four seed populations are evaluated according to their genetic structure, the seedHAS and seedBEY populations were closely located. This means, these two populations are genetically similar and share similar genetic background. The remaining two populations are located apart from the seedHAS and seedBEY (Figure 4.3).

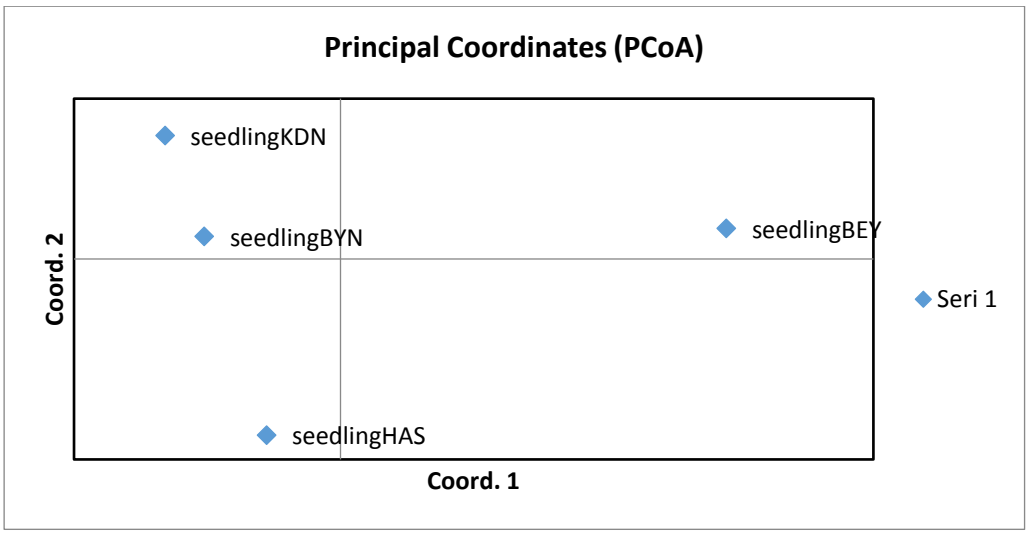


Figure 4.4 Principal coordinates analysis (PCoA) for seedling stage of the populations

Principle coordinate analysis done for the seedling stage of the populations revealed that the seedlingKDN and seedlingBYN populations are genetically close to each

other. According to coordinate 1, the seedlingBEY population is genetically distant from remaining seedling populations (Figure 4.4).

4.3.2. Genetic Relationships Among Populations

The constructed dendrograms for mature, seed and seedling stages of the populations of Anatolian black pine by neighbor-joining method Nei (1978) were given in Figure 4.5 a, b, c. In three stages of populations, it was clear that there were two major genetic groups. The first group included the HAS, KDN and BEY populations which are the most closely related while BYN population was distantly related to the others and formed other major group. It is clearly seen from neighbor-joining tree that genetic relationships among populations have been conserved through three life stages.

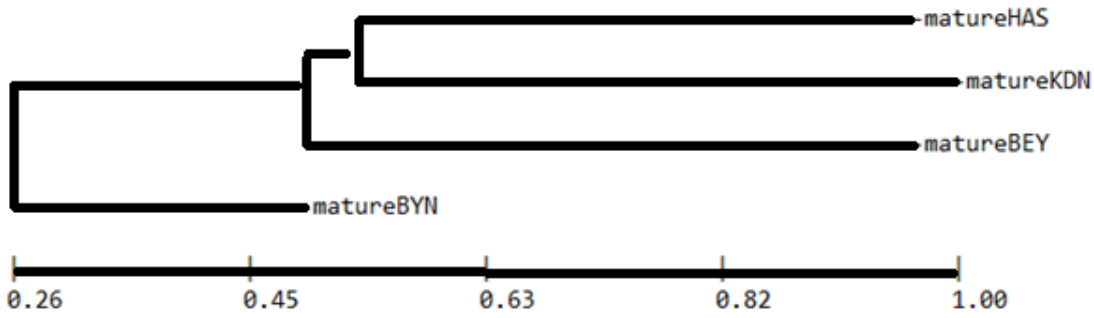


Figure 4.5a Dendrogram constructed among populations at mature stage

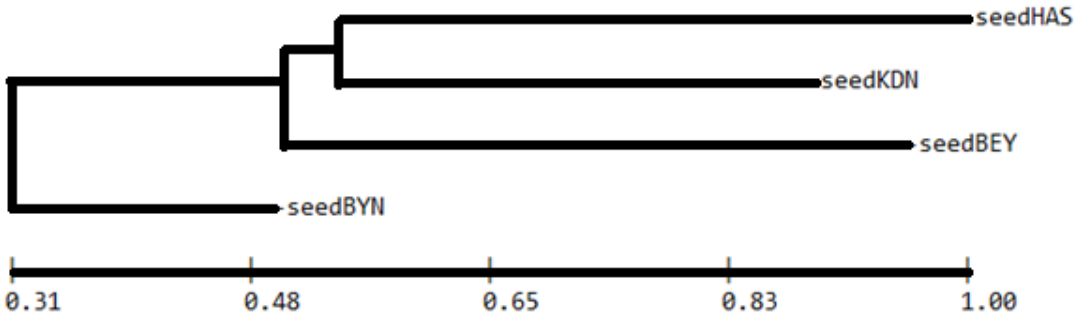


Figure 4.5b Dendrogram constructed among populations at seed stage

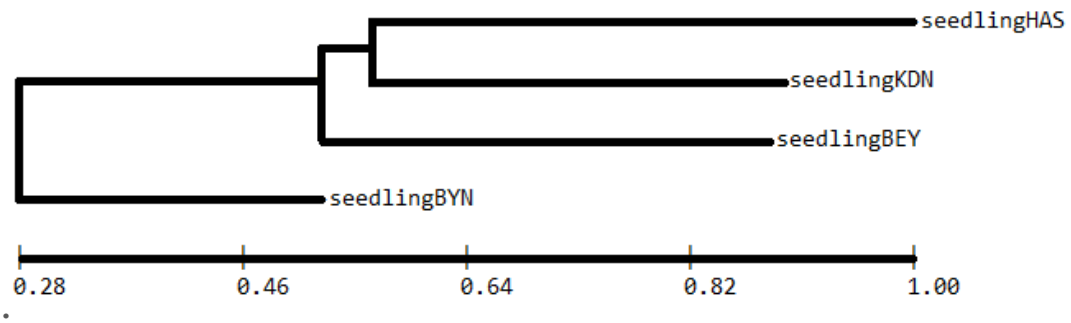


Figure 4.5c Dendrogram constructed *among populations at seedling stage*

4.3.3. Population Structuring and Clustering Patterns

Genetic structure analysis were performed by STRUCTURE v2.3.4 (Pritchard *et al.,* 2000) which use Bayesian clustering approach and determined actual population number (K) via clustering among individuals. After determining K, cluster pattern figure was revealed by another web-based program PopHelper (Francis, 2017). The clustering pattern was disclosed the given result mean that there are K number of separate gene pool present in these populations.

4.2.3.1. Populations of mature stage

In order to reveal the genetic structure of mature populations, STRUCTURE analysis was done, the K was calculated as three (Figure 4.6 and Table 4.6), and clustering pattern of these populations was given in Figure 4.7.

According to the results, mature *P. nigra* subsp. *pallasiana* populations from four different locations seem to be recruited from three gene pools that each population contains trees from these three gene pools. Geographical distance did not have much role in shaping the genetic clusters because high level of admixture among marginal Anatolian black pine populations. The BEY population was allocated to the cluster 1with the membership values of 0.4347, while the HAS population was placed in to cluster 2 with the highest estimated membership values (0.5163). BYN and KDN populations were allocated to third cluster with 0.4304 and 0.4403 membership values, respectively. In Figure 4.7. Blue, orange and white colors represent cluster1, cluster2, and cluster3, respectively.

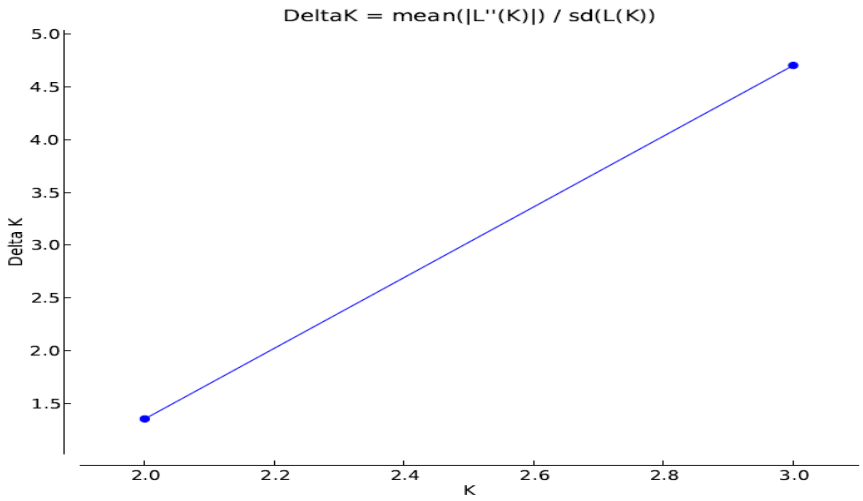


Figure 4.6 Graph of delta K values of mature stage of the without prior information

Table 4.6 Proportion of membership of each pre-defined population in each of the 3 clusters

Population	Inferred cluster1	Inferred cluster 2	Inferred cluster 3	Number of individuals
matureHAS	0.3180	0.5163	0.1657	20
matureBEY	0.4347	0.3209	0.2244	20
matureBYN	0.2412	0.3284	0.4304	20
matureKDN	0.3154	0.2443	0.4403	20

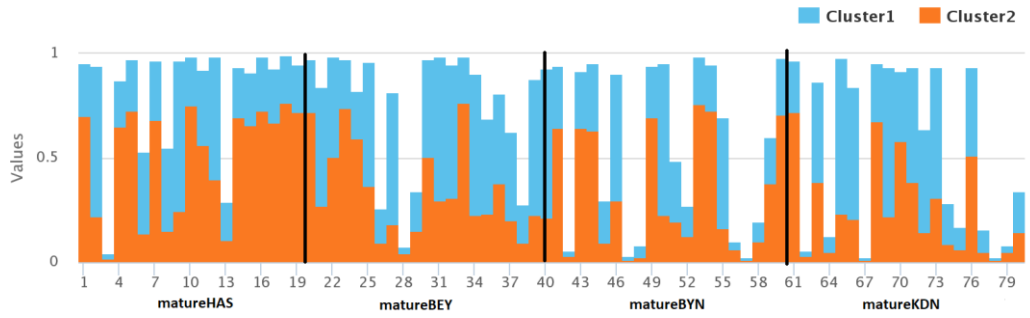


Figure 4.7. Genetic *STRUCTURE* analysis for mature stage of populations (K=3)

4.3.3.2. Seed Stage of Populations

At seed stage, the populations were analyzed with STRUCTURE and actual K value, was found to be two. (Figure 4.8 and Table 4.7) The results of this analysis was given in Figure 4.9. The admixture of the 2 gene pools is clearly revealed. The Blue

and red colors in Figure 4.9 represent cluster1 and cluster2, respectively. The SeedHAS(0.9542) and SeedBEY(0.9659) populations had very high membership values in cluster 2, whereas the SeedBYN(0.7398) and SeedKDN (0.5398) showed genetic pattern of admixture from both gene pools. The SeedBYN and seedKDN populations contain genetic material from the first cluster.

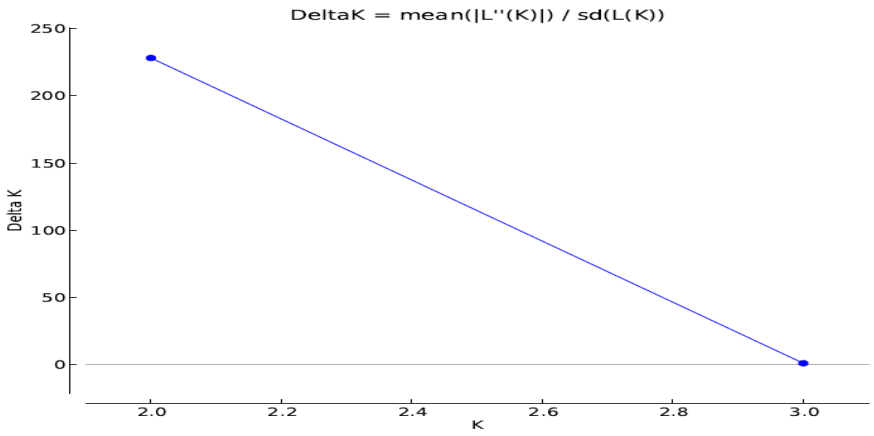


Figure 4.8 Graph of delta K values of seed stage of the without prior information

Table 4.7 Proportion of membership of each pre-defined population in each of the 2 clusters

Population	Inferred cluster1	Inferred cluster 2	Number of individuals
seedHAS	0.0458	0.9542	80
seedBEY	0.0341	0.9659	80
seedBYN	0.7398	0.2602	80
seedKDN	0.5398	0.4602	80

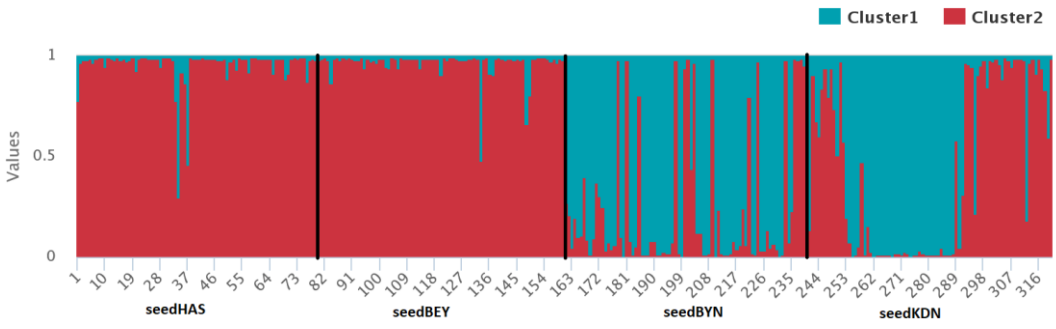


Figure 4.9 Genetic STRUCTURE analysis for seed stage of the populations (K=2)

4.3.3.3. Populations of seedling stage

Actual K value for seedling stage of the populations was found to be two (Table 4.8). The clustering pattern of seedling stage of the populations was given in Figure 4.11. This patterns have two clusters representing two gene pool as well. The Green and orange colors in Figure 4.8 represent cluster1 and cluster2, respectively. The KDN (0.9842) and BYN (0.8062) seedling populations had the highest estimated membership values and allocated to the cluster 1, while, seedlingBEY (0.9609) and seedling HAS (0.8852) similarly had high membership values and allocated to cluster 2. The Seedling KDN population is almost pure and distinguished from others. Although individuals have high rate of admixture of two gene pools, however SeedlingBEY population belonging to cluster 2 has more homogenous genetic structure.

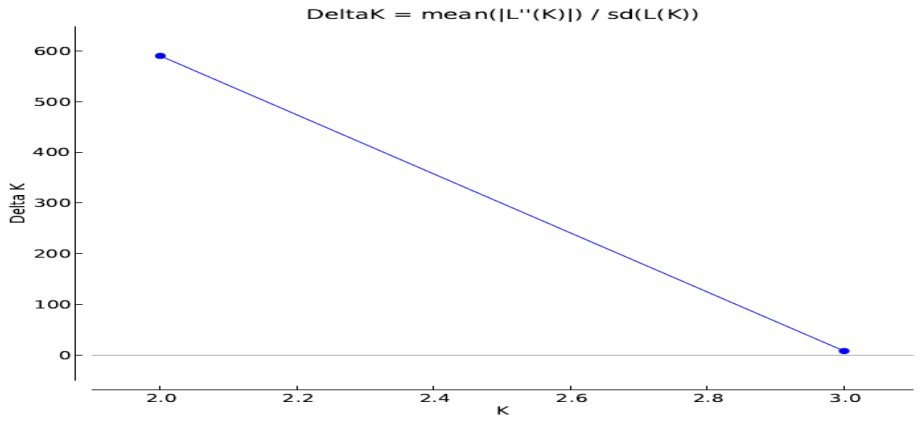


Figure 4.10 Graph of delta K values of seed stage of the without prior information

Table 4.8 Proportion of membership of each pre-defined population in each of the 2 clusters

Population	Inferred cluster1	Inferred cluster 2	Number of individuals
seedlingHAS	0.1148	0.95	80
seedlingBEY	0.0341	0.9659	80
seedlingBYN	0.7398	0.2602	80
seedlingKDN	0.5398	0.4602	80

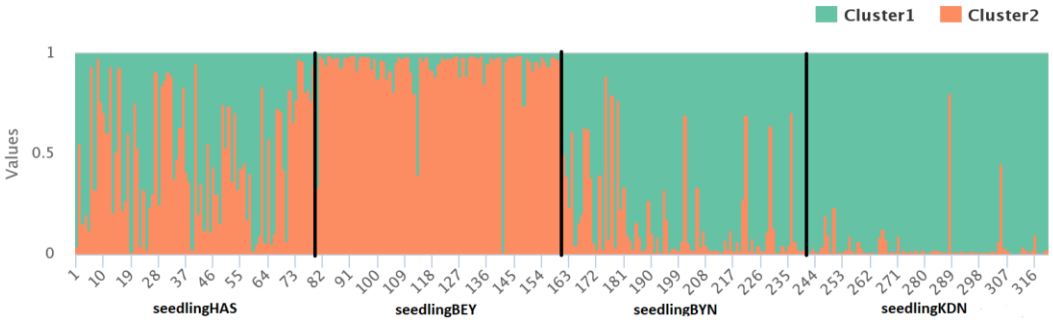


Figure 4.11 Genetic STRUCTURE analysis for seedling stage of the populations (K=2)

CHAPTER 5

DISCUSSION

In this study, genetic diversity and genetic structure of mature, seed and seedling stages of four marginally distributed *Pinus nigra subsp. pallasiana* populations from Central Anatolia were studied. The study is unique to reveal genetic structure of marginal populations of black pine at different life stages in Turkey. The obtained results could give some clues about natural selection effects on these type of populations, as well as their adaptive potential against global climate change.

Today, there are considerable number of marginal Anatolian Black pine populations existing around Central Anatolia. These populations are remaining parts of pre-existing larger Black pine populations in Central Anatolia from approximately 2000 BC (Wertime, 1983; Tsoumis, 1988). Therefore, the marginal populations survived under many environmental and human mediated forces during centuries, which are drought, temperature, fire, logging, urbanization and so on. Along with continuous climate changes, human pressure have caused disturbances on them and resulted in fragmentation. Additionally, they are geographically isolated with small population size. These features bring low range of gene flow and inbreeding depression, eventually may cause to collapse of the populations. Even under these conditions, they have been able to continue their existence. An assumption can be constructed based on this idea is that possibility of the genetic diversity they had. They may have some adaptively advantageous alleles to become selected and adapted the conditions through time. Therefore, the main aim of this study is proofing and revealing the genetic diversity of these populations. Moreover, to disclosure how the natural selection and adaptive process effect different stages of these trees.

5.1 Genetic diversity of microsatellite loci used in the study

The microsatellite markers used in this study are developed from different *Pinus* species *(Pinus taeda, Pinus sylvestris and Pinus nigra Arn*). None of the used

markers in the study was monomorphic. Therefore, all of the markers contributed to genetic diversity and structure of these populations.

Expected heterozygosity is an important measurement of genetic diversity. For each and every locus, estimated expected heterozygosity values were higher than those observed heterozygosity estimated. Similarly; the mean of expected heterozygosity values was higher than the mean of observed heterozygosity. In these populations, number of homozygote genotypes appear to be higher than heterozygote genotypes. Our result is consistent with the previous studies carried out in *Pinus rezedowskii* (Delgado *et al.,* 1999; He=0.21 and Ho=0.16), *Pinus albicaulis* (J. Krakowski *et al*., 2002; He=0.25 and Ho=0.21); and *Pinus sylvestris* L. (Pavia *et al*., 2014; He=0.81 and Ho=0.57). In these previous studies including in our study, all populations exhibited higher values of expected heterozygosity than observed ones.

5.2 Genetic diversity and the population structure of the Anatolian black pine

The aim of this study is not only determining genetic diversity and genetic relativeness of the marginally located populations, but also finding out factors affecting genetic diversity pattern at different life stages of the populations. To do so, mature, seed and seedling stages of the populations were studied.

Private allele number of the populations gives the information about genetic diversity of the populations. Higher private allele number was observed generally in the seed stage of the populations. These consequences might be caused by a reason that seed stage of the populations possess higher genetic diversity compared to seedling and mature stages of the populations. Specifically, seed BYN population possessed high number of private allele (Pa=8.0), effective allele (Ne=3.80) and both of expected (He=0.64) and observed (Ho=0.26) heterozygosity values. At the seed phase, since environmental conditions are not in effect, genetic potential of the populations may be maintained, but this potential may be reduced in later stages due to various selective forces operating. The other explanation of higher genetic diversity in seed stage of the Beynam population could be continuous afforestation activities called as ‘Green Belt Reforestation’ around Ankara. The activity covers 24.000 ha including METU, Beytepe and Bilkent Campuses. In those plantations black pine occupies majority of the planted area.(Ağaçlandırma Seferberliği Sonuç Raporu, 2012) In addition to current afforestation activity, METU campus alone covers 1650 ha of

Anatolian Black pine planted since 1956. These plantations are possible to affect the diversity structure of the black pine populations of Beynam as a result of pollen transfer which could be easily occur among these close locations. Therefore, high level of gene flow was seen among Beynam Population and plantations compared with other studied populations.

F_{IS} value is an important parameter for informing reproduction tendency whether mating are among homozygous or heterozygous individuals. It is calculated by using expected and observed heterozygosity values. If F_{IS} value is negative, population reproduction tendency is among heterozygous individuals. However, if F_{IS} is positive, that means homozygous individuals have higher tendency to contribution to reproduction. According to our estimation, The F_{IS} values were positive in all studied populations. Hence, it is clearly implicated that there is heterozygosity deficiency and inbreeding presence in these populations. This deficiency of heterozygosity indicates that population structure of these population may be changed by the presence of anthropogenic pressure, habitat destructions, urbanization, fragmentation.

When considering mature, seed and seedling stages of the populations of marginal black pine populations, it is clearly seen that the mean expected heterozygosity value for three different stages is higher than observed heterozygosity. Inbreeding coefficients for mature, seed and seedling populations were found to be positive with high values (mature F_{IS}=0.43; seed F_{IS} =0.46; seedling F_{IS} =0.36). These results indicate that marginally located black pine populations at all stages have been experiencing considerable inbreeding. Excess of homozygosity present in these populations may cause to deviation from HWE since homozygous trees have higher tendency to reproduce among themselves due to restricted gene flow. Moderate level of genetic differentiation for mature (F_{ST} =0.06) and high level of genetic differentiation for seed (F_{ST} =0.11) and seedling populations (F_{ST} =0.09) could be due to low level of gene flow, in turn increasing mating among relatives.

F_{ST} value is the clear indicator of the genetic differentiation among populations. The mean F_{ST} value among the three stages of four populations, varies from 0.06 in mature to 0.11 in seed stage. It can be said that genetic differentiation among populations is variable. The reasons for the mild level ($F_{ST=0.13)}$ and variability of

genetic differentiation are due to low level of gene flow among black pine populations from four locations. Geographic and genetic isolation may cause to low level of gene flow and high level of genetic differentiation among populations. As stated before, marginal populations are located on the edge of naturally distributed populations and have no geographic linkage with the core populations. These populations are geographically isolated with small population sizes. As a result of geographic isolation, gene flow among the core and marginal populations are not expected in a high level. Geographic isolation comes along with genetic isolation as well. The genetic isolation cause to reproduction of similar genotypes, and rising the frequencies of alike genes within the population. This event leads to increased inbreeding level within the marginal populations, as it is the case in the present study.

Pinus nigra evolved to maintain its genetic diversity with the mating strategies, which are wind pollination and outbreeding. These strategies ensure mating between individuals which are genetically different from each other. Furthermore, seed is the only product of sexual reproduction process, so, it carries the all genetic variation caused by both meiosis and sexual reproduction, which are main sources of genetic variation (Hamrick *et al*., 1979). Moreover, seeds themselves are not affected by environmental conditions before germination and also not eliminated by natural selection. Hence, in the light of these issues, it can be claimed that "seed" stage has a high portion of genetic variation. When comparing the results of three life stages, it is obvious that marginal populations at the seed stage have higher heterozygosity level and higher allelic richness. Even if the genetic richness present in the seed stage of the populations, it could not pass through the further life stages due to living on the edge of habitats. Despite the fact that, genetic diversity regenerated increasingly in the seed stage, as it could not be transmit to the other life stages effectively. Thus, marginal populations will be vulnerable to the changing environmental conditions in the future.

In the current study, the seeds and seedlings were faced to both natural and artificial selection before using. The reason of the artificial selection is imitating the natural selection on these individuals. By repetitive eliminations, it was tried to observe, how seedlings' genetic diversity and genetic structure were affected from selection. According to result of the seedling stage, it was observed that both mean expected

and observed heterozygosity values are lower compare to seed stage, and F_{IS} and F_{ST} values decline since seed stage of the populations carry some alleles with very low frequencies. So, elimination of some individuals bring with losing some alleles, along with heterozygosity. This situation is revealed because of the artificial and natural elimination process.

Finally, the mature stage of the four different population is evaluated. Just like seedling stage, mature stage faced with many selective forces as well. In fact, these selective forces could be both naturally and human mediated. Because of living on edge, all natural forces such as, temperature fluctuations, drought and gust; and human mediated; land clearing and using as firewood are in operation. In the light of this, it is expected to see a downfall of the genetic diversity on this stage of the populations. The findings of the current study support the expectations that low mean heterozygosity level and high differentiation level among three life stages were estimated.

In mature stage of the populations, genetic diversity parameters were calculated lower than the other two stages. This points out that, the populations experienced bottleneck and dramatic decrease of the founder population. The effective population size is small, so that mating occurred mostly between relative trees. Inbreeding is the natural consequences of the small and isolated populations and make the population more vulnerable against changing environmental conditions

When Principle Coordinate Analysis (PCoA) were performed with pairwise F_{ST} values, four populations (matureHAS, matureBEY, matureBYN, matureKDN) were completely separated, but seeds and seedling populations appear to be closely related populations. This situation possibly caused by F_{ST} values which are dissimilar from each other. In other words, having the highest F_{ST} value is resulted as genetically most distantly located, the others with lower F_{ST} are located as clusters. This was caused by the higher gene flow between individuals or populations. Although long-distance gene flow is not expected in marginal populations due to geographic isolation. As HAS and BYN populations are close to western Black-sea Anatolian Black pine populations, which are core populations, may have contributed to the gene flow. This phenomenon may lead to genetic similarities between populations at distant locations.

Revealing the phylogenetic relationship, dendrogram were drawn fallowed coancestry identities with Neighbor-joining method. The dendrogram give an information about relatedness of the populations. The findings were that populations the same clustering pattern at consistent each of three life stages. All of them shows HAS and KDN are closely related, BEY forms a group with them, and BYN populations are most distant to other three.

When Structure analysis were performed separately for all three life stages of the populations, mature stage showed difference in terms of having three, yet seed and seedling have two genetic groups. In addition, while mature populations showed complete genetic admixtures, other two life stages had some genetic admixtures as well. This situation may be caused by gene flow among core and marginal populations via long distant pollen dispersal (Ledig, 1997)

The most important feature of these populations is being marginal populations. Hence, before the conduction of the study there were some expectations about the results, especially at the mature stage. For being marginal, these populations are located edge of the distribution, geographically isolated, small population size, and high average age. Our expectations were that these populations may have high inbreeding, low observed heterozygosity and high genetic differentiation. Our findings supported these expectations. This was also supported by the findings of Pandey and Rajora (2012).

The diversity and differentiations results were found to be parallel to our expectations. Seed stage possess remarkably high genetic diversity as expected, despite challenging environmental conditions such as unequable temperature, drought, fire etc. Sustained genetic diversity at the seed stages of marginal populations could be an explanation why these marginal populations have not been extinct throughout many years experiencing with gigantic amount of selective forces. The genetic diversity probably will lead to survive these populations against future climatic and other environmental changes given that the current genetic diversity and population sizes are maintained. Therefore, seeds from marginally located Anatolian Black pine populations may be the good source for future afforestation activities in marginal lands of central Turkey.

CHAPTER 6

CONCLUSION

Anatolian black pine (*Pinus nigra* subsp. *pallasiana*) is one of the most important conifer species of Turkey in both ecological and economical values. The species natural distribution ranges from core area to the marginal habitats. Marginal populations are more vulnerable in terms of losing original genetic makeup and reduced survival rate in changing environmental and climatic conditions. Global climatic change is major threat to all species for rapid and unpredictable alterations. Marginal populations will be affected from these alterations as well.

In this thesis, 4 marginally located Anatolian black pine (*Pinus nigra* subsp. *pallasiana*) populations were studied to reveal the genetic diversity and relativeness at three different life stages (seed, seedling and mature). It was analyzed that how the natural selection and adaptation processes affected the genetic diversity of these life stages.

The results of the study demonstrated that low range of genetic diversity (He=0.44, Ho=0.22) and high level of inbreeding (F_{IS}=0.42) were observed in the four marginal populations. The results with respect to genetic diversity and inbreeding followed a similar pattern at each of the three stages. Moreover, genetic differentiation was determined as relatively high level (F_{ST}=0.13). The most diverse life stage was found to be the seed stage, which may be caused by contribution by rare and low frequency alleles and being not subjected to selective forces. It holds also the highest genetic differentiation due to inbreeding occurring in populations. The highest heterozygosity was estimated for the seed stage of Beynam population.

The results of the study were parallel to our expectation which based on the features of marginal population theory. The main consequences of the study are; (i) low level of genetic diversity, (ii) high level of inbreeding and (iii) high level of genetic differentiation. These were observed at three life stages of 4 marginal populations of *Pinus nigra* subsp. *pallasiana.*

In the light of these findings, the populations should be actively used for afforestation activities in available marginal habitats of central Turkey since these populations harbor vital genetic resources of the species, with the capability of adaptation to harsh conditions. These marginal Anatolian black pine populations are the remaining forests from the widely distributed forests in the past. Throughout the years, the populations adapted to the most extreme environmental conditions since, containing adaptive genetic diversity. They are valuable by means of adapting variable environmental conditions, including climatic as well. It is necessary that, the conservation status of these populations need to be elevated and negative anthropogenic factors need to be eliminated.

REFERENCES

Afzal- Rafii, Z. A. R. A., & Dodd, R. S. Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (Pinus nigra) in western Europe. *Molecular Ecology*, 16(4), 723-736., 2007.

Aguinagalde, I., Llorente, F., & Benito, C. Relationships among five populations of European black pine (Pinus nigra Arn.) using morphometric and isozyme markers. *Silvae Genetica*, 46(1), 1-5., 1997.

Ağaçlandırma Seferberliği Sonuç Raporu, retrieved from http://www.cem.gov.tr/erozyon/Files/yayinlar/AGM_FAALiYET_2008_2009_2010_2011.pdf, 2012. (in Turkish)

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis- McLane, S. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, *1*(1), 95-111, 2008.

Akkemik, Ü., Çinar Yilmaz, H., Kaya, A., Oral, D. Some Taxonomic Notes on *Pinus* L. (Pinaceae) Taxa Naturally Grow in Turkey. *Istanbul Üniversitesi Orman Fakültesi Dergisi*, 61(1), 63-78, 2010. (in Turkish)

Akkemik, Ü. (Editor). Türkiye'nin Doğal-Egzotik Ağaç ve Çalıları I. İ. Ü. Orman Genel Müdürlüğü Yayınlar, Ankara. 736 Pages, 2014. (in Turkish)

Alptekin, C. Ü. Anadolu karaçamı (*Pinus nigra* ssp. *pallasiana* Lamb*). İstanbul Üniversitesi, Orman Fakültesi Dergisi*, 36: 132- 154, 1986. (in Turkish)

Arbez, M., Bernard-Dagan, C., Filion, C. Intraspecific Variability of *Pinus nigra* Monoterpenes-Analyses of First Results. Ann. *Sci.For*. 31:57-70, 1974.

Bonnet-Masimbert, M., & Bikay-Bikay, V.. Variabilite intranspecifique des isozymes de la glutamate-oxaloacetate-transaminase chez Pinus nigru Arnold. Inter&t pour la taxonomie des sous especes. *Silvae Genetica*, 27(2), 71-79., 1978.

Buschbom, J., Yanbaev, Y., & Degen, B. Efficient long-distance gene flow into an isolated relict oak stand. *Journal of Heredity*, *102*(4), 464-472, 2011.

Cengel, B., Velioglu, E., Tolun, A.A,, Kaya, Z. Pattern and magnitude of genetic diversity in *Pinus nigra* Arnold subspecies *pallasiana* populations from Kazdağı: implications for in situ conservation. *Silvae Genetica* 49(6):249–256, 2000.

Cengel, B., Tayanc¸ Y., Kandemir, G., Velioglu, E., Alan, M., Kaya, Z. Magnitude and efficiency of genetic diversity captured from seed stands of *Pinus*

nigra (Arnold) subsp. *pallasiana* in established seed orchards and plantations. *New Forests* 43:303–317, 2012.

Davis, P. H. 1988: Flora of Turkey and the east Aegean Islands vols. 1-10. *Edinburgh: Edinburgh UniversityPress*, 1965.

Debreczy, Z. And Racz, I. Conifers Around the World: Conifers of the Temperate Zones and Adjacent Regions. Dendro Press, 1089 pages, 2012.

Delgado, P., Pin~ ero, D., Chaos, A., Pe´ rez- Nasser, N., & Alvarez- Buylla, E. R. . High population differentiation and genetic variation in the endangered Mexican pine Pinus rzedowskii (Pinaceae). *American Journal of Botany*, *86*(5), 669-676.,1999.

Doğan, B., Özer, A.S. Gülbaba, A.G. Velioğlu,E. Doerksen, A.H., Adams, W.T. Kazdağlarından örneklenen Karaçam (Pinus nigra Arnold.) populasyonlarında kalıtım ve allellerin bağlılığı. EOAED Vol:1, 1997.(in Turkish)

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

Ducci, F., Garavaglia, V., and Monteverdi, M.C. Conserving marginal populations of forest trees in Euprope. Unasylva 242, Vol 65: 51-55, 2014.

Echeverría, C., Newton, A. C., Lara, A., Benayas, J. M. R., & Coomes, D. A. Impacts of forest fragmentation on species composition and forest structure in the temperate landscape of southern Chile. *Global Ecology and Biogeography*, *16*(4), 426-439, 2007.

Economou, A. . Growth intercept as an indicator of site quality for planted and natural stands of *Pinus nigra* var. pallasiana in Greece. *Forest Ecology and Management*, 32(2-4), 103-115., 1990.

Enescu, C. M., de Rigo, D., Caudullo, G., Mauri, A., Houston Durrant, T.. *Pinus nigra* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), *European Atlas of Forest Tree Species*. Publ. Off. EU, Luxembourg, pp. e015138+,2016.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, *10*(3), 564-567.

Evanno, G., Regnaut, S., & Goudet, J. . Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*(8), 2611-2620., 2005.

Faaliyet Raporu ,Retreived from" http://www.dsi.gov.tr/docs/stratejik-plan/dsi-2013-faaliyet-raporu.pdf?sfvrsn=2", 2013(in Turkish)

Farjon, A. A Handbook of the World's Conifers (2 vols.): Revised and Updated Edition. Brills, pg 1112., 2010.

Finesci-ii, S. Determination of the Origin of an Isolated Group of Trees of Pinus nigra through Enzyme Gene Markers. *Silvae Genetica*, *33*, 4-5, 1984.

Francis, R. M. . POPHELPER: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*(1), 27-32., 2017.

Fukarek, P. Notes on P. nigra sl. *Radovi Poljoprivrednog Sumarskog Fakulteta Univerziteta u Sarajevu*, (3), 3-92, 1958

Giovannelli, G., Roig, A., Spanu, I., Vendramin, G. G., & Fady, B. . A new set of nuclear microsatellites for an ecologically and economically important conifer: the European black pine (Pinus nigra Arn.). *Plant Molecular Biology Reporter*, 35(3), 379-388., 2017.

González-Martínez, S.C., Robledo-Arnuncio, J.J., Collada, C., Díaz,A., Williams, C. G., Alía, R., Cervera, M.T. "Cross-amplification and sequence variation of microsatellite loci in Eurasian hard pines" *Theoretical and Applied Genetics*, 109(1), 103-111., 2004.

GÖKER, Y. Dursunbey ve elekdağ karaçamlari (pinus nigra var: pallasina)'nin fiziksel, mekanik özellikleri ve kullaniş yerleri hakkinda araştirmalar. *Journal of the Faculty of Forestry Istanbul University| İstanbul Üniversitesi Orman Fakültesi Dergisi*,1969. (in Turkish)

Hamrick, J. L., Linhart, Y. B., & Mitton, J. B. (1979). Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. *Annual review of ecology and systematics*, *10*(1), 173-200. ,1979.

Hamrick, J. L., & Godt, M. W. Allozyme diversity in plant species. *Plant population genetics, breeding, and genetic resources.* 43-63,1990.

Isajev, V., Fady, B., Semerci, H., & Andonovski, V. *EUFORGEN Technical Guidelines for genetic conservation and use for European black pine (Pinus nigra). Bioversity International,2004.*

Işik, K. 1990: Seasonal Course of Height and Needle Growth in *Pinus nigra* Grown in Summer-Dry Central Anatolia. *Forest Ecol. Man*. 35: 261-270, 1990.

Jakobsson, M., & Rosenberg, N. A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801-1806., 2007.

Jaramillo- Correa, J. P., Grivet, D., Terrab, A., Kurt, Y., De- Lucas, A. I., Wahid, N., ... & González- Martínez, S. C. The Strait of Gibraltar as a major biogeographic barrier in Mediterranean conifers: a comparative phylogeographic survey. *Molecular Ecology*, *19*(24), 5452-5468, 2010.

Kaya, Z., Ching, K.K., Ve Stafford, S.G. 1985: A Statistical Analysis of European Black Pine (*Pinus* nigra Arnold.) from Different Sources, *Silvae Genetica* 34:148-156,1985.

Kaya, Z., and Neale, D.B. Random Amplified Polymorphic DNA (RAPD) Polymorphısms in *Pinus* nigra subsp. *pallasiana* and *Pinus* brutia. Doğa-Tr *J. of Agriculture and Forestry* 17 pp. 295- 306, 1993.

Kaya, Z. & Temerit, A. Genetic Structure of Marginally Located *Pinus nigra* var *pallasiana* Populations in Central Turkey. *Silvae Genetica* 43, 5/6 pp 272-275,1994.

Koski, V., Antola, J. National tree breeding and seed production programme for Turkey: 1994 - 2003. Prepared in cooperation with Research Directorate of Forest Tree Seeds and Tree Breeding. ENSO Forest Development OY LTD, 1993.

Krakowski, J., Aitken, S. N., & El-Kassaby, Y. A. (2003). Inbreeding and conservation genetics in whitebark pine. *Conservation Genetics*, *4*(5), 581-593., 2003.

Ledig, F. T. (1998). Genetic diversity in tree species: with special reference to conservation in Turkey and the eastern Mediterranean. In *The proc. of international symposium on in situ conservation of plant genetic diversity, Central research institute for field crops,* Ankara, Turkey (pp. 231-248), 1998.

Lewontin, R.C. The genetic basis of evolutionary change.,Columbia University Press, New York City, New York. 346., 1974,

Linares, J. C., & Tíscar, P. A. Climate change impacts and vulnerability of the southern populations of Pinus nigra subsp. salzmannii. *Tree physiology*, *30*(7), 795-806, 2010.

Lindner, M., Maroschek, M., Netherer, S., Kremer, A., Barbati, A., Garcia-Gonzalo, J., ... & Lexer, M. J. Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. *Forest ecology and management*, *259*(4), 698-709, 2010.

Matziris, D. I. Variation in Growth and Branching Characters in Black Pine (Pinus nigra Arnold) of Peloponnesos. *Silvae Genetica*, *38*(3-4), 77-81, 1989.

Mayr, E. Populations, species, and evolution, Harvard University Press, Cambridge, Massachusetts,453, 1970.

Naydenov, K. D., Tremblay, F. M., Fenton, N. J., & Alexandrov, A. . Structure of Pinus nigra Arn. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: provenance tests. *Biochemical Systematics and Ecology*, 34(7), 562-574., 2006.

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

Nei, M. F- statistics and analysis of gene diversity in subdivided populations. *Annals of human genetics*, *41*(2), 225-233, 1977.

Nikolić, D., & Tucić, N. Isoenzyme variation within and among populations of European black pine (Pinus nigra Arnold). *Silvae Genetica,* 32(3-4), 80-89, 1983.

Ormanİstatistikleri,Retreivedfrom https://www.ogm.gov.tr/ekutuphane/Sayfalar/Istatistikler.aspx, 2015. (in Turkish)

Ortego, J., Bonal, R., & Muñoz, A. Genetic consequences of habitat fragmentation in long-lived tree species: the case of the Mediterranean holm oak (Quercus ilex, L.). *Journal of Heredity*, *101*(6), 717-726, 2010.

Pandey, M., & Rajora, O. P. Genetic diversity and differentiation of core vs. peripheral populations of eastern white cedar, Thuja occidentalis (Cupressaceae). *American Journal of Botany*, *99*(4), 690-699, 2012.

Pandey, M., & Rajora, O. P. Higher fine-scale genetic structure in peripheral than in core populations of a long-lived and mixed-mating conifer-eastern white cedar (Thuja occidentalis L.). *BMC Evolutionary Biology*, *12*(1), 48, 2012.

Parida, S. K., Dalal, V., Singh, A. K., Singh, N. K., & Mohapatra, T. Genic non-coding microsatellites in the rice genome: characterization, marker design and use in assessing genetic and evolutionary relationships among domesticated groups. *BMC genomics*, *10*(1), 140, 2009.

Pavia, I., Carvalho, A., Rocha, L., Gaspar, M. J., & Lima-Brito, J. Physical location of SSR regions and cytogenetic instabilities in Pinus sylvestris chromosomes revealed by ND-FISH. *Journal of genetics*, *93*(2), 567-571, 2014.

PE, P. R. S. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, *28*, 2537-2539, 2012.

Portfaix, C. Exploration of Genetic-Variability of 5 Natural Stands of Corsican Pine (*Pinus nigra* ssp. *laricio* var *corsicana*) Ann. Sci. For. 46:217-232, 1989.

Pritchard, J. K., Stephens, M., & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959, 2000.

Rajora, O. P., Rahman, M. H., Buchert, G. P., & Dancik, B. P. Microsatellite DNA analysis of genetic effects of harvesting in old- growth eastern white pine (Pinus strobus) in Ontario, Canada. *Molecular ecology*, *9*(3), 339-348, 2000.

Ruiz- Labourdette, D., Nogués- Bravo, D., Ollero, H. S., Schmitz, M. F., & Pineda, F. D. Forest composition in Mediterranean mountains is projected to shift along the entire elevational gradient under climate change. *Journal of Biogeography*, *39*(1), 162-176, 2012.

Scaltsoyiannes, A., Rohr, R., Panetsos, K. P., & Tsaktsira, M. Allozyme frequency distributions in five European populations of black pine (Pinus nigra Arnold).(I) Estimation of genetic variation within and among populations.(II) Contribution of isozyme analysis to the taxonomic status of the species. *Silvae genetica*, 43(1), 20-30, 1994.

Scaltsoyiannes, A., Rohr, R., Panetsos, K. P., & Tsaktsira, M. Allozyme frequency distributions in five European populations of black pine (Pinus nigra

Arnold).(I) Estimation of genetic variation within and among populations.(II) Contribution of isozyme analysis to the taxonomic status of the species. *Silvae genetica*, 43(1), 20-30, 2000.

Sebastiani, F., Pinzauti, F., Kujala, S. T., González-Martínez, S. C., & Vendramin, G. G. Novel polymorphic nuclear microsatellite markers for Pinus sylvestris L. *Conservation Genetics Resources*, 4(2), 231-234, 2012.

Silin, A. E. Ve Goncharenko, G.G. Allozyme variation in natural populations of Eurasian pines: IV. Population structure and genetic variation in geographically related and isolated populations of *Pinus nigra* Arnold on the Crimean Peninsula. *Silvae Genetica*, 45, 67-75, 1996.

Şimşek, Y., Erkuloğlu, Ö.S. Ve Tosun, S. Türkiyede karaçam (*Pinus nigra* Arn. ssp. *pallasiana* (Lamb) Holmboe) Orijin Denemelerinin İlk Sonuçları, OAE Teknik Bülten No: 247, 1995. (in Turkish)

Tolun, A.A, Velioglu, E., Cengel, B., Kaya, Z. Genetic structure of black pine (*Pinus nigra* Arn. subsp. *pallasiana*) populations sampled from the Bolkar Mountains. *Silvae Genetica* 49(3):113–119, 2000.

Tsoumis, G. T. *The depletion of forests in the Mediterranean region: a historical review from ancient times to the present*. Aristotelian University of Thessaloniki,1988.

Üçler, A. Ö., & Gülcü, S. A Study on the Variations of Cone and Seed Morphology of Some Natural Anatolian Black Pine (Pinus nigra Arnold. subsp. pallasiana Lamb. Holmboe) Populations in Isparta Lake District. In *Proceedings of the 1st International Symposium on Protection of Natural Environment and Ehrami Karacam* (pp. 332-340), 1999.

Velioğlu, E., Çengel, B., İçgen, Y., Kandemir, G., Kaya, Z. Determination of Genetic Diversity of Anatolian Black pine (*Pinus nigra* Arnold) Populations in Kırklareli-Kasatura Bay Nature Conservation Area by Molecular Markers. Forest Tree Seeds and Tree Breeding Research Institute Directorate, Technical Bulletin No: 33, Ankara-TURKEY, 2004.

Velioğlu, E., Çengel, B., Kaya, Z. Kazdağları Doğal Karaçam (*Pinus nigra* Arnold subspecies *pallasiana*) populasyonlarında genetik çeşitliliğin yapılanması, Forest Tree Seeds and Tree Breeding Research Institute Directorate, Technical Bulletin No: 1, Ankara-TURKEY, 1999. (in Turkish)

Velioğlu, E., Kandemir, G., Tayanç, Y., Çengel, B., Alan, M., Kaya, Z. Determination of Genetic Diversity and Gene Conservation Strategies for Oriental Sweetgum (*Liquidambar orientalis* Miller) Populations in Turkey by Molecular Markers. Forest Tree Seeds and Tree Breeding Research Institute Directorate, Technical Bulletin No: 20, Ankara-TURKEY, 2008.

Vidakovic, M. Genetics of European Black Pine (Pinus nigra Arn.). In *Annales forestales* Vol. 6, No. 3, pp. 57-86, 1974.

Vidaković, M. *Conifers: morphology and variation*. Grafičko Zavod Hrvatske, 1991.

Vieira MLC, Santini L, Diniz AL, Munhoz CF , Microsatellite markers: what they mean and why they are so useful. Genet. Mol. Biol. 39(3):312-328, 2016.

Wertime, T. A. The furnace versus the goat: the pyrotechnologic industries and Mediterranean deforestation in antiquity. *Journal of Field Archaeology*, 445-452, 1983.

Yaltirik, F. Dendroloji Ders Kitabı. İstanbul Üniversitesi Orman Fakültesi Yayınları, 1988. (in Turkish)

Yeh, F. C., & Boyle, T. POPGENE version 1.2, Microsoft Windows-based software for population genetics analysis. *University of Alberta, Alberta, Canada*, 1997.

APPENDICES

APPENDIX A

BUFFERS CHEMICALS AND EQUIPMENTS

Buffers and solutions for DNA isolation

2X CTAB: 2 gr CTAB (Cetyl Trimethyl Ammonium Bromide), (SIGMA)

4 ml (pH:8) 0.5 M EDTA, (FLUKA)

10 ml (pH:8) Tris HCL, (AppliChem)

28 ml NaCl is completed with 100 mL distilled water

Chloroform isoamil alcohol, (FLUKA) : (24/1)

β mercapto ethanol, (SIGMA) : 17,5 ml β mercapto ethanol is completed with 250 ml with distilled water

Isopropanol, (FLUKA): Pure Isopropanol, ice cold

Tris buffer pH:8(1M): AppliChem

DNA Quantification: BioDrop μlite

Buffers and solutions for PCR

Sterile water

5x HOT FIREPol® Blend Master Mix;

- HOT FIREPol® DNA polymerase
- Proofreading enzyme
- 5x blend Master Mix Buffer
- 15mM $MgCl_2$
- 2mM dNTPs of each
- BSA
- Blue dye
- Yellow dye
- Compound that increases sample density for direct loading

DNA: 20ng/ μl

Primer Pairs: 10μM

Agarose Gel Electrophoresis Buffers and Gel System

1x TAE ((AppliChem 50X)

Agarose:

Running buffer: 1x TAE ((AppliChem 50X)

Thermo GeneRuler 50bp

Bioshop Bio-View White

Equipments

Autoclave: Nüve OT-90L

Centrifuge: Sigma

Electrophoresis System: Thermo Scientific

Thermocyclers: Eppendorf- Mastercycler

Deepfreezer: Vestel- Freezer

Magnetic Stirrer: JKI

Refrigerator: Arçelik

UV Transilluminator: Vilbor Lourmant

Vortex: Labnet International INC

Water Bath: Nüve

Micropipettes: Thermo

APPENDIX B

ASSAY PROCEDURE WAS DONE BY THE BM LABOSIS (Çankaya, Ankara)

1. Registration of pcr product (by Customer)

2. pcr product + Hi-Di formamide + size standard * size standard type : 120LIZ, 350ROX, 400HD, 500LIZ, 600LIZ, 1200LIZ

3. denaturation

4. 3730xl running by using Dye set : DS-30 set for internal standard size marker 400HD , DS-33 set for internal standard size marker 400HD

5. Genemapper v.5 analysis

APPENDIX C

A PART OF EXCEL MATRIX SHOWING GENOTYPES ALLELE SIZES

Q	R	S	T	U	V	W
CK 16						
	94	97	103			
fKDN-1.1-C			1	103/103		
fKDN-1.2-C	1		1	94/103		
fKDN-1.3-C			1	103/103		
fKDN-1.4-C			1	103/103		
fKDN-2.1-C			1	103/103		
fKDN-2.2-C			1	103/103		
fKDN-2.3-C			1	103/103		
fKDN-2.4-C			1	103/103		
fKDN-3.1-C			1	103/103		
fKDN-3.2-C			1	103/103		
fKDN-3.3-C	1		1	94/103		
fKDN-3.4-C			1	103/103		
fKDN-4.1-C	1		1	93/103		
fKDN-4.2-C			1	103/103		
fKDN-4.3-C			1	103/103		
fKDN-4.4-C			1	103/103		
fKDN-6.1-C		1	1	97/103		
fKDN-6.2-C			1	103/103		
fKDN-6.3-C			1	103/103		
fKDN-6.4-C			1	103/103		
fKDN-8.1-C		1	1	97/103		

EY-C-raw | tBYN-C | tKDN-C-raw | fHAS-C-raw | fBEY-C-raw | fBYN-C-raw | **fKDN-C-raw**

APPENDIX D

EXAMPLES OF DATA FILE FORMATS

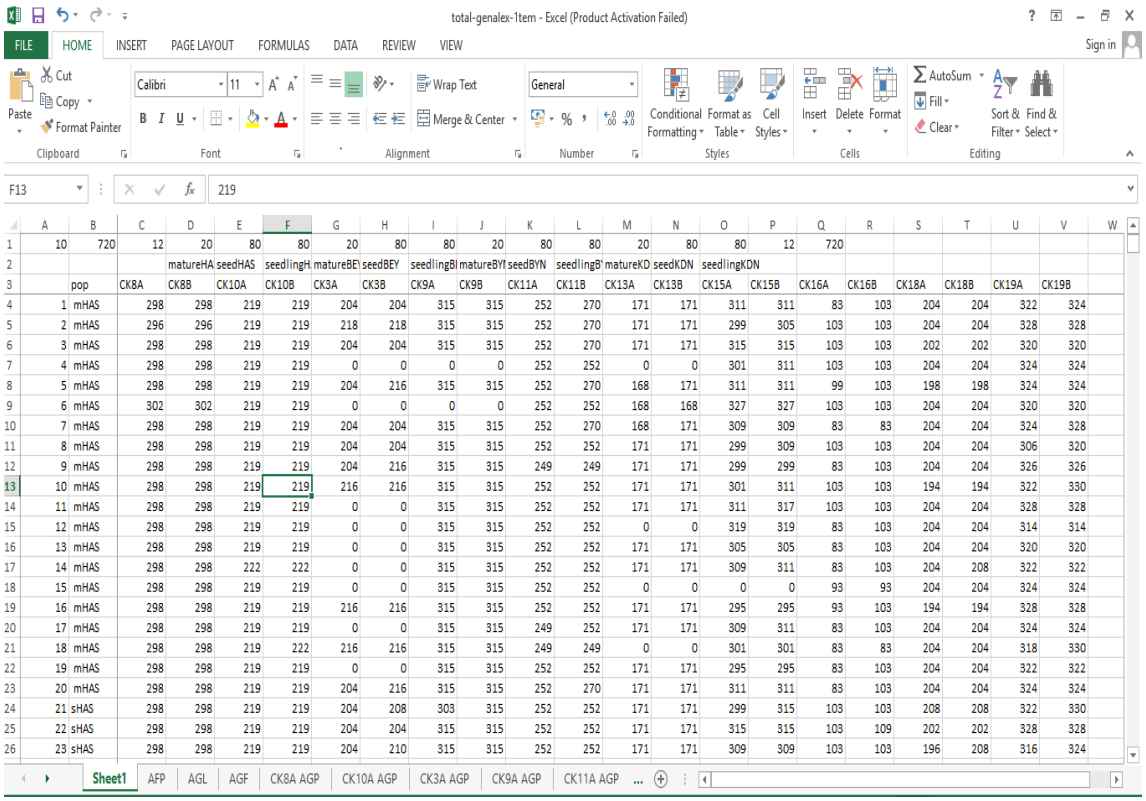
Genepop data format

PopulationData-27haz - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

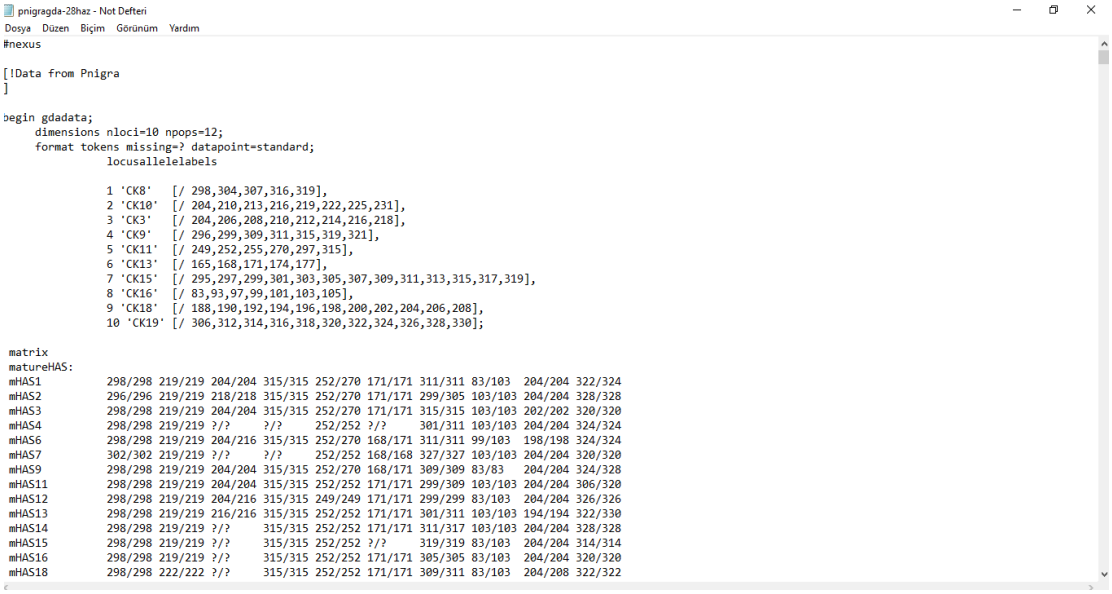
```
Title line:  "PopulationData.txt"
CK8
CK10
CK3
CK9
CK11
CK13
CK15
CK16
CK18
CK19
Pop
1_      ,       298298 219219 204204 315315 252270 171171 311311 083103 204204 322324
1_      ,       296296 219219 218218 315315 252270 171171 299305 103103 204204 328328
1_      ,       298298 219219 204204 315315 252270 171171 315315 103103 202202 320320
1_      ,       298298 219219 000000 000000 252252 000000 301311 103103 204204 324324
1_      ,       298298 219219 204216 315315 252270 168171 311311 099103 198198 324324
1_      ,       302302 219219 000000 000000 252252 168168 327327 103103 204204 320320
1_      ,       298298 219219 204204 315315 252270 168171 309309 083083 204204 324328
1_      ,       298298 219219 204204 315315 252252 171171 299309 103103 204204 306320
1_      ,       298298 219219 204216 315315 249249 171171 299299 083103 204204 326326
1_      ,       298298 219219 216216 315315 252252 171171 301311 103103 194194 322330
1_      ,       298298 219219 000000 315315 252252 171171 311317 103103 204204 328328
1_      ,       298298 219219 000000 315315 252252 000000 319319 083103 204204 314314
1_      ,       298298 219219 000000 315315 252252 171171 305305 083103 204204 320320
1_      ,       298298 222222 000000 315315 252252 171171 309311 083103 204208 322322
1_      ,       298298 219219 000000 315315 252252 000000 000000 093093 204204 324324
1_      ,       298298 219219 216216 315315 252252 171171 295295 093103 194194 328328
1_      ,       298298 219219 000000 315315 249252 171171 309311 083103 204204 324324
1_      ,       298298 219222 216216 315315 249249 000000 301301 083083 204204 318330
1_      ,       298298 219219 000000 315315 252252 171171 295295 083103 204204 322322
1_      ,       298298 219219 204216 315315 252270 171171 311311 083103 204204 324324
Pop
2_      ,       298298 219219 204208 303315 252252 171171 299315 103103 208208 322330
2_      ,       298298 219219 204204 315315 252252 171171 315315 103109 202202 328328
2_      ,       298298 219219 204210 315315 252252 171171 309309 103103 196208 316324
2_      ,       298298 219219 204212 315315 252252 171171 299299 103103 198204 316326
```

GenAIEx data format



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	10	720	12	20	80	80	20	80	80	20	80	80	20	80	80	12	720						
2				matureHA	seedHAS	seedlingH	matureBE	seedBEY	seedlingBI	matureBY	seedBYN	seedlingB	matureKD	seedKDN	seedlingKDN								
3		pop	CK8A	CK8B	CK10A	CK10B	CK3A	CK3B	CK9A	CK9B	CK11A	CK11B	CK13A	CK13B	CK15A	CK15B	CK16A	CK16B	CK18A	CK18B	CK19A	CK19B	
4	1	mHAS	298	298	219	219	204	204	315	315	252	270	171	171	311	311	83	103	204	204	322	324	
5	2	mHAS	296	296	219	219	218	218	315	315	252	270	171	171	299	305	103	103	204	204	328	328	
6	3	mHAS	298	298	219	219	204	204	315	315	252	270	171	171	315	315	103	103	202	202	320	320	
7	4	mHAS	298	298	219	219	0	0	0	0	252	252	0	0	301	311	103	103	204	204	324	324	
8	5	mHAS	298	298	219	219	204	216	315	315	252	270	168	171	311	311	99	103	198	198	324	324	
9	6	mHAS	302	302	219	219	0	0	0	0	252	252	168	168	327	327	103	103	204	204	320	320	
10	7	mHAS	298	298	219	219	204	204	315	315	252	270	168	171	309	309	83	83	204	204	324	328	
11	8	mHAS	298	298	219	219	204	204	315	315	252	252	171	171	299	309	103	103	204	204	306	320	
12	9	mHAS	298	298	219	219	204	216	315	315	249	249	171	171	299	299	83	103	204	204	326	326	
13	10	mHAS	298	298	219	219	216	216	315	315	252	252	171	171	301	311	103	103	194	194	322	330	
14	11	mHAS	298	298	219	219	0	0	315	315	252	252	171	171	311	317	103	103	204	204	328	328	
15	12	mHAS	298	298	219	219	0	0	315	315	252	252	0	0	319	319	83	103	204	204	314	314	
16	13	mHAS	298	298	219	219	0	0	315	315	252	252	171	171	305	305	83	103	204	204	320	320	
17	14	mHAS	298	298	222	222	0	0	315	315	252	252	171	171	309	311	83	103	204	208	322	322	
18	15	mHAS	298	298	219	219	0	0	315	315	252	252	0	0	0	0	93	93	204	204	324	324	
19	16	mHAS	298	298	219	219	216	216	315	315	252	252	171	171	295	295	93	103	194	194	328	328	
20	17	mHAS	298	298	219	219	0	0	315	315	249	252	171	171	309	311	83	103	204	204	324	324	
21	18	mHAS	298	298	219	222	216	216	315	315	249	249	0	0	301	301	83	83	204	204	318	330	
22	19	mHAS	298	298	219	219	0	0	315	315	252	252	171	171	295	295	83	103	204	204	322	322	
23	20	mHAS	298	298	219	219	204	216	315	315	252	270	171	171	311	311	83	103	204	204	324	324	
24	21	sHAS	298	298	219	219	204	208	303	315	252	252	171	171	299	315	103	103	208	208	322	330	
25	22	sHAS	298	298	219	219	204	204	315	315	252	252	171	171	315	315	103	109	202	202	328	328	
26	23	sHAS	298	298	219	219	204	210	315	315	252	252	171	171	309	309	103	103	196	208	316	324	

GDA data format



```
#nexus

[!Data from Pnigra
]

begin gdadata;
     dimensions nloci=10 npops=12;
     format tokens missing=? datapoint=standard;
               locusallelelabels

               1 'CK8'   [/ 298,304,307,316,319],
               2 'CK10'  [/ 204,210,213,216,219,222,225,231],
               3 'CK3'   [/ 204,206,208,210,212,214,216,218],
               4 'CK9'   [/ 296,299,309,311,315,319,321],
               5 'CK11'  [/ 249,252,255,270,297,315],
               6 'CK13'  [/ 165,168,171,174,177],
               7 'CK15'  [/ 295,297,299,301,303,305,307,309,311,313,315,317,319],
               8 'CK16'  [/ 83,93,97,99,101,103,105],
               9 'CK18'  [/ 188,190,192,194,196,198,200,202,204,206,208],
               10 'CK19' [/ 306,312,314,316,318,320,322,324,326,328,330];

 matrix
 matureHAS:
 mHAS1          298/298 219/219 204/204 315/315 252/270 171/171 311/311 83/103  204/204 322/324
 mHAS2          296/296 219/219 218/218 315/315 252/270 171/171 299/305 103/103 204/204 328/328
 mHAS3          298/298 219/219 204/204 315/315 252/270 171/171 315/315 103/103 202/202 320/320
 mHAS4          298/298 219/219 ?/?     ?/?     252/252 ?/?     301/311 103/103 204/204 324/324
 mHAS6          298/298 219/219 204/216 315/315 252/270 168/171 311/311 99/103  198/198 324/324
 mHAS7          302/302 219/219 ?/?     ?/?     252/252 168/168 327/327 103/103 204/204 320/320
 mHAS9          298/298 219/219 204/204 315/315 252/270 168/171 309/309 83/83   204/204 324/328
 mHAS11         298/298 219/219 204/204 315/315 252/252 171/171 299/309 103/103 204/204 306/320
 mHAS12         298/298 219/219 204/216 315/315 249/249 171/171 299/299 83/103  204/204 326/326
 mHAS13         298/298 219/219 216/216 315/315 252/252 171/171 301/311 103/103 194/194 322/330
 mHAS14         298/298 219/219 ?/?     315/315 252/252 171/171 311/317 103/103 204/204 328/328
 mHAS15         298/298 219/219 ?/?     315/315 252/252 ?/?     319/319 83/103  204/204 314/314
 mHAS16         298/298 219/219 ?/?     315/315 252/252 171/171 305/305 83/103  204/204 320/320
 mHAS18         298/298 222/222 ?/?     315/315 252/252 171/171 309/311 83/103  204/208 322/322
```

Structure data format

mature-structure-28haz - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

		CK8	CK10	CK3	CK9	CK11	CK13	CK15	CK16	CK18	CK19
1	1	298	219	204	315	252	171	311	083	204	322
1	1	298	219	204	315	270	171	311	103	204	324
2	1	296	219	218	315	252	171	299	103	204	328
2	1	296	219	218	315	270	171	305	103	204	328
3	1	298	219	204	315	252	171	315	103	202	320
3	1	298	219	204	315	270	171	315	103	202	320
4	1	298	219	-9	-9	252	-9	301	103	204	324
4	1	298	219	-9	-9	252	-9	311	103	204	324
5	1	298	219	204	315	252	168	311	099	198	324
5	1	298	219	216	315	270	171	311	103	198	324
6	1	302	219	-9	-9	252	168	327	103	204	320
6	1	302	219	-9	-9	252	168	327	103	204	320
7	1	298	219	204	315	252	168	309	083	204	324
7	1	298	219	204	315	270	171	309	083	204	328
8	1	298	219	204	315	252	171	299	103	204	306
8	1	298	219	204	315	252	171	309	103	204	320
9	1	298	219	204	315	249	171	299	083	204	326
9	1	298	219	216	315	249	171	299	103	204	326
10	1	298	219	216	315	252	171	301	103	194	322
10	1	298	219	216	315	252	171	311	103	194	330
11	1	298	219	-9	315	252	171	311	103	204	328
11	1	298	219	-9	315	252	171	317	103	204	328
12	1	298	219	-9	315	252	-9	319	083	204	314
12	1	298	219	-9	315	252	-9	319	103	204	314
13	1	298	219	-9	315	252	171	305	083	204	320
13	1	298	219	-9	315	252	171	305	103	204	320
14	1	298	222	-9	315	252	171	309	083	204	322
14	1	298	222	-9	315	252	171	311	103	208	322
15	1	298	219	-9	315	252	-9	-9	093	204	324
15	1	298	219	-9	315	252	-9	-9	093	204	324
16	1	298	219	216	315	252	171	295	093	194	328
16	1	298	219	216	315	252	171	295	103	194	328
17	1	298	219	-9	315	249	171	309	083	204	324
17	1	298	219	-9	315	252	171	311	103	204	324
18	1	298	219	216	315	249	-9	301	083	204	318
18	1	298	222	216	315	249	-9	301	083	204	330

APPENDIX E

A SUMMARY OF THE STATSTICS USED IN THESIS

The statistics used in GenAlEx 6.5

Number of different alleles (Na)

Determined by direct count. GenAlEx also provides the arithmetic mean across loci.

Effective number of alleles (Ne)

Ne represents an estimate of the number of equally frequent alleles in an ideal population. Ne enables meaningful comparisons of allelic diversity across loci with diverse allele frequency distributions. The formula is as follows;

$$Ne = \frac{1}{1-He}$$

Ne via *Frequency* is calculated by locus from He for each population.

No. of private alleles

Equivalent to the number of alleles unique to a single population in the data set.

Expected heterozygosity (He)

He is the Expected Heterozygosity or Genetic Diversity within a population. Calculated per locus as 1 minus the sum of the squared allele frequencies, pi^2. The formula is as follows;

$$He = 1 - \sum pi^2$$

Expected Heterozygosity averaged across populations (Mean He)

The average *He* or genetic diversity per population, also called *Hs* and used in the calculation of *F*- statistics. Where *Hes* is the expected heterozygosity in the *s*-th population; *k* is the number of populations. The formula is as follows;

$$Hs = \overline{He} = \frac{\sum H_{Es}}{k}$$

Observed Heterozygosity (Ho)

Observed heterozygosity for a single locus within a population, where the number of heterozygotes is determined by direct count, N = sample size. The formula is as follows;

$$Ho = \frac{No.\_of\_Hets}{N}$$

Observed heterozygosity, averaged across populations (Mean Ho)

The average observed heterozygosity of a collection of populations, used in the calculation of F-statistics. Here, Hos is the observed heterozygosity in the s-th population; k is the number of populations. The formula is as follows;

$$\boldsymbol{\overline{H}_o = \frac{\sum H_{Os}}{k}}$$

Fixation Index (F)

Calculated on a per locus basis. GenAlEx also provides the arithmetic mean across loci. Values close to zero are expected under random mating, while substantial positive values indicate inbreeding or undetected null alleles. Negative values indicate excess of heterozygosity, due to negative assortative mating, or heterotic selection. The formula is as follows;

$$F = \frac{H_e - H_o}{H_e}$$

F statistics

F statistics show the statistically expected level of heterozygosity in a population. It measures the degree of a reduction in heterozygosity relative to Hardy-Weinberg expectations (Allendorf and Luikart, 2007).

There are three indices used in F statistics (Wright, 1951);

F_{IS}, the measure of reduction of heterozygosity of an individual due to nonrandom mating within subpopulations. The formula is;

$$F_{IS} = 1 - \left(\frac{H_I}{H_S}\right)$$

F_{IT}, the measure of reduction of heterozygosity of an individual in relation to the total population and the formula is;

$$F_{IT} = 1 - \left(\frac{H_I}{H_T}\right)$$

F_{ST}, (the inbreeding coefficient within subpopulations, relative to the total) the degree of reduction in heterozygosity of a subpopulation due to genetic drift. F_{ST} provides a measure of the genetic differentiation among populations. That is, the proportion of the total genetic divergence that separates the populations. F_{ST} is typically greater than or equal to zero (but can be slightly negative). If all subpopulations are in Hardy-Weinberg equilibrium with the same allele frequencies, $F_{ST} \approx 0$. (Note that the s used for subpopulations in the notation for F_{ST}). The formula is as follows;

$$F_{ST} = 1 - \left(\frac{H_S}{H_T}\right)$$

The genetic structure of populations can be estimated by the formula

$$(1 - F_{IT}) = (1 - F_{IS})(1 - F_{ST})$$

where H_I represents the average observed heterozygosity in subpopulations, H_S represents the average expected heterozygosity estimated from each subpopulation and H_T represents the total gene diversity or expected heterozygosity in the total population as estimated from the pooled allele frequencies (Allendorf and Luikart, 2007).

Number of Migrants (Nm)

Where F_{ST} represents the degree of population genetic differentiation.

$$Nm = \frac{\left[\left(\frac{1}{F_{ST}}\right)-1\right]}{4}$$

Nei's Genetic Distance

Nei's genetic distance D, where I is Nei's Genetic Identity.

$$D = -\ln(I)$$

Nei's Genetic Identity

p_{ix} and p_{iy} are the frequencies of the i^{th} allele in populations x and y. For multiple loci, J_{xy}, J_x and J_y are calculated by summing over all loci and alleles and dividing by the number of loci. These average values are then used to calculate I**.** (Nei, 1972;1978)

$$I = \frac{J_{xy}}{\sqrt{(J_x J_y)}} \ , J_{xy} = \sum_{i=1}^{k} P_{ix}P_{iy} \ , J_x = \sum_{i=1}^{k} P_{ix}^2 \ , \ J_y = \sum_{i=1}^{k} P_{iy}^2$$

Standard Error

SE is the standard error of the mean and is widely reported when the arithmetic mean and other summary statistics are reported. Where s is the standard deviation and n is the sample size.

$$SE = \frac{s}{\sqrt{n}}$$

Statistics used to select K

To evaluate the K, ad hoc quantity (ΔK) was calculated. When ΔK shows the top peak it means the true value of K is obtained (Evanno *et al.,* 2005).

Ln P(D) shows the log likelihood for each K in STRUCTURE output obtained by first computing the log likelihood of the data at each step of the Markov Chain Monte Carlo(MCMC). The LnP(K) gives the mean likelihood over 10 runs for each K, the average of 10 values of Ln P(D). Steps for evaluating ΔK were as below;

The mean difference between successive likelihood values of K was plotted;

$$Ln'(K)=LnP(K)-LnP(K-1)$$

First the difference between values L′(K) were calculated and then, absolute value was taken;

$$|\ Ln''(K)=Ln'(K+1)-Ln'(K)\ |$$

Estimated ΔK as the mean of the absolute values of Ln″(K) was averaged over 10 runs and divided by the standard deviation of LnP(K);

$$\Delta K=\ (m|Ln''(K)|)/(s|LnP(K)|\ )$$