GENETIC STRUCTURE OF *POPULUS EUPHRATICA* POPULATIONS IN GÖKSU RIVER ECOSYSTEM

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

BY

ÇİĞDEM KANSU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
BIOLOGY

APRIL 2018

Approval of the thesis:

GENETIC STRUCTURE OF *POPULUS EUPHRATICA* POPULATIONS IN GÖKSU RIVER ECOSYSTEM

submitted by **ÇİĞDEM KANSU** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Biology Department, Middle East Technical University** by,

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

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

Prof. Dr. Zeki Kaya _____
Supervisor, **Biological Sciences Dept., METU**

Examining Committee Members:

Prof. Dr. Musa Doğan _____
Biological Sciences Dept., METU

Prof. Dr. Zeki Kaya _____
Biological Sciences Dept., METU

Prof. Dr. Füsun Eyidoğan _____
Faculty of Education, Başkent University

Prof. Dr. Sertaç Önde _____
Biological Sciences Dept., METU

Assist.Prof.Dr. Fatih Temel _____
Forestry Engineering, Düzce University

17.04.2018

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: Çiğdem Kansu

Signature:

ABSTRACT

GENETIC STRUCTURE OF *Populus euphratica* POPULATIONS IN GÖKSU RIVER ECOSYSTEM

Kansu, Çiğdem
Ph.D., Department of Biological Sciences
Supervisor: Prof. Dr. Zeki Kaya
April 2018, 91 Pages

Poplar is one of the economically most important plants growing in Anatolia. Developing and growing world population and ascending demand for energy sources increased the genetic studies on poplars. There are four poplar species that have economical value and distributed naturally in Turkey; *Populus nigra, Populus tremula, Populus alba* and *Populus euphratica*. *Populus euphratica,* which can survive in salty and calciferous soils and known as Euphrates poplar, is resistant for low moisture and long summer drought. Having economical value for these respects, this species is distributed in East Mediterranean and Southeastern Anatolia. The species possesses great importance for both renewable energy resources and persistence of a healthy river ecosystem. Due to decreasing water resources with increased population and habitat destructions by human activities, the distribution area of this species become narrower gradually as well as loss of gene sources. Hence, searching for potential genetic diversity present in species' genetic resources is of great importance in terms of conservation (*in situ* and *ex situ*), breeding and use.

In this PhD Thesis, genetic structure and diversity of *Populus euphratica* populations in the Göksu river ecosystem were studied with 21 microsatellite DNA (SSR: single sequence repeats) markers. Results demonstrated reduced level of genetic diversity with low heterozygosity values (Ho:0.50±0.07, uHe:0.49±0.06). Severe past reductions in population sizes have resulted in loss of genetic variation for the species. Native populations of this species in the Göksu river are experiencing gene pool shrinkage and are in great danger of collapsing, mainly because of anthropogenic pressures. With great number of private alleles and slightly higher heterozygosity

values The Euphrates river population is shown to be an effective candidate of genetic resource for future conservation studies.

Effective breeding and conservation programs including *in situ* conservation of Euphrates poplar stands in the Göksu river and *ex situ* conservation by establishing gene banks and clone banks in order to preserve germplasm resources should be initiated immediately. The conservation of these genetic resources is of great importance and priority should be given for fast and reproducible actions.

Keywords: *Populus euphratica*, Microsatellite, SSR, genetic structure, genetic diversity

ÖZ

GENETIC STRUCTURE OF *POPULUS EUPHRATICA* POPULATIONS IN GÖKSU RIVER ECOSYSTEM

Kansu, Çiğdem
Doktora, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. Zeki Kaya
Nisan 2018, 91 Sayfa

Kavak, Anadolu'da yetişen en önemli ekonomik bitkiler arasında yer almaktadır. Türkiye'de ekonomik öneme sahip ve doğal olarak bulunan 4 kavak türü vardır; *Populus nigra*, *Populus tremula*, *Populus alba* ve *Populus euphratica*. Fırat kavağı olarak bilinen ve tuzlu, kireçli topraklarda yaşayabilen *Populus euphratica* düşük hava rutubetine, uzun yaz kuraklığına dayanıklıdır. Bu açıdan ekonomik bir önem taşıyan bu tür Akdeniz bölgesinin doğusu ve Güneydoğu Anadolu bölgesinde yayılış göstermektedir. Tür, yenilenebilir enerji kaynağı olması açısından ve sağlıklı bir nehir ekosisteminin devamlılığı için büyük öneme sahiptir. Artan nüfus ile birlikte azalan su kaynakları ve insan aracılığıyla yapılan habitat tahribatı nedeniyle türün yayılış alanı gittikçe daralmakta ve gen kaynakları kaybolmaktadır. Bu nedenle, türün genetic kaynaklarında bulunan potansiyel genetik çeşitliliğin araştırılması koruma (*in situ* ve *ex situ*), ıslah ve kullanım açısından çok önemlidir.

Bu doktora tezinde, Göksu nehir ekosistemindeki *Populus euphratica* popülasyonlarının genetic yapısı ve çeşitliliği 21 mikrosatelit DNA (SSR: basit tekrar dizileri) markörü ile çalışılmıştır. Sonuçlar oldukça düşük heterozigotluk değerleri ile azalmış genetik çeşitlilik seviyesi olduğunu göstermiştir (Ho:0.50±0.07, uHe:0.49±0.06). Geçmişte geçirilen şiddetli popülasyon küçülmeleri türün genetic varyasyonunun kaybolması ile sonuçlanmıştır. Türün Göksu nehrindeki doğal populasyonları gen havuzu daralması yaşamakta ve antropojenik baskılar yüzünden çok büyük bir çökme tehlikesi ile karşı karşıya kalmaktadır. Fırat nehri popülasyonunun çok sayıda özel allele sahip olması ve kısmen daha yüksek

heterozigotluk değeri ile gelecek koruma çalışmaarı için etkili bir gen kaynağı adayı olduğu öngörülmüştür.

Göksu nehri Fırat kavağı meşçerelerinin *in situ* korunması ile gen kaynaklarını korumak için gen bankası ve klon bankası kurulumuyla yapılacak *ex situ* koruma içeren etkili ıslah ve koruma programları derhal başlatılmalıdır. Bu genetik kaynakların korunması büyük önem arz etmektedir, bu yüzden hızlı ve tekrarlanabilir eylemlere öncelik verilmelidir.

Anahtar Kelimeler: *Populus euphratica*, Mikrosatelit, SSR, genetik yapılanma, genetik çeşitlilik

To my son

ACKNOWLEDGMENTS

Foremost, I would like to express my deepest gratitude to my supervisor Prof. Dr. Zeki Kaya for his guidance, encouragements and advice throughout the research. I would also like to express my regards to Prof. Sertaç Önde for his support and guidance all through my thesis study.

I express my sincere appreciation to all members of thesis examining committee for their comments and criticism.

I compassionately express my special thanks to Asiye Çiftçi and Funda Özdemir Değirmenci owing to their precious help with their broad knowledge about the topic, lovely attitude in writing this thesis and for their support and encouragements. I am also thankful to all my labmates for their friendship and help.

I am also thankful to Sadi Şıklar, Prof.Dr. Hayri Duman and Prof.Dr. Abdullah Hasbenli for their help in field studies.

I am grateful to my brother for his love and emotional support in my life. I would like to express my deepest thanks to my father for his love, patience, insight, support and encouragements along the way.

Finally, I would also like to express my special thanks to my spouse, Tuna for standing by me all through my life, his endless support, patience and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

%P	Percentage of Polymorphic Loci
AMOVA	Analysis of Molecular Variance
Ar	Allelic Richness
bp	Base pairs
CLUMPP	Cluster Matching and Permutation Program
CTAB	Cetyl Trimethyl Ammonium Bromide
EDTA	Ethylenediaminetetraaceticacid disodium salt
F	Inbreeding Coefficients
F_{CT}	Difference Among Groups For The Total Population
F_{IS}	Inbreeding Coefficient Within Individuals
F_{IT}	Inbreeding Coefficient Within Total Population
F_{SC}	Differences Among Population Within Groups
F_{ST}	Inbreeding Coefficient Within Subpopulations
GDA	Genetic Data Analysis
G-W Index (M)	Modified Garza-Williamson Index
uHe	Unbiased Expected Heterozygosity
Ho	Observed Heterozygosity
HWE	Hardy–Weinberg Equilibrium
I	Shannon's Information Index
MCMC	Markov Chain Monte Carlo
$MgCl_2$	Magnesium Chloride
Na	Number of Alleles
Ne	Number of Effective Alleles
Nm	Number of Migrants
PCoA	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
PI	Probability of Identity
PIC	Polymorphism Information Content
R_{ST}	Inbreeding Coefficient Within Subpopulations (based on size)
SSR	Simple Sequence Repeats
Ta	Annealing temperature
TE	Tris EDTA

TUBIVES	Türkiye Bitkileri Veri Servisi
UPGMA	Unweighted Pair-Group Method with Arithmetic Mean

CHAPTER 1

INTRODUCTION

Poplars and willows are species of Salicaceae family, which are deciduous trees or shrubs. The family comprises two genera with approximately 400–500 species, *Populus* (poplars, cottonwoods and aspens) and *Salix* (willows, sallows and osiers), respectively. Salicaceae was traditionally a temperate climate family consisting of poplars and willows. However, according to recent plastid DNA analysis, some tribes of the tropical family Flacourtiaceae and the Salicaceae have a common ancestry (Chase et al. 2002). Thus, the Angiosperm Phylogeny Group (2003) recommended a reordering of these taxa and the family is recognized as Salicaceae *sensu lato*, including tribes *Saliceae* (*Populus* and *Salix*), *Flacourtieae*, *Samydeae*, *Homalieae*, *Scolopieae*, *Prockieae*, *Abatieae* and *Bembiciea*. Yet, this classification did not have a predominant acceptance in literature.

Members of the family is distributed mainly in the temperate, boreal and tundra regions of the northern hemisphere, but there are species native to the subtropical and tropical regions of North America, Africa and Asia and South America as well (Isebrands and Richardson 2014)

According to Skvortsov (1999) and Dorn (1976) *Populus* is evolutionarily more primitive than *Salix*, which were diverged about 60–65 million years ago from a common paleotetraploid ancestor after the "salicoid" duplication event (Dai et al. 2014; Tuskan et al. 2006). After genome sequencing of *Populus*, it is understood that lineages of *Populus* and *Salix* shared the same large-scale genomic history (Tuskan et al. 2006).

1.1. *Populus* L.

Genus *Populus* has key species that are dioceous, wind pollinating and distributed especially along river ecosystems (Braatne et al. 1996). Members of the genus are suitable for industrial wood production worldwide. The number of species in genus is controversial in literature as few as 22 to as many as 85 (Eckenwalder 1996) due to broad distribution of the species, extensive phenotypic variation and presence of hybrids. According to the most commonly accepted classification, there are 29 species in the genus, which is divided into six sections, *Abaso*, *Turanga*, *Leucoides*, *Aigeiros*, *Tacamahaca* and *Populus*, respectively (Eckenwalder 1996).

In Turkey, genus *Populus* was identified with four native species in The Flora of Turkey and the Aegean Islands. These are namely, *Populus nigra* L. (black poplar), *Populus tremula* L. (aspen poplar), *Populus alba* L. (white poplar) and *Populus euphratica* Oliv. (Euphrates poplar) (Browicz and Yaltırık 1982). In addition to these, there is an introduced North American species *Populus deltoides* Marsh., which can make hybridizations with black poplar frequently.

Populus trees reach reproductive maturity at an age of 10–15 years under favorable conditions in natural populations (Stanton and Villar 1996). *Populus* flowers are catkins and flowering usually starts in early spring, lasts for 1-2 weeks and after that leaf emergence follows. (Braatne et al. 1996; Eckenwalder 1996). Wind Pollination continues one or two months (Braatne et al. 1996) and the distance pollen can reach during pollination is highly variable (Tabbener and Cottrell 2003). Fruits are capsules with very small seeds produced with cotton-like appendages in great numbers (Braatne et al. 1996).

Besides sexual reproduction, the genus has high capacity to propagate vegetatively via root suckers. Also, individuals have the ability to undergo vegetative propagation naturally by the help of broken branches, enabling genotypes to spread over large distances (Legionnet et al. 1997). This allows a successful genotype with desired characteristics to be distributed in favorable environment readily.

1.2. *Populus euphratica* Oliv.

1.2.1. Biology and Ecology

Euphrates poplar (*Populus euphratica*) is a dioceous, broadleaved, bushy tree which belongs to the section *Turanga* of genus *Populus*. It can attain a height of about 15 m and a girth of 2.5 m under favorable conditions. It is one of the pioneer species in riparian forests in arid areas due to its tolerance to drought, high salt concentration and dust storms. Furthermore, it is well known for its phreatophytic habit of growing in deserts (Gries et al. 2005). The species is diploid and possesses a basic haploid chromosome number of n = 19 just like the other poplar species.

This species has glabrous leaves which are glaucous-green on both surfaces. Leaves are heteromorphic, thus there is a leaf shape polymorphism in different parts of the crown. Young plants and shoots have lanceolate to elongate-ovate leaves with generally entire margins, resembling willow, while older trees and shoots have ovate-rhombic, elliptic-orbicular or reniform leaves with shallowly dentate margins in upper part (Figure 1.1) (Browicz and Yaltırık 1982; Mamıkoğlu 2007).

Female and male flowers are catkins on different individuals and the species is wind-pollinated (Figure 1.1). Flowering time depends on the geographic and climatic conditions, but in Turkey it usually starts in late March (Mamıkoğlu 2007). Fruits are capsules with very short-lived seeds enveloped in silky hairs for efficient wind dispersal. The species produces high number of seeds which could germinate on wet floodplains. Thus, seed dispersal generally coincides with the annual high flood. Germination depends on wet and exposed sites found at riverbanks, especially alluvial soil. As a result, *P. euphratica* is distributed widely along riverbanks (Westermann et al. 2008). Moreover, after germination, seedlings should have continuous access to ground water to survive (Gries et al. 2005).

Euphrates poplar has the ability to regenerate by root suckers besides seed/seedlings and propagating via broken branches is rare, if occurs. Clonal growth by root suckering starts when the plants reach at an age of 11–15 years and root suckers can bridge distances up to 40 m from the parent tree (Wiehle et al. 2009).

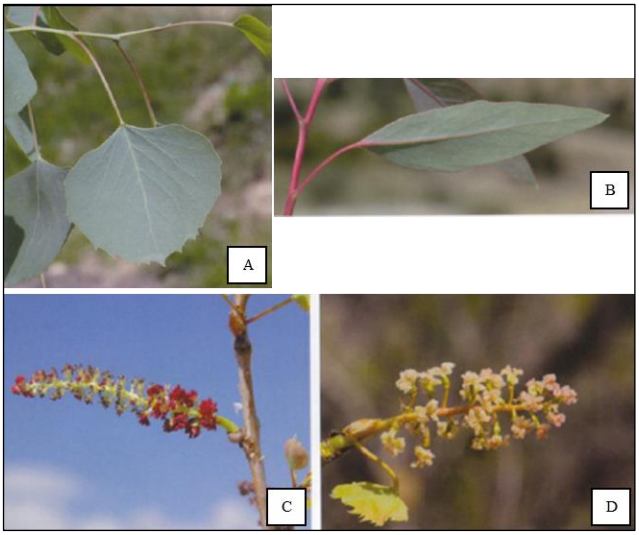

Figure 1.1: Representative pictures of leaves and flowers of Populus euphratica A&B: Polymorphic leaf shapes C: Male catkins D: Female catkins (Mamıkoğlu 2007).

The species is known for its remarkable survival under unfavorable conditions such as alkaline soils, drought and salt stress, and extreme temperatures. Euphrates poplar individuals survive very cold winters (down to –40°C) and hot summers (up to 43°C) with low rainfall and high evaporation rates.

Although it is highly tolerant to drought, water transport system of Euphrates poplar is damaged by cavitation if there is water deficit (Hukin et al. 2005). This is compensated by possessing a deep root system. Euphrates poplar is a phreatophyte, meaning that it is a deep-rooted plant that has to get water from a permanent ground supply or from the groundwater table. Therefore, survival of the species is dependent on access to deep water tables such as those occurring on riverbanks. This species is also well- known for its ability to cope with high levels of salinity (Chen et al. 2001; Gu et al. 2004).

Hybridization is common in *Populus* species and inter-specific hybridization accompanied by clonal selection is important for domestication of species with economically important traits. Hybridization trials of the species were done with *P. deltoides*, *P. nigra* and *P. simonii* (Zsuffa et al. 1996), and *P. alba* (Mofidabadi et al.

1998) from which successful hybrids were obtained. In Turkey, hybridization studies between *P. euphratica* and other fast growing, high wood content poplars were conducted in 1968-1971, but the attempts did not produce viable pollens or seeds (Tunçtaner 2008)

1.2.2. Distribution

The distribution range of *Populus euphratica* extends from northern Africa over the Middle East to Central Asia, northern India and China (Figure 1.2). There is also an isolated population with anthropogenic origin in Spain (Fay et al. 1999). The largest forests of Euphrates poplar are found in Kazakhstan and China (Thevs et al. 2008; Wang et al. 1996). This extraordinary species has latitudinal and altitudinal ranges from 48 to 49° North in Kazakhstan to 15° North in Yemen, and from 390 m below sea level in the Dead Sea depression to 4500 m in Kashmir (Browicz 1977).

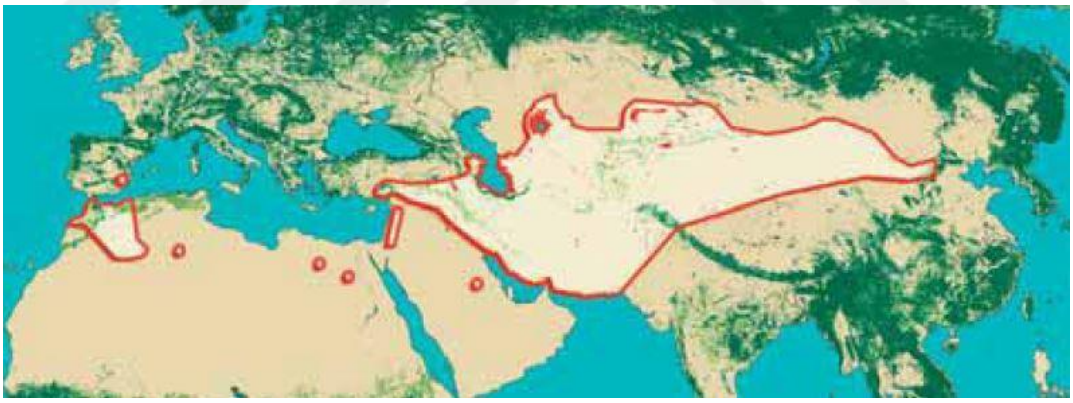

Figure 1.2: Natural range of *Populus euphratica* in Eurasia and Africa (Browicz 1977)

In Turkey, the species is distributed naturally throughout South and Southeastern Anatolia. The Eastern border of distribution is the Dicle river and its tributaries and Western border is the Bozyazı River in Anamur (Figure 1.3). The densest stands of this species are found in the Euphrates and the Göksu river basins.

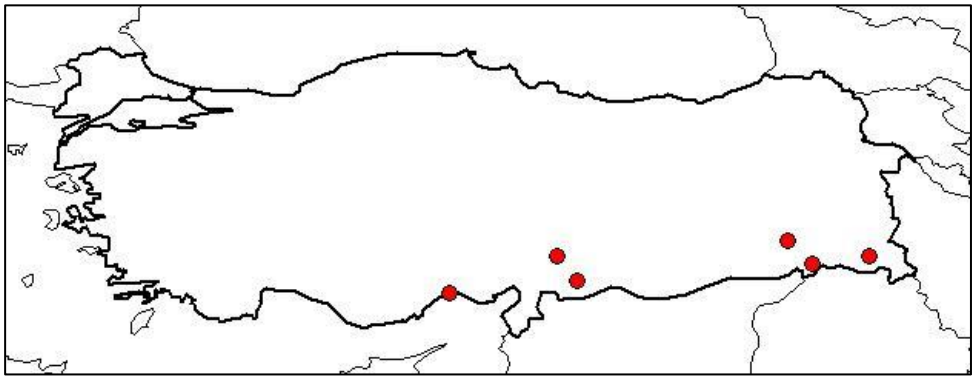

Figure 1.3: Natural distribution of *Populus euphratica* in Turkey (TUBIVES database http://www.tubives.com/index.php?sayfa=1&tax_id=8512)

1.2.3. Importance of Euphrates Poplar and Threats to the Speices

Populus euphratica, together with *Populus tremula, Populus nigra* and *Populus alba,* is important for wood and timber production for rural people since antique ages. It is also used in grazing sheep and goats, planted for windbreaks and stabilization of soil in its natural range as an ecosystem service. Besides, Euphrates poplar is one of the indicators to determine sustainable usage of river potentials. River ecosystems provide natural habitats for many socially, economically and ecologically important species. Since the mass production of the species is problematic due to groundwater access requirement for survival of seedlings, its role in desert and river ecosystems should be well studied and conserved.

Due to its natural survival capacity in extreme environmental conditions, projects aiming to obtain transcript profiles and expression studies were increased remarkably (Brosché et al. 2005; Gu et al. 2004; Qiu et al. 2011). Euphrates poplar is also drawing attention for hybridization programs because of its tolerance to high temperatures, salinity and drought (Calagari et al. 2004; Mofidabadi and Modir-Rahmati 2000).

Although the species does not have an assigned IUCN Red List category yet, the number of populations and individuals of this species has dropped considerably in recent years to the point where it should be now considered to be endangered all over the world (Bruelheide et al. 2004). In addition, The Forest Tree Genetic Resources Panel of the Food and Agriculture Organization of the United Nations (1995) declared that *P. euphratica* was one of the threatened boreal species.

This species is one of the most important poplar species in northwest China, where there is an *in situ* conservation program for *P. euphratica* forests in the Tarim river reserve in Xinjiang Autonomous Region (Yimit et al. 2006). Moreover, *P. euphratica* forests in Taklimakan desert in Xinjiang Uygur Autonomous Region were nominated for World Heritage List in 2010 (UNESCO World Heritage Center, 2010).

In Turkey, Euphrates poplar stand in Birecik, Şanlıurfa is a natural breeding area for Pallid Scops Owl (*Otus brucei*), which is in first priority conservation program in Europe (Karacadağ Development Agency, 2014). On the other hand, dam constructions in rivers for meeting the increased energy demand diminishes flooding and lowering groundwater tables, thus preventing survival of the seedlings. In some localities, habitats are extirpated for agricultural practices or fragmented for urbanization.

1.2.4. Euphrates poplar in Turkey

Besides its natural range in Turkey, it is known that Euphrates poplar was planted in Konya Karapınar for preventing erosion. Moreover, there is one plantation in Birecik, Şanlıurfa in the district of Birecik Forest Nursery Directorate.

The first study related with Euphrates poplar was giving information about identification and morphological characters of the species (Acatay 1961) Later on, there were hybridization and cloning trials (Gülbaba 1991) and studies investigating wood structure and anatomical features of the species (Acar 1973). Greenaway et al. (1991) studied the phenolics of bud exudates from the individuals sampled from Euphrates river. There is also a master's thesis studying morphological characters of the species (Karatay 2003). However, there is no study investigating genetic diversity and structure of *Populus euphratica* populations naturally found in Turkey. Although

recently the Karacadağ Development Agency of Ministry of Development organized a workshop on drawing public attention and suggesting conservation projects on the species, there exists no sufficient genetic data on genetic resources in order to initiate a comprehensive conservation strategy for populations in Turkey.

1.3. Microsatellite Markers

Microsatellites, short tandem repeats (STRs) or simple sequence repeats (SSRs) are tandem repeat motifs of 1-6 bp length. These repeat regions are present both in prokaryotes (Field and Wills 1998) and eukaryotes (Tautz and Renz 1984). They are among the most variable types of DNA sequence in nuclear (nSSR) or organellar genomes (cpSSR, mtSSR) (Provan et al. 2001; Rajendrakumar et al. 2007). They are length polymorphisms, which occurs both in coding and noncoding regions of genomes (Zane et al. 2002). However, majority of microsatellites occurs in intergenic sequence or in the introns, which are non-coding DNA. These repeat regions are the ones that are preferentially used as genetic markers due to their neutral evolution.

Length changes or allelic patterns in microsatellite DNA can stem from replication slippage and point mutations (Ellegren 2002; Kruglyak et al. 1998; Levinson and Gutman 1987). If these changes in microsatellite DNA are not corrected by mismatch repair system, they end up as new microsatellite mutations (Strand et al. 1993). The rate of these mutations can vary according to the loci and species under consideration. It is known as the higher the number of repeats, the higher the mutation rate is (Ellegren 2000). Presence of retrotransposons, which are repetitive DNA, is also related with generation of new microsatellite regions or new allelic combinations (Nadir et al. 1996; Temnykh et al. 2001).

1.3.1. Microsatellite markers in Population Genetics

Microsatellites are multiallelic markers with unique properties such as hyper variability, codominant inheritance, reproducibility, relative abundance, extensive genome coverage, chromosome specific locations, automated and high throughput genotyping (Parida et al. 2009). Thus, they have various applications in genome mapping, forensics, parentage analysis, population genetics and conservation genetics (Kalia et al. 2011). Nuclear and organellar microsatellite or SSR markers are used in genetic mapping (Gaudet et al. 2008; Han et al. 2004), identification of genetic

resources and diversity (Powell et al. 1995; Prasad et al. 2000) and conservation biology studies (Chase et al. 1996) frequently.

Although currently there are new high throughput genotyping techniques, genome wide abundant distribution and inexpensive analysis techniques of these markers ensure their significance in generating sufficient data to assess genetic variation. Hence, SSR markers are used to estimate the level of genetic diversity in various studies (Dayanandan et al. 1998; Fossati et al. 2003; Rajora et al. 2000; Rathmacher et al. 2010; van de Ven and McNicol 1996).

1.4. Population genetics studies of Euphrates poplar

Molecular and population genetics studies of *Populus euphratica* are very recent and limited in number. One of the very first study was assessing genetic variability of Israel populations with isozyme markers (Rottenberg et al. 2000). After that, there were studies investigating genetic diversity and structure of populations in China with RAPD and AFLP markers (Bruelheide et al. 2004; Saito et al. 2002). Later on, microsatellite markers were in action and some multiplex PCR systems with primers designed for other poplar species were tried to investigate genetic variation, sex ratio effects and clonal growth (Eusemann 2010; Eusemann et al. 2009, 2013; Petzold et al. 2013; Xu et al. 2013). While SSR primers designed for other poplar species were in use, new species specific markers were developed and used in genetic variation studies (Wang, Li et al. 2011; Wang et al. 2015; Wu et al. 2008). Besides microsatellite markers from non-coding regions of genome, EST-SSRs were developed to be used in population genetics, comparative genomics, linkage mapping, QTL, and marker-assisted breeding (Du et al. 2013).

CHAPTER 2

JUSTIFICATION AND OBJECTIVES OF THE STUDY

Euphrates poplar is a substantially important species for river ecosystems owing to its survival in stress conditions such as low humidity, high temperature, salinity and drought. It is one of the native species of poplars in Turkey and possesses great importance for both renewable energy resources and persistence of a healthy river ecosystem. Hence, investigating potential genetic diversity present in species' genetic resources is of great importance in terms of conservation (*in situ* and *ex situ*), breeding and use.

As a consequence of increased number of dams and levees to meet gradually increasing water and energy demand, the species' habitats are either fragmented or disappeared. For the Göksu river basin in particular, there are two dams (Ermenek and Gezende) and five hydroelectric power stations on the Göksu river, which cause flow decrease and lower groundwater tables. Moreover, agricultural activities are one of the major sources of income for the rural population and riverbanks have been cleared to extend the land for cultivation. Hence, in Göksu river basin genetic resources of the species are in great danger due to anthropogenic pressures. Consequently, the populations with extensive gene pool that could contribute to species diversity should be determined and included in conservation programs immediately. Hence, the goals of this study are;

- to determine reproducible and informative microsatellite loci for studying genetic variation of Euphrates poplar,
- to investigate magnitude of genetic diversity and structure of Euphrates poplar populations in the Göksu river basin with microsatellite markers,
- to generate information for initiating efficient conservation programs and preserving and use of genetic resources in the Göksu river basin.

Furthermore, an additional population from the Euphrates river was included to the study in order to compare genetic diversity of two different river systems and to identify whether the gene pool of the populations in Göksu are enriched by gene flow from an outsider population.

CHAPTER 3

MATERIALS AND METHODS

3.1. Plant Material and Field Work

The field studies for sampling was held on in August 2014. For DNA extraction and further molecular analysis young leaf samples were gathered from *Populus euphratica* trees. There were five sampling locations (populations) in two main rivers, Göksu and Euphrates, respectively (Table 3.1). Their corresponding locations are shown on Turkey rivers map in Figure 3.1. Göksu populations were determined to represent upstream tributaries (Gökcay and Ermenek), middle (Mut) and downstream (Silifke) of the river. There is only one population representing the Euphrates river which is the closest portion to the Turkey-Syria border. In field, leaf samples were collected from individuals at least approximately 200 m apart in order to prevent clonal sampling. The number of individuals in each population varied according to the size of the population. Leaf samples from each tree were gathered and placed in silica gel containing zipper bags until DNA extraction to avoid decay. Pictures from field work are given by Figure 3.2.

Table 3.1: Sampling locations and corresponding populations in two rivers

River	Population ID	Altitude Range	Number of Individuals
Göksu	Gokcay	250-288m	46
	Ermenek	331-359m	23
	Mut	86-112m	39
	Silifke	30-92m	32
Euphrates	Euphrates	33-351m	20

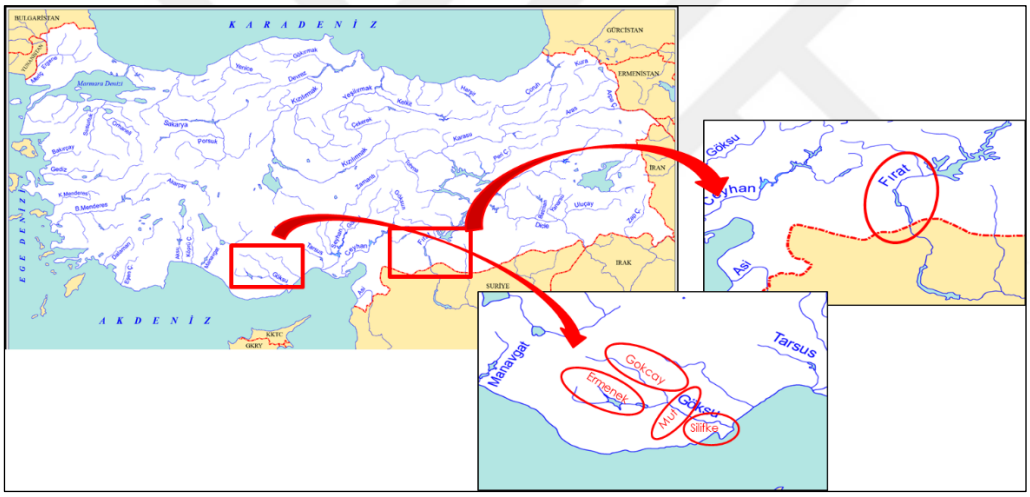

Figure 3.1: Sampling locations and corresponding populations in Turkey river map (adopted from Cografya Harita website, http://cografyaharita.com/turkiye_hidrografya_haritalari.html)

Figure 3.2: Photographs from field work study areas. A and B: Views from the study area along the Göksu river C: Young Populus euphratica individuals in the Göksu river basin D: A view from the study area along the Euphrates river

3.2. Molecular Studies

3.2.1. DNA Extraction

For the DNA extraction, leaf samples were first ground with mortar and pestle using liquid nitrogen (-196°C). From the powder obtained, ~0.1g was used for DNA extraction. Total DNA extraction was performed according to Doyle (1991) CTAB extraction protocol with some modifications (Appendix A). Visual confirmation of total DNA was done with Agarose gel electrophoresis and concentration and purity of the obtained DNA samples were measured using NanoDrop 2000/2000c Spectrophotometer (Thermo Scientific).

3.2.2. PCR Amplification of Microsatellite Loci

For the population genetics study, first a panel of thirty-two microsatellite loci were used to check the amplification success of the corresponding primers. Among those, twenty-two of the loci were successfully amplified. The repeat motif, primers and their reference, and expected product size for each locus are given in Table 3.2. In order to perform microsatellite genotyping for the individuals, DNA Fragment Analysis by Capillary Electrophoresis was done. For this purpose, one of the primers was labelled with a fluorescent dye for each locus prior to amplification (Table 3.4). The Polymerase chain reaction conditions for the loci used are given in Table 3.3. In PCR reactions 5x HOT FIREPol® Blend Master Mix Ready to Load (Solis BioDyne, Tartu, Estonia), which included 15mM $MgCl_2$. was used. The presence of products was verified by running 3% Agarose gel electrophoresis.

Table 3.2: Characteristics of 22 microsatellite loci used

Locus ID	Repeat Motif	Left Primer (5′–3′)	Right Primer (5′–3′)	Product size	Reference
WPMS5	GT	TTCTTTTTCAACTGCCTAACTT	TGATCCAATAACAGACAGAACA	280	Smulders et al. 2001; van der Schoot et al. 2000
WPMS7	GT	ACTAAGGAGAATTGTTGACTAC	TATCTGGTTTCCTCTTATGTG	230	
WPMS10	GT	GATGAGAAACAGTGAATAGTAAAGA	GATTCCCAACAAGCCAAGATAAAA	250	
WPMS12	GT	TTTTTCGTATTCTTATCTATCC	CACTACTCTGACAAAACCATC	170	
WPMS14	CGT	CAGCCGCAGCCACTGAGAAATC	GCCTGCTGAGAAGACTGCCTTGAC	245	
WPMS15	CCT	CAACAAACCATCAATGAAGAAGAC	AGAGGGTGTTGGGGGTGACTA	193	
WPMS18	GTG	CTTCACATAGGACATAGCAGCATC	CACCAGAGTCATCACCAGTTATTG	245	
WPMS20	TTCTGG	GTGCGCACATCTATGACTATCG	ATCTTGTAATTCTCCGGGCATCT	252	
PMGC14	CTT	TTCAGAATGTGCATGATGG	GTGATGATCTCACCGTTTG	210	The International *Populus* Genome Consortium
PMGC2163	GA	CAATCGAAGGTAAGGTTAGTG	CGTTGGACATAGATCACACG	220	
PMGC2889	GA	CCCAAGATCCGATTTTTGGG	CACAATGTACAAATCGCTGTC	207	
PMGC93	CTT	ATCATGCGTTCGGCTACAGC	CTCAAACTCCAACTGTTATAAC	350	
Pe2	$(CT)_9(CA)_{11}$	CGCTATCATCCTCATCTCTGATCTC	CCATGCTGGGATGAAATATGTAAAC	128–147	Wu et al. 2008
Pe5	TC	ACCCACCCAATGTGCAGCCCTGCAA	CTCGCCCTCTATATATCTCTATGAA	164–206	
Pe6	TG	GAACATGGGTCCATATCAAGTAG	CACCCCACACACCCACAC	175–233	
Pe7	TG	ACTATGGAGCACGCAAAATGCCAAG	TAGTTCCTGCTCTGCTTATCGTC	170–228	
Pe8	$(TG)_{12}(GA)_{11}$	AACACGGAAGCAAGAAAAAATGAAG	CACACATTCCACCCTCCACACCACT	146–197	
Pe13	$(CT)_6(GT)_5$	TTCAACTTGACTAGTTGTAACTCTC	CACTTTCCCAGCTATCCCTTTCTAA	121–137	
Pe14	CT	TCGAAATGGGAGATCTGTAGAGGTG	CACCACAACAGCGTACAGAAATGAA	137–178	
Pe15	AG	TGTTTAAGAAGAGATCAAAAGGGGA	TTTCTCAGGTATCCAAGCGATGCTG	124–161	
Pe17	AC	TAACATTGCCATCATCACTATCA	GATGCTAAAATGGTGGTAGGAGG	160–211	
Popeu13	AT	TCTGGAGCACAATTGAATTTTTG	TCTTTGAAATTCATGCTACTTCA	270–330	Wang et al. 2015

Table 3.3: PCR conditions and cycling parameters

Loci	Ingredients	Volume (µL)	PCR Cycling Parameters		
WPMS10	dH2O	8.9	Initial Denaturation 3min.		94°C
	Master Mix	6.5	28 cycle	Denaturation 30 sec.	94°C
	Primers (10µM)	0.8+0.8		Annealing 30 sec.	Ta
	DNA (10ng/µL)	8		Extension 40 sec.	72°C
	Total	25	Final Extension 20 min.		72°C
Pe2	dH2O	11	Initial Denaturation 3min.		94°C
	Master Mix	6	28 cycle	Denaturation 30 sec.	94°C
	Primers (10µM)	0.5+0.5		Annealing 30 sec.	Ta
	DNA (10ng/µL)	7		Extension 40 sec.	72°C
	Total	25	Final Extension 20 min.		72°C
Pe17 and Popeu13	dH2O	10.6	Initial Denaturation 3min.		94°C
	Master Mix	5	28 cycle	Denaturation 30 sec.	94°C
	Primers (10µM)	0.7+0.7		Annealing 30 sec.	Ta
	DNA (10ng/µL)	8		Extension 40 sec.	72°C
	Total	25	Final Extension 20 min.		72°C
WPMS5, WPMS7, WPMS14, WPMS15, WPMS18,WPMS20 PMGC14, PMGC2168, PMGC2889, PMGC93	dH2O	14	Initial Denaturation 3min.		94°C
	Master Mix	5	25 cycle	Denaturation 30 sec.	94°C
	Primers (10µM)	0.5+ 0.5		Annealing 30 sec.	Ta
	DNA (10ng/µL)	5		Extension 40 sec.	72°C
	Total	25	Final Extension 20 min.		72°C
Pe5, Pe6, Pe7, Pe8, Pe9, Pe13, Pe14, Pe15	dH2O	12	Initial Denaturation 3min.		94°C
	Master Mix	5	25 cycle	Denaturation 30 sec.	94°C
	Primers (10µM)	0.5+ 0.5		Annealing 30 sec.	Ta
	DNA (10ng/µL)	7		Extension 40 sec.	72°C
	Total	25	Final Extension 20 min.		72°C

Table 3.4: Annealing temperatures of the primers for the microsatellite loci and the fluorescent dyes used

Locus ID	Ta (°C)	Fluorescent Dye
WPMS5	55	6-FAM
WPMS7	55	6-FAM
WPMS10	51	HEX
WPMS12	50	TAMRA
WPMS14	60	HEX
WPMS15	60	6-FAM
WPMS18	57	TAMRA
WPMS20	54	6-FAM
PMGC14	55	6-FAM
PMGC2163	55	HEX
PMGC2889	55	HEX
PMGC93	55	TAMRA
Pe2	59	VIC
Pe5	57	NED
Pe6	59	VIC
Pe7	57	PET
Pe8	60	6-FAM
Pe13	57	PET
Pe14	60	6-FAM
Pe15	52	NED
Pe17	53	PET
Popeu13	52	VIC

3.2.3. DNA Fragment Analysis for Microsatellite Genotyping

After amplification of products with labelled primers, the fragment analysis by Capillary Electrophoresis was done in BM Laboratory Systems Facilities, Ankara. Assay procedure for fragment analysis was done with the Applied Biosystems 3730 XL DNA Analyzer (Applied Biosystems, Foster City, CA, USA) using an internal standard size marker The GeneScan ROX labelled 400HD. In this analysis the software uses the size standard to obtain a standard curve after amplified fragments were separated by size using capillary electrophoresis together with size standard fragments. Then it determines the relative size of each fragment in the sample studied by comparing fragments with the standard curve and assigns allele calls of the sample for the loci under evaluation. The result of this analysis is an electropherogram, an

example of which is shown in Figure 3.3. In order to obtain this electropherogram, one of the primers for each locus was labelled with a fluorescent dye, so that each locus has fragment peaks in different colors (Table 3.4).

Base callings were manually checked in Peak Scanner 2.0 software (Applied Biosystems) and allele sizes were recorded for each individual per locus to obtain microsatellite genotypes.

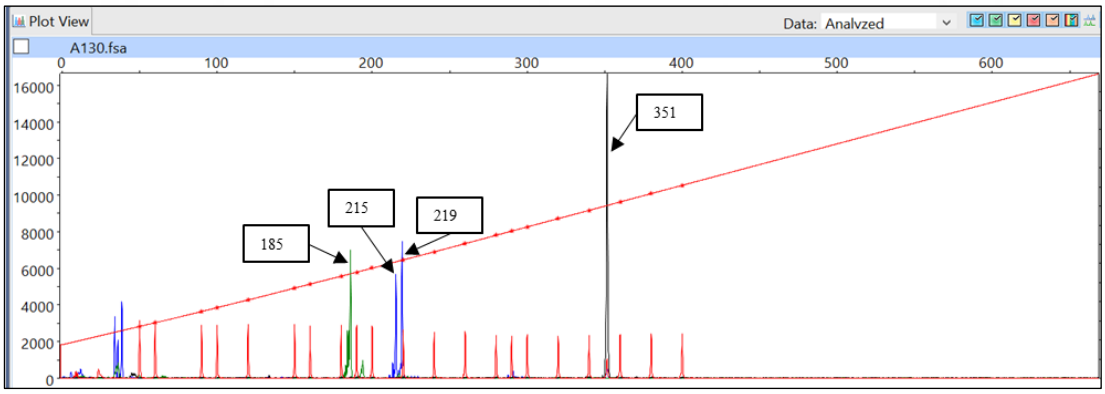

Figure 3.3: A microsatellite electropherogram example of 3 different SSR loci used in the study (y-axis: Relative Fluorescence Units, x-axis: Size of the fragment, numbers in black boxes are showing the sizes of the corresponding fragments)

3.3. Population Genetics Analyses

The formula for population genetics statistics were given in appendix B and representative file formats were shown in appendix C

3.3.1. Detection of Clones

Salicaceae family has the characteristics of clonal reproduction and genus *Populus* has the ability to propagate vegetatively via root suckers or by the help of broken branches. As a member of the genus, clonal propagation by root suckering in *P.euphratica* sets in when the plants reach an age of 11–15 years and root suckers can reach at distances up to 40 m (Wiehle et al. 2009)

The data obtained from the fragment analysis of 22 microsatellite loci were analyzed with GenClone 2.0 software (Arnaud-Haond and Belkhir 2007) in order to determine any clonal sampling among genotypes in the populations.

3.3.2. Null allele presence

One of the problems affecting analysis of microsatellite loci is the presence of null alleles that are caused by (i) mutations in the primer binding region, thus preventing amplification of true alleles (Pemberton et al. 1995) (ii) the preferential amplification of short alleles due to DNA template quality, and (iii) DNA slippage during PCR amplification (Chapuis and Estoup 2007; Shinde et al. 2003).

The presence of null alleles could appear as an excess of homozygotes which in turn leads to deviations from Hardy-Weinberg equilibrium (Carlsson 2008). Thus, null alleles may result in overestimation of genetic differentiation and affect population genetic analyses (Chapuis and Estoup 2007; de Sousa et al. 2005). In literature many different estimating methods are present (Brookfield 1996; Chakraborty et al. 1992; Dempster et al. 1977; Kalinowski et al. 2007; Marshall et al. 1998; Van Oosterhout et al. 2004; Weir 1996). There are numerous computer programs that are estimating null allele frequency (r) based on those different methods such as Genepop (Rousset 2008), Cervus 3.0.7 (Marshall et al. 1998), Microchecker (Van Oosterhout et al. 2004) and ML-NullFreq (Kalinowski and Taper 2006).

For the null allele presence, Brookfield's (1996) estimator method implemented in Genepop 4.2 (Raymond and Rousset 1995) was first used to detect null allele

frequencies for each locus in each population. But Dabrowski et al. (2014) suggested that the null allele detection may be improved by combining results of several methods. Thus, subsequent analysis was done with MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) to check scoring errors and allele dropouts in addition to presence of null alleles. The program estimates the frequency of null alleles together with identification of short allele dominance (large allele dropout) and the scoring of stutter peaks. The null allele frequency (r) is estimated using the methods described by Chakraborty et al. (1992) and Brookfield (1996). Further additional analysis was performed with Cervus 3.0.7, in which null allele frequency is estimated by analyzing deviations from Hardy–Weinberg equilibrium using a chi-square goodness-of-fit test (Marshall et al. 1998) and results are shown in Appendix D.

3.3.3. Microsatellite Polymorphism and Genetic Diversity Parameters

For each locus and population, number of alleles (Na) and mean effective number of alleles (Ne), observed (Ho) and unbiased expected (uHe) heterozygosity, Fixation index (F), Shannon's Information Index (I) and Private alleles and their frequencies were assessed by GenAlEx 6.503 (Peakall and Smouse 2012). In addition to number of alleles and mean effective number of alleles, allelic richness (Ar) was computed by FSTAT version 2.9.3.2 (Goudet 1995). It is representing the number of alleles standardized according to the smallest sample size so that excluding the sample size bias in the study. Variation in sampling numbers would bias estimates of allelic richness. Thus, while estimating Ar, FSTAT is incorporating a rarefaction option, which is used to standardize Ar to the smallest sample size in the analysis (Petit et al. 1998).

Polymorphism information content (PIC) at each locus was calculated by Cervus version 3.0.7 (Kalinowski et al. 2007). In addition, for each population proportion of polymorphic loci (%P) and the probability of identity (PI) were estimated with GenAlEx 6.503 (Peakall and Smouse 2012). Genepop 4.2 (Raymond and Rousset 1995; Rousset 2008) was used to asses F statistics (F_{IS}, F_{ST} and F_{IT}) for each locus among populations and to implement exact tests for testing the Hardy-Weinberg deviations. A Markov chain (MC) algorithm with default parameters is used to estimate without bias the exact p-value of this test (Dememorization number:1000, Number of batches:100, Number of iterations per batch:1000) (Guo and Thompson

1992). Gene flow was estimated from F_{ST} obtained from GenAlEx 6.503 (Peakall and Smouse 2012) ($Nm= 0.25 (1-F_{ST})/F_{ST}$).

To test for any experienced population reductions in effective size of populations, Garza-Williamson Index (Garza and Williamson 2001) was calculated for all loci and populations with the software Arlequin version 3.5.1.2 (Excoffier and Lischer 2010).

To estimate partition of the variation among rivers, between and within populations, the analysis of molecular variance (AMOVA) was carried out for all loci as implemented in Arlequin version 3.5.1.2 (Excoffier and Lischer 2010). There are two different analyses for AMOVA in Arlequin software; namely, the number of different alleles (F_{ST}) based (the infinite allele model) and the sum of squared size difference (R_{ST}) based (the stepwise mutation model). Both of them were applied to assess the variance of the data.

3.3.4. Population Genetic Structure

Genetic differentiation between studied populations was examined by pairwise F_{ST} values. Pairwise F_{ST} matrix and their corresponding Nm values were obtained by using Arlequin version 3.5.1.2 (Excoffier and Lischer 2010).

To show the genetic differences over populations, the pairwise F_{ST} values were used to generate a principal coordinates analysis (PCoA) based on the covariance matrix with data standardization as implemented in GenAlEx 6.503 (Peakall and Smouse 2012).

The relationships among the populations were assessed with a phenogram based on coancestry identities with GDA (Genetic data analysis) software (Lewis and Zaykin 2001). As the distance method UPGMA (Unweighted Pair-Group Method with Arithmetic Averaging) algorithm was used.

To identify the genetic structure of populations, a Bayesian iterative algorithm was used to assign individuals to clusters as implemented in STRUCTURE v2.3.4 (Pritchard et al. 2000). Admixture model was assumed and run settings were as follows: a burn-in of 50,000 and 250,000 Markov chain Monte Carlo iterations, possible cluster numbers (K) tested from K=1 to K=5 for Göksu populations only and K=1 to K=10 for all populations including Euphrates. ten replications were performed for each K.

Web-based tool Structure Harvester (Earl and vonHoldt 2012) was used to assess the most likely value of K (true K) from Structure run results for the detection of the number of genetic groups present in the data by implementing Evanno method (Evanno et al. 2005). Multiple runs for the true K value were analyzed with CLUMPP software (Jakobsson and Rosenberg 2007) to identify the best alignment to the replicate results of the cluster analysis. Lastly, Pophelper (Francis 2016) was used to visualize the output generated from the admixture analysis.

CHAPTER 4

RESULTS

4.1. Detection of Clones and Null Allele Presence

In the field studies of the project to prevent sampling of clonal individuals, leaf samples were gathered at least 200m apart. Further, to be certain that there are no clones among the sampled individuals in populations, we analyzed the data with GenClone 2.0 software (Arnaud-Haond and Belkhir 2007) and found no repeated sampling.

In the analysis of the data in this study Genepop 4.2 was first used to detect null allele frequencies for each locus in each population. Further analysis was done with MICRO-CHECKER 2.2.3 to check scoring errors and allele dropouts in addition to presence of null alleles. The results provided by these programs are shown in Table 4.1 and 4.2, respectively. Additional results obtained by using Cervus 3.0.7 are shown in Appendix D.

Table 4.1: Null allele frequencies of the studied 22 microsatellite loci (Brookfield's 1996 method)

Locus	**Populations**				
	Gokcay	**Ermenek**	**Mut**	**Silifke**	**Euphrates**
WPMS5	*0.3254*	0.0130	*0.2639*	0.0168	*0.3294*
WPMS7	*0.8362*	No inf	*0.8165*	*0.9186*	*0.2500*
WPMS10	0.0389	No inf	0.0000	No inf	0.0000
WPMS12	0.0000	0.0000	0.0646	0.0165	*0.2357*
WPMS14	0.1618	0.0486	0.0000	0.0942	0.6000
WPMS15	0.0000	0.0000	0.0000	0.0000	0.0000
WPMS18	*0.2452*	*0.2475*	0.0000	*0.3311*	*0.2858*
WPMS20	No inf	No inf	No inf	No inf	No inf
PMGC14	No inf	*0.1879*	No inf	No inf	0.0000
PMGC2163	0.0000	0.0000	0.0000	0.0000	0.0173
PMGC2889	0.0000	0.0926	0.0000	0.0220	0.0000
PMGC93	0.0821	0.0928	0.1195	0.0000	0.0000
Pe2	0.0000	0.0000	0.0000	0.0000	0.0000
Pe5	0.0000	0.0000	0.0090	0.0000	0.0000
Pe6	0.0460	*0.2482*	*0.2097*	*0.1853*	*0.5577*
Pe7	No inf	No inf	No inf	No inf	No inf
Pe8	0.0000	0.0000	0.0000	0.0000	0.1083
Pe13	*0.4080*	*0.4670*	*0.3725*	*0.5422*	*0.3000*
Pe14	0.0000	0.0595	0.0515	0.0000	0.0000
Pe15	0.0000	0.0000	0.0477	0.0000	0.0000
Pe17	*0.2508*	*0.3009*	*0.1090*	*0.2942*	*0.2111*
Popeu13	0.1808	0.4200	0.0323	0.0996	0.0860

High null allele frequencies are shown in italics.

Table 4.2: Null allele frequency estimations for each locus (MICROCHECKER 2.2.3) (Brookfield 1996; Chakraborty et al. 1992) *Possible null allele possessing loci

Population / Locus	Gokcay		Ermenek		Mut		Silifke		Euphrates	
	Chakraborty	Brookfield	Chakraborty	Brookfield	Chakraborty	Brookfield	Chakraborty	Brookfield	Chakraborty	Brookfield
WPMS5	0,1967*	0,1292*	0,0213	0,013	0,1139	0,0802	0,0785	0,039	0,1144	0,0742
WPMS7	0,8959*	0,268*	0	0	1*	0,3077*	1*	0,2087*	-0,2308	-0,1915
WPMS10	0,4875	0,0389	0	0	-0,0065	-0,0003	0	0	-0,0423	-0,0137
WPMS12	0,0528	0,0137	-0,0135	-0,0075	0,1686	0,0646	0,0413	0,0165	0,764*	0,2357*
WPMS14	0,034	0,0254	-0,1522	-0,1401	-0,0688	-0,0556	-0,0319	-0,0258	-0,0047	-0,0027
WPMS15	-0,011	-0,0009	-0,011	-0,0009	-0,0196	-0,0036	-0,0364	-0,0112	-0,2012	-0,1676
WPMS18	0,1138	0,0809	0,0673	0,0446	-0,2359	-0,2244	0,1071	0,0681	0,2547*	0,1362*
WPMS20	0	0	0	0	0	0	0	0	0	0
PMGC14	0	0	0,4743	0,0699	0	0	0	0	-0,0127	-0,0012
PMGC2163	-0,1648	-0,1169	-0,2105	-0,1664	-0,04	-0,0249	-0,1963	-0,1494	0,0328	0,0173
PMGC2889	-0,0065	-0,0054	0,1238	0,0926	-0,0606	-0,0562	0,0261	0,022	-0,0847	-0,074
PMGC93	0,161*	0,1105*	0,0838	0,059	0,0337	0,0265	-0,0279	-0,0233	-0,2167	-0,1881
Pe2	-0,0095	-0,0079	-0,0685	-0,0615	-0,0189	-0,016	-0,0427	-0,0371	-0,0345	-0,0326
Pe5	-0,1819	-0,1591	-0,1396	-0,107	0,0136	0,009	-0,1064	-0,0874	-0,1215	-0,1144
Pe6	-0,17	-0,161	-0,0007	-0,0006	-0,0376	-0,0299	-0,0307	-0,0249	0,1919	0,1096
Pe7	0	0	0	0	0	0	0	0	0	0
Pe8	-0,0465	-0,0175	-0,1153	-0,0691	-0,1033	-0,0677	-0,019	-0,0109	0,034	0,0244
Pe13	0,121	0,0719	0,0737	0,0441	0,1291	0,0761	-0,0769	-0,0455	-0,2121	-0,1684
Pe14	-0,0797	-0,0693	0,0891	0,0595	0,065	0,0515	-0,0657	-0,0598	0,0344	0,0273
Pe15	-0,0331	-0,0294	-0,1227	-0,1062	0,0584	0,0477	-0,0215	-0,0182	0,0361	0,0311
Pe17	0,3934*	0,2508*	0,5736*	0,3009*	0,2515*	0,1389*	0,5486*	0,2942*	0,3694*	0,233*
Popeu13	0,2722*	0,1725*	0,6923*	0,36*	0,1212	0,0881	0,0596	0,0422	0,0382	0,0336

Since it is not possible to identify which individuals are true homozygotes and which are heterozygotes for the null allele without sequencing, markers/loci which are prone to null alleles should be used with extreme caution. The null allele possessing locus Pe17, which is also showing significant HW deviation, was discarded from further analysis since there occurred prominent changes in effective population size, observed and expected heterozygosity values and population F_{IS} values when it was included in the analyses. Besides, MICROCHECKER 2.2.3 results showed that this particular locus may have had large allelic dropouts, which is the failure of amplification of one or both allelic copies at a locus by the polymerase chain reaction. For the other loci showing high null allele frequency especially for particular populations, the results were also checked but there are no notable variations in effective population size, observed and expected heterozygosity values and F statistics when they have been discarded.

4.2. Genetic Diversity

4.2.1. Microsatellite Polymorphism and Genetic Diversity Parameters of Loci

The alleles detected for each locus are shown in Table 4.3. The polymorphism genetic variation parameters of the twenty-one SSR loci are presented in Table 4.4. The F statistics and number of migrants for each locus over all populations are shown in Table 4.5.

Among the analyzed twenty-one microsatellite loci, two are monomorphic for the populations under study. The rest nineteen loci are polymorphic having a mean of 3.50±0.19 alleles per locus. The least variable ones are WPMS7 and Pe13 loci and the most variable one is Pe2 locus, displaying two and ten alleles, respectively.

The number of alleles per locus (Na) ranged from 1 to 7.8, and number of effective alleles per locus (Ne) ranged from 1 to 4.02. The allele size ranged from 97 bp to 369 bp (Table 4.3). PI (Probability of identity) is demonstrating that two individuals drawn at random from a population will have the same genotype at multiple loci (Waits et al., 2001). WPMS10 and WPMS15 showed higher PI values, but the rest of the loci had sufficiently low PI values. The PIC (polymorphism Information Content) values varied from 0 (WPMS20 and Pe7) to 0.740 (Pe15). Eleven SSR loci were highly informative (PIC>0.5) and two of the twenty-one loci were monomorphic for the studied

populations. The value of allelic richness (Ar) ranged from 1 to 7.36. Genetic variability for loci studied is represented by I (Shannon's information index) values and it was moderate with the mean value of 0.81±0.05. The moderate value of Shannon's index represents the effectiveness of microsatellite loci to reveal the variation present.

The mean observed and unbiased expected levels of heterozygosity were 0.44±0.03 and 0.45±0.03, respectively (Table 4.4). Ho was higher than uHe at eight SSR loci, showing heterozygote excess in accordance with the negative fixation indices (F) of the corresponding loci. Five loci showed significant departures from HWE ($p<0.05$, $p<0.01$ or $p<0,001$) in the natural populations.

The analysis of twenty-one loci revealed that eleven loci had a negative F_{IS} value across all populations (Table 4.4) showing heterozygote excess (outbreeding) compared with HW expectations. F_{ST} values ranged between 0 and 0.28 with a mean of 0.06 over all loci and all populations.

Table 4.3: Observed alleles in each locus studied.

Locus	Alleles detected (bp)
WPMS5	287, 289, 291
WPMS7	215, 219
WPMS10	220, 222, **246**, **248**
WPMS12	**175**, 178, 180, 182, **186**
WPMS14	210, 213, 219, 231, 237
WPMS15	**179**, 182, 185, 194
WPMS18	226, 229, 232
WPMS20	214
PMGC14	**188**, 197, **200**
PMGC2163	195, 201, **203**, **205**
PMGC2889	179, 181, 185, 187, 189, 193
PMGC93	348, 351, 354, 357, 369
Pe2	**97**, 99, 101, 103, 105, 107, 109, **111**, 120, **124**
Pe5	160, 166, 168, 170, 172, **176**
Pe6	155, 157, 163
Pe7	172
Pe8	143, 156, 158, 160
Pe13	129, 133
Pe14	139, 144, 146, 148, 150, 152, **154**
Pe15	130, 132, 134, 138, 140, 142, 144, 148, **153**
Popeu13	280, 282, 284, 286, 288, 290, 292, 296, 298

Private alleles are shown in bold.

Table 4.4: Genetic diversity parameters of the 21 microsatellite loci

Locus	N	Na	Ne	Ar	PI	PIC	I	Ho	uHe	F	HW
WPMS5	32±4.85	2.8±0.20	2.27±0.25	3.00	0.308	0.530	0.86±0.10	0.43±0.03	0.54±0.06	0.19±0.05	***
WPMS7	32±4.85	1.8±0.20	1.54±0.16	2.00	0.564	0.297	0.46±0.12	0.15±0.15	0.32±0.09	0.59±0.35	***
WPMS10	32±4.85	1.8±0.37	1.06±0.04	1.80	0.901	0.049	0.12±0.07	0.05±0.04	0.06±0.04	0.18±0.18	ND
WPMS12	29.2±3.56	3±0.32	1.43±0.08	2.90	0.536	0.259	0.54±0.07	0.21±0.06	0.30±0.04	0.26±0.16	ND
WPMS14	32±4.85	4.2±0.20	2.49±0.22	4.12	0.238	0.601	1.06±0.08	0.65±0.08	0.59±0.05	-0.10±0.07	NS
WPMS15	32±4.85	2.4±0.24	1.28±0.18	2.55	0.743	0.172	0.30±0.11	0.23±0.13	0.18±0.09	-0.13±0.09	ND
WPMS18	32±4.85	3±0	2.33±0.15	2.99	0.276	0.514	0.92±0.04	0.53±0.10	0.57±0.02	0.06±0.18	NS
WPMS20	31.8±4.97	1±0	1±0	1.00	1.000	0.000	0.00	0 .00	0.00	-	ND
PMGC14	32±4.85	1.4±0.24	1.04±0.03	1.41	0.937	0.025	0.07±0.05	0.02±0.01	0.04±0.02	0.31±0.21	ND
PMGC2163	32±4.85	2.4±0.24	1.74±0.04	2.56	0.414	0.354	0.65±0.01	0.55±0.06	0.43±0.01	-0.29±0.12	*
PMGC2889	31.8±4.97	5.6±0.24	3.71±0.25	5.40	0.115	0.728	1.46±0.03	0.73±0.05	0.74±0.02	-0.01±0.07	NS
PMGC93	32±4.85	4.2±0.37	2.79±0.21	4.43	0.200	0.614	1.15±0.09	0.63±0.06	0.64±0.03	-0.02±0.14	NS
Pe2	31.6±4.78	6±0.63	4.02±0.56	6.54	0.117	0.720	1.48±0.10	0.79±0.03	0.75±0.03	-0.07±0.02	NS
Pe5	32±4.85	4.8±0.20	2.40±0.26	4.81	0.240	0.545	1.08±0.08	0.71±0.07	0.58±0.04	-0.25±0.08	***
Pe6	32±4.85	3±0	2.64±0.15	3.00	0.221	0.572	1.02±0.04	0.65±0.09	0.63±0.03	-0.05±0.12	*
Pe7	32±4.85	1±0	1±0	1.00	1.000	0.000	0 .00	0.00	0.00	-	ND
Pe8	31.8±4.94	3.8±0.20	1.73±0.19	3.84	0.416	0.409	0.74±0.10	0.43±0.05	0.40±0.06	-0.11±0.06	NS
Pe13	31.8±4.94	2±0	1.87±0.07	2.00	0.400	0.368	0.65±0.02	0.46±0.06	0.47±0.02	-0.03±0.15	NS
Pe14	32±4.85	5±0.55	3.24±0.27	5.37	0.151	0.685	1.33±0.08	0.68±0.06	0.69±0.03	0.01±0.07	NS
Pe15	30.8±4.58	6.4±0.40	3.88±0.43	6.43	0.117	0.740	1.49±0.10	0.75±0.02	0.74±0.03	-0.04±0.07	NS
Popeu13	30.8±5.01	7.8±0.49	3.93±0.68	7.36	0.117	0.701	1.57±0.10	0.48±0.10	0.73±0.05	0.33±0.14	***
Grand Mean	31.69±0.95	3.50±0.19	2.26±0.11	-		-	0.81±0.05	0.44±0.03	0.45±0.03	0.02±0.03	

N:Sample size, Na:mean allele number, Ne: effective number of alleles, Ar:allelic richness, PI: Probability of Identity PIC:polymorphism information content, I:Shannon's Information Index, Ho:observed heterozygosity, uHe:unbiased expected heterozygosity, F: inbreeding coefficient, HW:Hardy Weinberg deviation (ND: non-deviating NS: non-significant,***: p<0,001,**: p<0,01,*: p<0,05)

Table 4.5: F-Statistics and estimates of Nm over *Populus euphratica* populations for each locus

Locus	F_{IS}	F_{IT}	F_{ST}	N_m
WPMS5	0.23***	0.29	0.08	2.46
WPMS7	0.71***	0.73	0.07	2.14
WPMS10	0.22	0.24	0.03	5.31
WPMS12	0.28***	0.28	0.00	12.02
WPMS14	-0.07	0.04	0.11	1.65
WPMS15	-0.24	0.10	0.28	0.68
WPMS18	0.03***	0.06	0.02	7.65
WPMS20	Monomorphic			
PMGC14	0.48	0.50	0.03	5.97
PMGC2163	-0.29**	-0.28	0.01	8.63
PMGC2889	0.00	0.04	0.04	4.10
PMGC93	0.07*	0.08	0.02	7.70
Pe2	-0.05	-0.01	0.04	4.33
Pe5	-0.24***	-0.19	0.04	5.45
Pe6	-0.11***	-0.07	0.03	6.18
Pe7	Monomorphic			
Pe8	-0.10	0.11	0.19	1.01
Pe13	0.07	0.10	0.04	6.11
Pe14	0.00	0.06	0.06	3.06
Pe15	-0.01*	0.04	0.05	3.31
Popeu13	0.34***	0.36	0.03	4.35
Mean	0.03	0.09	0.06	4.39±0.68

(HWE deviations ***: p<0,001,**: p<0,01,*: p<0,05)

4.2.2. Population Genetic Diversity

Summary of private alleles for each population are shown in Table 4.6. The mean allelic patterns across populations and their schematic representation are shown in Table 4.7 and Figure 4.1, respectively. The polymorphism information and diversity parameters of the natural populations studied are presented in Table 4.8.

The mean number of alleles per locus (Na) ranged between 3 and 3.81 with a mean value of 3.50±0.19 and effective number of alleles was high with regard to number of alleles, with a mean of 2.26±0.11. There were no private alleles observed for Silifke

population and the highest number of private allele was occured in Euphrates population over all loci (0.429±0.15).

Average percentages of polymorphic loci among populations was 84.76%±1.78%, with highest value observed in Euphrates population. Observed heterozygosity (Ho) was lower than unbiased expected heterozygosity for all populations except Euphrates (Ho:0.50±0.07, uHe:0.49±0.06), however, these values were explicitly lower. F_{ST} value, which is the overall differentiation among populations was 0.075±0.014.

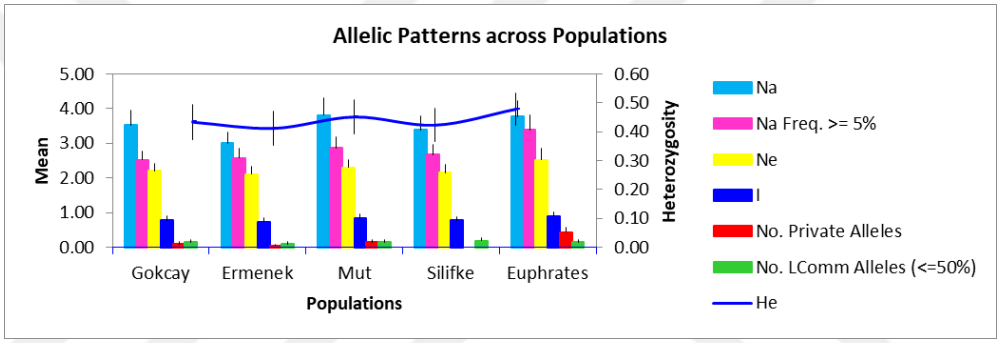

Figure 4.1: Schematic representation of mean allelic patterns and expected heterozygosity across populations (Na: No. of Different Alleles, Na (Freq≥5%): No. of Different Alleles with a Frequency≥ 5%, Ne: No. of Effective Alleles, I: Shannon's Information Index, No. LComm Alleles (≤50 %):No. of Locally Common Alleles (Freq.≥5%) Found in 50% or Fewer Populations He: expected heterozygosity)

Table 4.6: Summary of private alleles by population and locus

Population	Locus	Allele	Frequency
Gokcay	WPMS12	186	0.014
Gokcay	PMGC2163	205	0.011
Ermenek	PMGC14	200	0.065
Mut	WPMS12	175	0.013
Mut	WPMS15	179	0.026
Mut	Pe2	97	0.013
Euphrates	WPMS10	246	0.075
Euphrates	WPMS10	248	0.025
Euphrates	PMGC14	188	0.025
Euphrates	PMGC2163	203	0.125
Euphrates	Pe2	111	0.150
Euphrates	Pe2	124	0.050
Euphrates	Pe5	176	0.025
Euphrates	Pe14	154	0.025
Euphrates	Pe15	153	0.050

Table 4.7: Mean allelic patterns across populations

Population	Na	Na Freq. ≥ 5%	Ne	I	No. Private Alleles	No. LComm Alleles (≤50%)	Ho	uHe
Gokcay	3.52±0.42	2.52±0.25	2.21±0.21	0.78±0.11	2	0.143±0.08	0.41±0.07	0.44±0.06
Ermenek	3±0.33	2.57±0.28	2.10±0.22	0.73±0.11	1	0.095±0.07	0.40±0.07	0.42±0.06
Mut	3.81±0.50	2.86±0.33	2.29±0.23	0.84±0.12	3	0.143±0.08	0.43±0.07	0.46±0.06
Silifke	3.38±0.42	2.67±0.29	2.15±0.22	0.78±0.11	0	0.190±0.09	0.43±0.07	0.43±0.06
Euphrates	3.76±0.46	3.38±0.43	2.53±0.34	0.90±0.12	9	0.143±0.08	0.50±0.07	0.49±0.06

Na:No. of Different Alleles, Na (Freq≥5%):No. of Different Alleles with a Frequency≥ 5%, Ne:No. of Effective Alleles, I:Shannon's Information Index, No. LComm Alleles (≤50%) :No. of Locally Common Alleles (Freq.≥5%) Found in 50% or Fewer Populations Ho:observed heterozygosity, uHe:unbiased expected heterozygosity

Table 4.8: Genetic diversity parameters of *Populus euphratica* populations

Population	N	Na	Ne	I	%P	G-W index (M)	Ho	uHe	F	F_{ST}	
Gokcay	45.38±0.53	3.52±0.42	2.21±0.21	0.78±0.11	85.71%	0.394± 0.133	0.41±0.07	0.44±0.06	0.087±0.08	0.031	0.075*
Ermenek	22.81±0.11	3±0.33	2.10±0.22	0.73±0.11	80.95%	0.383± 0.165	0.40±0.07	0.42±0.06	0.021±0.08		
Mut	38.71±0.24	3.81±0.50	2.29±0.23	0.84±0.12	85.71%	0.387± 0.126	0.43±0.07	0.46±0.06	0.049±0.07		
Silifke	31.76±0.15	3.38±0.42	2.15±0.22	0.78±0.11	80.95%	0.375± 0.126	0.43±0.07	0.43±0.06	0.006±0.07		
Euphrates	19.81±0.11	3.76±0.46	2.53±0.34	0.90±0.12	90.48%	0.363± 0.135	0.50±0.07	0.49±0.06	-0.037±0.08		
Mean	31.70±0.95	3.50±0.19	2.26±0.11	0.81±0.05	84.76%±1.78%	-	0.44±0.03	0.45±0.03	0.025±0.03		

N: Sample size, Na: mean allele number, Ne: effective number of alleles, I: Shannon's Information Index, %P: Percentage of Polymorphic loci, G-W: Garza-Williamson, Ho: observed heterozygosity, uHe: unbiased expected heterozygosity, F: the fixation index over loci F_{ST}: the proportion of the diversity in the sample that's due to allele frequency differences among populations *: F_{ST} over all populations

Garza and Williamson (2001) showed that M, the mean ratio of the number of alleles to the range in allele size calculated from a population sample of microsatellite loci can be used to detect past reductions in population sizes. Thus, to test for any possible bottlenecks occurred in the populations, Garza-Williamson index, M values were calculated for populations and shown in Table 4.8. Also, the Garza-Williamson indices for each polymorphic locus were presented in a bar graph shown in Figure 4.2. According to the critical values, if M<0.68, it indicates a bottleneck and if M>0.80, it indicates no reduction of effective population size. Calculated M values ranged between 0.363 and 0.394 (Table 4.8), which are much smaller than the critical value 0.68 for all of the populations in the study.

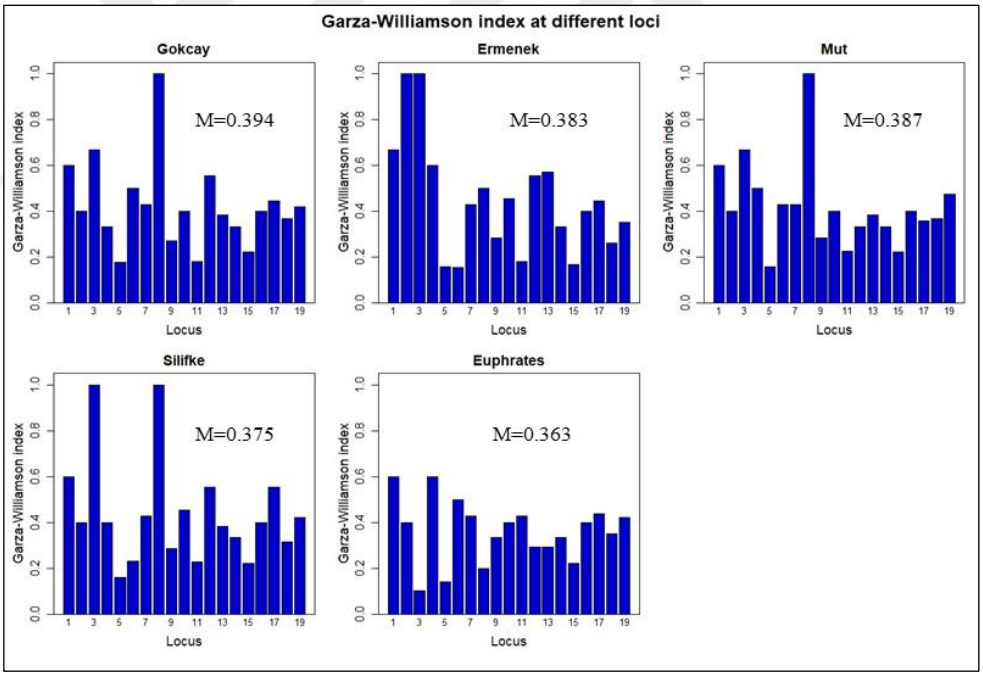

Figure 4.2: The Garza-Williamson index at polymorphic loci in each population

4.2.2.1. Analysis of Molecular Variance (AMOVA)

AMOVA was done as implemented in Arlequin ver 3.5.2.2 (Excoffier and Lischer 2010). AMOVA was carried out with F_{ST}-based on the infinite allele model (using the number of different alleles) and Rst- based on the stepwise mutation model (using the sum of squared size difference) were shown in Table 4.9.

On the one hand, F_{ST} based AMOVA of the two rivers showed that only 11.73 % of the variation is significantly due to the differences among the Göksu and the Euphrates rivers (F_{CT}= 0.117, p=0.000) and only 1.87% of the variation is significantly attributed to differences among populations within river basins (F_{SC}=0.021, p=0.000). Of the total genetic variance, 86.40 % is explained significantly by the differences among individuals within populations (F_{ST}=0.136, p=0.000).

On the other hand, Rst based AMOVA of the two rivers showed different partitioning of the variation with varied percentages, but with similar results. 28.57% of the variation is significantly attributed to the differences among Goksu and Euphrates rivers (F_{CT}= 0.286, p=0.000) and only 1.45% of the variation is significantly de to the differences among populations within river basins (F_{SC}=0.020, p=0.000). Of the total genetic variance, 69.98% is explained by the differences among individuals within populations and it is significant (F_{ST}=0.300, p=0.000).

Table 4.9: Analysis of molecular variance (AMOVA) for *Populus euphratica* using twenty-one microsatellite loci between two groups/regions (The Goksu river and the Euphrates river)

	Source of variation	Sum of squares	Variance components (Va, Vb, Vc)	Percentage of variation	Fixation indices (F_{CT}, F_{SC},F_{ST})
Based on F_{ST}	Among rivers	53.227	0.63541	11.73	0.117***
	Among populations within rivers	34.741	0.10132	1.87	0.021***
	Within Populations	1458.036	4.67866	86.40	0.136***
	Total	1546.004	5.41539	-	-
Based on Rst	Among rivers	9768.124	131.61214	28.57	0.286***
	Among populations within rivers	2314.126	6.68287	1.45	0.020***
	Within Populations	100529.858	322.38320	69.98	0.300***
	Total	112612.109	460.67820	-	-

(***:p<0,001,**:p<0,01,*:p<0,05)

4.3. Population Genetic Structure

4.3.1. Pairwise F_{ST} matrix and Principal Coordinate Analysis

Pairwise F_{ST} values were calculated and presented in Table 4.10 and Figure 4.3. It is clear that Euphrates population is highly differentiated from the Göksu river populations. Among the Göksu river populations, the most distant one is Ermenek population.

Table 4.10: Pairwise F_{ST} matrix of populations with corresponding Nm in parenthesis

	Gokcay	**Ermenek**	**Mut**	**Silifke**	**Euphrates**
Gokcay	-				
Ermenek	0.033*** (7.33)	-			
Mut	0.012*** (20.04)	0.035*** (6.79)	-		
Silifke	0.024*** (10.04)	0.020*** (12.35)	0.016*** (15.13)	-	
Euphrates	0.141*** (1.52)	0.157*** (1.34)	0.119*** (1.85)	0.135*** (1.61)	-

(***: p<0,001, **: p<0,01, *: p<0,05)

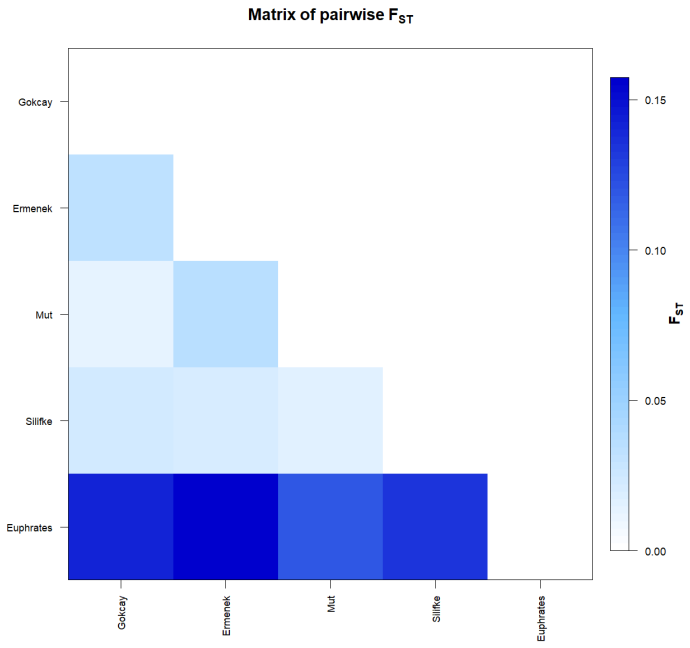

Figure 4.3: Schematic representation of pairwise F_{ST} matrix among populations.

Based on the Pairwise F_{ST} matrix, Principal Coordinates analysis was performed with GenAlEx 6.503 (Peakall and Smouse 2012) and shown in Figure 4.4. The PCoA revealed overall differences among populations. The first principal component, clearly discriminating the two rivers, explained 64.57% of the variance and the second principal component explained 19.40% of the variation among populations of the rivers.

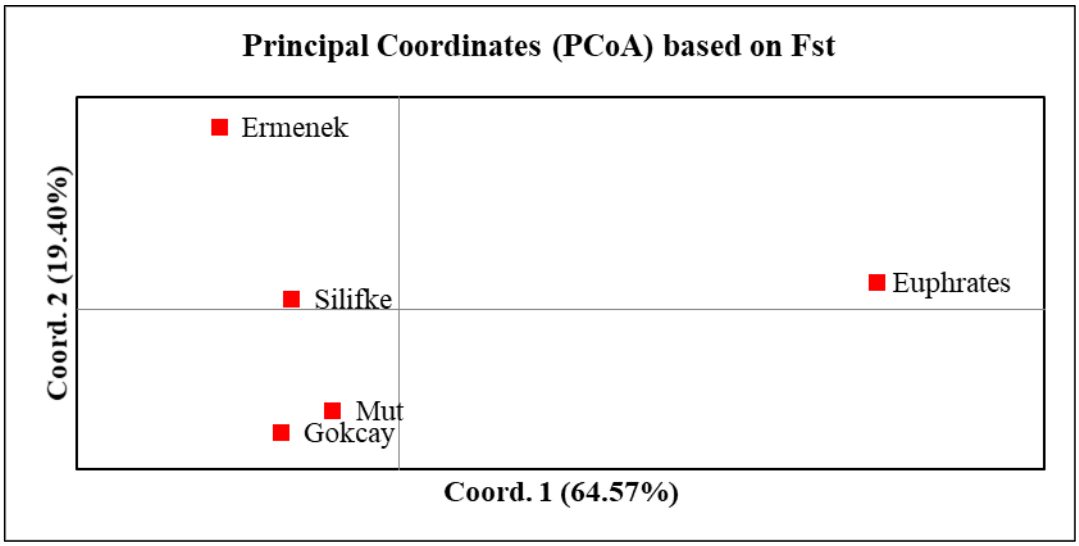

Figure 4.4: Principal component analysis based on population pairwise F_{ST} values

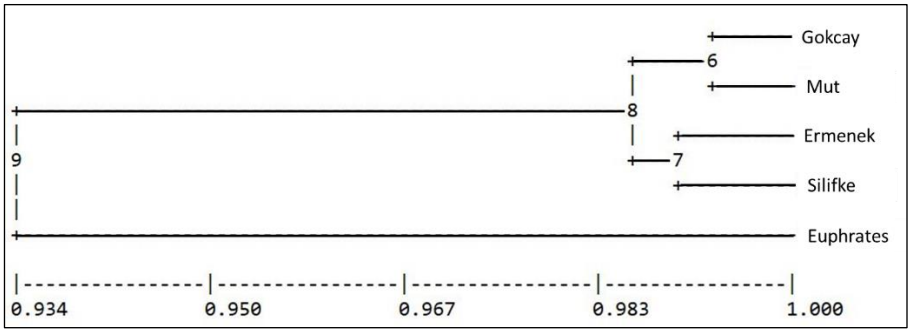

Figure 4.5: UPGMA tree of all populations

4.3.2. Phenetic relationships among populations

A phenogram showing the relationships between the populations was drawn based on coancestry identity with UPGMA method (Figure 4.5). It is evident that Euphrates population is distant to Göksu populations, among which Silifke and Ermenek being the closest ones to Euphrates, as shown in PCoA.

4.3.3. Clustering patterns

Further genetic structure analysis was done with STRUCTURE v2.3.4 (Pritchard et al. 2000) which is implementing a Bayesian iterative algorithm and assigning individuals into clusters to identify genetic substructures. Genetic structure patterns of populations were assesed with prior information about the localities of individuals and without prior information. Both analysis gave similar results, thus, only the results for prior information given data were shown.

First, genetic structure of Göksu populations was determined. The web-based Structure Harvester program identified true K with highest ΔK and lowest standard deviation as three with the Evanno Method (Evanno et al. 2005) (Table 4.11 and Figure 4.6). The clustering patterns for individuals and populations are shown in Figure 4.7. Assignment of individuals showed three genetic clusters representing three separate gene pools. However, it is evident that there is high amount of admixture in Göksu populations.

Table 4.11: Evanno table showing ΔK values for clustering analysis of Göksu populations

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	10	-4876.670000	0.170294	—	—	—
2	10	-4863.180000	12.193332	13.490000	24.630000	2.019957
3*	10	-4825.060000	18.337890	38.120000	117.730000	6.420041
4	10	-4904.670000	24.843244	-79.610000	71.990000	2.897770
5	10	-5056.270000	53.515482	-151.600000	—	—

ΔK calculated as ΔK = m|L″(K)|/ s[L(K)] (Evanno et al. 2005, Appendix B). The true K is shown highlighted with asterisk (*)

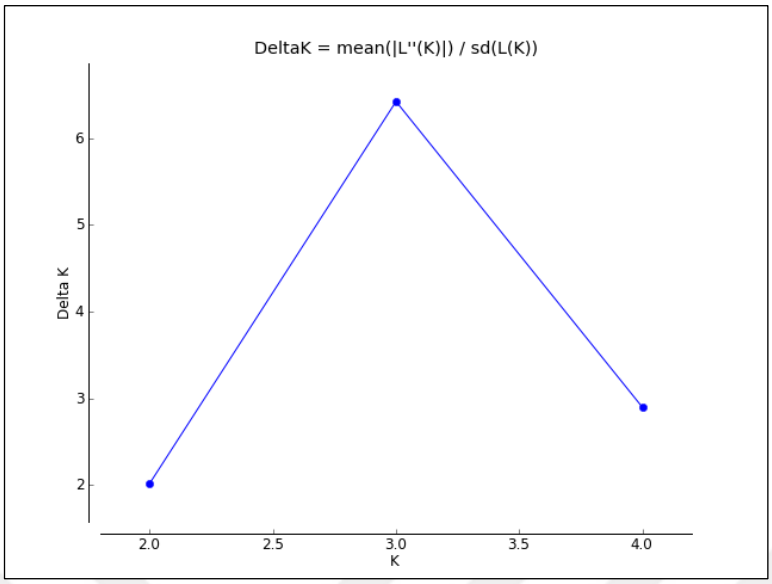

Figure 4.6: The estimated K value by delta-K method of Evanno et al. (2005) for Göksu populations with prior information given

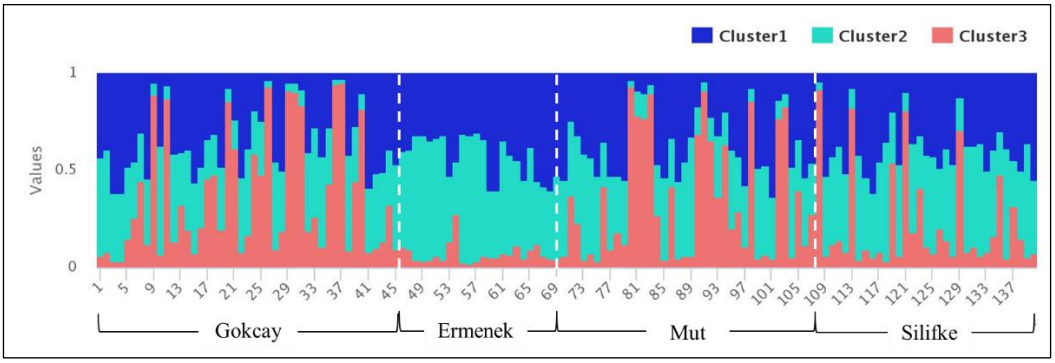

Figure 4.7: Genetic STRUCTURE analyses for 140 *Populus euphratica* genotypes in the Göksu river (K=3). Each color represents a different cluster shown in the figure with prior information given.

Then, genetic structure of Göksu populations were determined together with the population from the Euphrates river. True K with highest ΔK and lowest standard deviation was determined as 3 again with the Evanno Method (Evanno et al. 2005) (Table 4.12 and Figure 4.8). The clustering patterns for individuals and populations are shown in Figure 4.9 and the corresponding membership groups are given in Table 4.13. Assignment of individuals showed 3 genetic clusters representing 3 separate gene pools. Euphrates population was the most distant one among all populations having a homogeneous gene pool as distinctly seen in Figure 4.9. Göksu populations possess 3 different gene pools that is evident from the three major clusters observed, but the gene pool which is evidently coming from Euphrates population is rather ineffective. Among the Göksu river populations, individuals from Gokcay population displayed the highest admixture as clearly observed in the figure. The admixture analysis results were in accordance with the Principal coordinate analysis for the five populations under consideration (Figure 4.4).

Table 4.12: Evanno table showing ΔK values for clustering analysis of Göksu and Euphrates populations

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	10	-5948,42	0.234758	—	—	—
2	10	-5753,33	4,734049	195,090000	60,680000	12,817780
3*	10	-5618,92	1,135097	134,410000	93,820000	82,653753
4	10	-5578,33	5,250196	40,590000	0,380000	0,072378
5	10	-5538,12	3,461470	40,210000	39,790000	11,495115
6	10	-5537,70	7,686207	0,420000	43,590000	5,671198
7	10	-5580,87	29,361276	-43,170000	12,030000	0,409723
8	10	-5612,01	44,711034	-31,140000	17,860000	0,399454
9	10	-5661,01	54,598992	-49,000000	43,670000	0,799832
10	10	-5753,68	73,195489	-92,670000	—	—

ΔK calculated as ΔK = m|L″(K)|/ s[L(K)] (Evanno et al. 2005). The true K is shown highlighted with asterisk (*)

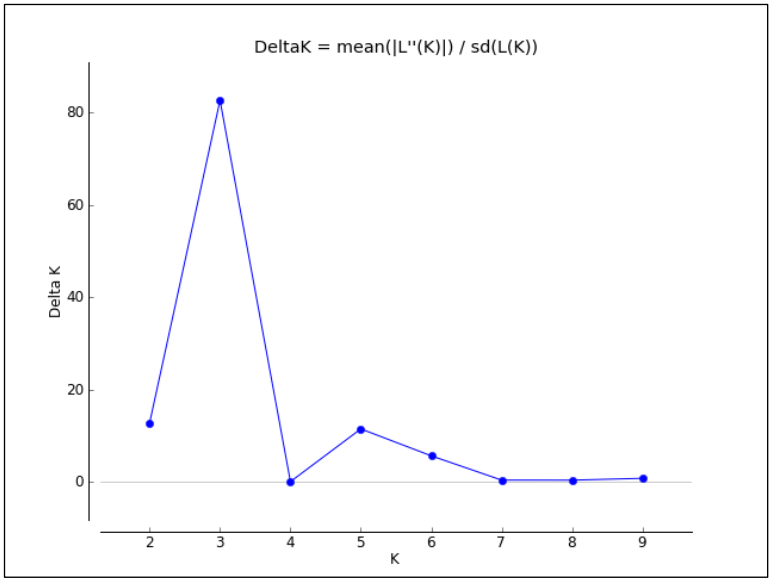

Figure 4.8: The estimated K value by delta-K method of Evanno et al. (2005) for Göksu and Euphrates populations with prior information given

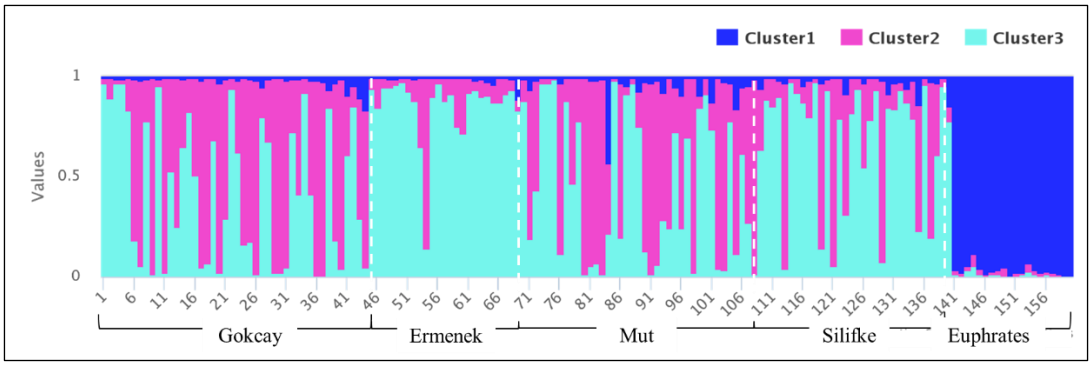

Figure 4.9: Genetic STRUCTURE analyses for 160 *Populus euphratica* genotypes of four Göksu and one Euphrates populations (K=3). Each color represents a different cluster shown in the figure with prior information given.

Table 4.13: The inferred membership values of the 160 *Populus euphratica* genotypes sampled from five populations in Göksu and the Euphrates rivers

Clusters / Populations	Cluster 1			Cluster 2			Cluster 3		
	<0.5	0.5-0.800	0.801-0.99	<0.5	0.5-0.800	0.801-0.99	<0.5	0.5-0.800	0.801-0.99
Gokcay	46	-	-	23	7	16	23	10	13
Ermenek	23	-	-	21	-	2	1	3	19
Mut	39	-	-	18	9	12	22	7	10
Silifke	32	-	-	25	3	4	7	8	17
Euphrates	-	-	20	20	-	-	20	-	-

CHAPTER 5

DISCUSSION

In this thesis the current genetic diversity and genetic structure of Euphrates poplar in the Göksu river populations were studied and presented together with one population from the Euphrates river. The results obtained in this study are unique and precious in that it is the first population genetic study conducted with Euphrates poplar populations in Turkey. Analysis of microsatellite data revealed the genetic potential that the species possess and will provide fundamental information for future conservation studies.

5.1. Genetic diversity of the microsatellite loci used in the study

The microsatellite markers used in this study are all with *Populus* origin but only markers encoded with "Pe" prefix were generated from the species itself (species-specific markers). WPMS and PMGC are cross-species markers that were developed in *Populus nigra* and *Populus trichocarpa*, respectively. While deciding on WPMS and PMGC loci, we paid attention them to be representatives of whole genome, thus tried to choose loci on different chromosomes. For the species-specific loci there is no information about their chromosome locations.

Among the analyzed 21 microsatellite loci, WPMS 20 and Pe7 are monomorphic for the populations under study. These are polymorphic loci in literature with high number of alleles, thus the reason of monomorphism could be the low number of individuals in this study. Moreover, based on the geographical region that the sampling was done, there could be primer binding site mutations in flanking sequences of the repeat regions which can prevent amplification. For the polymorphic markers the allele sizes are generally in consistence with the literature.

Probability of identity (PI) is demonstrating that two unrelated individuals randomly drawn from a population will have the same genotype at multiple loci (Waits et al.

2001). WPMS10, WPMS15 and PMGC14 have high PI values together with low PIC (Polymorphism Information Content) and low I (Shannon's information index) values meaning that they are not sufficiently informative for analyzing the genetic diversity if they are used one by one. The resolution of rest of the markers is high with sufficiently low PI values. This is also in accordance with their PIC values indicating acceptable discrimination power. Thus, these markers could be considered as informative for *P. euphratica* populations.

Genetic variability for loci studied is represented by I values, which was moderate with the mean value of 0.81±0 .05. This moderate value of Shannon's index represents the effectiveness of markers used to reveal the variation present in Euphrates poplar populations.

Some authors argued that allelic richness may indicate populations' long-term potential more effectively than heterozygosity does (Allendorf 1986; Petit et al. 1998). Thus, we prefer to use both Ar and heterozygosity values for the loci studied. Among the polymorphic markers, WPMS14, PMGC2889, PMGC93, Pe2, Pe5, Pe14, Pe15 and Popeu13 have higher Ar values (>3.0). These loci also have the highest He and PIC values. Consequently, those aforementioned markers are the most informative and effective ones with high polymorphism and heterozygosity.

Expected heterozygosity is an important measurement of genetic diversity. Heterozygosity values obtained in this study are substantially lower than the ones found in literature (Wang, Li et al. 2011; Wang, Wu et al. 2011; Xu et al. 2013). Among nineteen polymorphic loci, eleven of them have uHe values higher than Ho and five of those loci are showing significant deviation from HWE. The uHe is lower than Ho in remaining eight loci, of which five of them are showing significant deviation from HWE. The cause for departures from the Hardy-Weinberg equilibrium could be anthropogenic pressures on the habitats. The human mediated practices, such as constructions of dams, habitat destructions for agriculture, urbanization lead to fragmentation and so nonrandom mating as well as reduced effective population sizes.

5.2. Genetic diversity and the population structure of *Populus euphratica*

Breeding new variants with adaptive traits is one of the main goals of plant breeding studies, which are mainly relying on the existing sources of variation in natural populations. Thus, investigating magnitude and structure of genetic diversity will produce information needed for preserving and use of gene resources to conserve, manage and utilize Euphrates poplar gene resources efficiently.

Average number of alleles and mean effective number of alleles per locus are low as compared to similar studies (Wang, Li et al. 2011). In addition, observed and expected heterozygosity values in the current study are substantially low compared to literature (Eusemann et al. 2009; Wang, Li et al. 2011), indicating highly reduced level genetic of diversity. Similar lower heterozygosity values were also found in other poplar species (Du et al. 2012; Lexer et al. 2005; Namroud et al. 2005). Critically low genetic diversity is pointing out gene pool shrinkage of Euphrates poplar populations which are under the threat of collapsing. Considering low number of alleles observed for the species, there could be two possible explanations for gene pool shrinkage. First, although microsatellite markers are neutral in their nature, some of them could have selectively significant function and be under the effect of natural selection (Li et al. 2004). Stabilizing selection or directional selection may have occurred and eliminated different alleles. However, we do not have any convincing proof for this scenario. The second possible explanation is that Euphrates poplar populations have experienced a bottleneck, retaining small number of alleles. This is consistent with the Garza-Williamson index values of the populations, which are indicating severe past reductions in population sizes. Garza and Williamson (2001) argued that following a reduction in population size, genetic drift will be in action and the rare alleles are lost by drift more often than common alleles. This can also result in HWE deviations, which are observed for ten out of twenty-one loci (Table 4.5). Bottlenecks can further increase rate of inbreeding and loss of genetic variation, which in turn can reduce adaptive potential of populations and result in high probability of extinction in the future.

Private and rare alleles are crucial to adapt future environmental changes. They are more prone to decrease in number than reduction in level of heterozygosity (Luikart and Cornuet 1998). Private alleles are also strongly influencing Ar estimates. Among

Göksu populations number of private alleles are quite low, in fact there is none in Silifke population. This is most probably due to high admixture because it is the downstream part of the river. The highest number of private alleles were observed in Euphrates population. The presence of high amount of private alleles in Euphrates nominates it as an effective candidate of possible genetic resource for conservation in future with great adaptation capacity.

In Göksu river populations, positive F_{IS} values are indicating the occurrence of inbreeding, which is a result of mating of related individuals. This could be explained by the fact that at some point in past there occurred intensive reproduction by root suckers because germination and survival of the seedlings were failed due to flow declines. However, the situation is better for Euphrates population with excess of heterozygotes.

In literature, F_{ST} values are low as a characteristic for Salicaceae family members. For Euphrates poplar, Wang, Li et al. (2011) showed F_{ST} as 0.093 with high level of diversity for distant populations. Our results showed low differentiation among Göksu populations (F_{ST}= 0.031). This is also supported by the AMOVA results. Only 1.87 percent of the total variation is attributed to among population diversity (F_{SC}= 0.021). When Euphrates population is included in the analysis, differentiation increased slightly (F_{ST}= 0.075). It is also obvious from the number of migrants between the pairs of populations, which is inversely proportional with F_{ST} values. These results showed clearly that Euphrates population is highly differentiated from the Göksu river populations. Actually, populations from different river basins could be genetically similar if geographical barriers that are preventing gene flow are absent. However, the presence of the Taurus mountains and Anatolian diagonal between Göksu and the Euphrates rivers generates a geographical barrier between the rivers that results in generation of different gene pools. The low differentiation in the Göksu river basin is an expected situation because of wind-pollinating, outcrossing and clonal reproduction nature of the species throughout the river basin. Pollens could be dispersed over long distances, so that, gene flow is high in those populations with no geographical barrier, which in turn reducing the interpopulation differentiation as it is the case in Göksu populations. When Principal Coordinate Analysis was performed with pairwise F_{ST} values first principal component, clearly discriminated Euphrates population and explained 64.57% of the variance. In addition, Ermenek is the most distant one among

the Göksu river basin. This is presumably because there is a slight elevation difference between Ermenek and other three population localities, causing a climatic barrier. Also Gezende dam located between Ermenek population and Mut provides a geographical barrier hindering gene flow.

Population clustering analysis showed that there is a lack of genetic substructure in the Göksu river basin but the individuals are assigned to have three separate gene pools. In consistence with the PCoA, Ermenek is differentiated with low admixture. High admixture in downstream population Silifke was expected to see because of potential gene flow via pollen and seeds along with river flow and possible clonal reproduction. However, Euphrates poplar does not reproduce via broken branches carried with river flow, rather prefers suckers for clonal reproduction. In addition, there is different seedling establishment success at different population sites due to variation in flooding, so this could have probably diminished diversity in downstream part of the Göksu river. The river dynamics in junction of two tributaries of Göksu, where Mut population is located, provide a large area suitable for colonization of the species, that is, the river bed is extending with meanders in river which are suitable for germination and seedling establishment during flooding. Thus, highest intensity of admixture is observed in Mut population. In addition, high admixture is observed due to human effect in Mut together with Gökcay population. Indeed, these populations inhabit the locations where the most intensive farming is practiced.

Based on the results, one can argue that there are no barriers to gene flow in Göksu river basin and Göksu populations form one large metapopulation, which is groups of local populations/subpopulations in a patchy environment. This is also supported by the studies of Eusemann et al. (2013) and Wang, Li et al. (2011). Local extinctions and recolonization events occur frequently in metapopulations (Harrison and Taylor 1997). However, habitat fragmentations due to human activities as present in Göksu river basin can hinder migration/gene flow between subpopulations, thus could prevent recolonization and result in population extinctions. Consequently, it is urgent to take actions against habitat destructions and for conservation strategies.

When Euphrates population is included, population structure analyses support the presence of two genetically separated groups. Under the circumstances, we could have two possible scenarios explaining the clustering patterns. The first one is that the founder population of the Göksu river is coming from the Euphrates, as it is explicit

that Euphrates individuals are possessing the third gene pool of Göksu homogeneously. However, for this possible explanation to be acceptable, gene pool coming from the founder population should be represented with higher membership values in all Göksu populations, yet it is not. The second and the more probable scenario is that the populations of Euphrates poplar in two river ecosystems have two different founder populations possessing very similar genotypes few in number and there occurred strong inbreeding between genetically related individuals. This is also explaining the low heterozygosity values compared to literature. In both cases, the two river populations of Euphrates poplar may have been evolving as two different gene pools that have experienced little contact through low level of gene flow. Then, in the Göksu river, populations have further evolved to have two other gene pools with high admixture.

5.3. Conservation of *Populus euphratica* Genetic Resources

The disappearance of natural Euphrates poplar genetic resources is a notable indicator of habitat degradation and habitat lost for the particular river ecosystem and thus proper measures should be taken to protect them.

One of the major problems in preserving natural populations is human ignorance, that is the species is not considered to be a forest tree in Turkey. Since not directly used as a timber source, it is underestimated by forestry officers. Research should be directed to inform public about the value of the species. In addition, a conservation management plan should be initiated, including strategies to implement general policies and to establish objectives for the preservation of natural resources. Implementing a conservation plan is a collective responsibility of rural people, government and stakeholders. Thus, a major attempt should be educational workshops to inform public and enhance communication between the partners of the plan.

Since it is a riparian species, reduced water flow is another major drawback in *P. euphratica* forests. Access to groundwater ensures continuous growth for Euphrates poplar trees and triggers rejuvenation in disturbed habitats because germination and seedling growth depends heavily on wet and exposed sites found at riverbanks (Westermann et al. 2008). Ling et al. (2015) found that flooding should happen two to three times per year with a duration of 15-20 days and an intensity of 25-30 m^3/s for

regenerating *P. euphratica* forests. This is necessary for both seed germination and seedling growth. Moreover, groundwater table is increasing by these floodings, which is crucial because as the groundwater depth increases, survival and growth rates of *P. euphratica* decline (Xu et al. 2009). In field studies of this thesis, it is observed that many tributaries of the Göksu river in the watershed were dried out. Thus, planned floodings that are coinciding with seed germination and seedling emergence will be important for regeneration of the populations as an *in situ* conservation strategy for the Göksu river basin. This could be achieved by releasing water from the dams in a controlled manner. A successful application of this strategy is applied in the Tarim river basin in China as a restoration program by water diversions (Aishan et al. 2015).

Conservation of genotypes by providing identified material for planting, breeding and further experimental research is also crucial. The results of this study demonstrated that Euphrates population has the potential to be an important resource for *ex situ* conservation programs of Euphrates poplar. Gene banks and clone banks are promising for future research and breeding experiments. An effective example *ex situ* conservation program for poplars in Turkey is the one initiated for European black poplar by Poplar and Fast-Growing Forest Trees Research Institute. European black poplar clone banks were established as a part of *ex situ* conservation programme. This could also be achieved with Euphrates poplar and replicated gene banks or clone banks would be promising in order to preserve germplasm of the species.

CHAPTER 6

CONCLUSION

Populus euphratica is a pioneer species for river ecosystems in south and southeastern regions of Turkey. It has substantial potential for future breeding programmes especially aiming to develop drought and salt stress resillient genotype plantations on dry and saline sites.

Results showed that using species-specific markers impact the estimates of genetic diversity. Pe loci, which have shown the highest Ar, He and PIC values, are consequently the most informative and effective markers. Thus, we recommend to use species-specific Pe loci to assess genetic diversity and population structure of Euphrates poplar.

Population structuring showed that there were two different founder populations for the Göksu and the Euphrates rivers. Besides, Euphrates poplar populations had experienced bottlenecks which caused severe past reductions in population sizes and increased inbreeding. Although observed and expected heterozygosity values did not have remarkable difference, they indicated reduced level of genetic diversity when compared to other poplar species. In addition, populations have low number of private alleles except Euphrates population. In future, the populations could survive in current habitats if there will be no human intervention. However, it is well known that the loss of genetic diversity is not desirable because it reduces the ability of species to cope with environmental changes.

As a consequence, the results demonstrated that native populations of this species in the Göksu river are in great danger of extinction mainly because of human mediated practices. Possessing high number of private alleles and slightly higher heterozygosity values make Euphrates population be an effective candidate for genetic resource conservation programs.

Likewise, it is urgent to start *in situ* conservation of Euphrates poplar stands especially in the Göksu river. In situ conservation on natural reserves seems to be easier, cheaper and sustainable than running breeding programs for the species at first, however, public awareness and strict management for agricultural areas are required for the conservation program to be successful. For the Göksu river, in particular, reintroduction of the species is not necessary at present, thus in situ conservation would be straightforward and promising if initiated immediately.

The conservation of these genetic resources is of great importance and priority should be given for fast and reproducible actions. Trade-offs between biodiversity conservation and economic growth is challenging but governments should be control centers for implementing politically, socially and economically practicable solutions to ensure sustainable conservation plans.

REFERENCES

Acar O (1973) *Populus euphratica* Oliv. odununun anatomik ve teknolojik özellikleri. Kavak ve Hızlı Gelişen Orman Ağaçları Araştırma Enstitüsü, Yıllık Bülten 8:1-177

Acatay G (1961) *Populus euphratica* Oliv. İstanbul Üniversitesi Orman Fakültesi Dergisi 11(2): 47-50

Aishan T, Halik Ü, Kurban A, Cyffka B, Kuba M, Betz F, Keyimu M (2015) Eco-morphological response of floodplain forests (*Populus euphratica* Oliv.) to water diversion in the lower Tarim River, northwest China. Environmental Earth Sciences 73 (2): 533-545

Allendorf FW (1986) Genetic drift and the loss of alleles versus heterozygosity. Zoo Biology 5:181–190

Arnaud-Haond S, Belkhir K (2007) GENCLONE: A computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. Molecular Ecology Notes 7(1): 15–17

Braatne JH, Rood SB, Heilman PE (1996) Life history, ecology and conservation of European seahorses. In: Stettler R, Bradshaw T, Heilman P, Hinckley T (ed) Biology of *Populus* and Its Implications for Management and Conservation. Ottowa, NRC Research Press, pp 57–86

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5: 453–455

Brosché M, Vinocur B, Alatalo ER et al (2005) Gene expression and metabolite profiling of *Populus euphratica* growing in the Negev desert. Genome biology 6:R101

Browicz K (1977) Chorology of *Populus euphratica* Oliver. Arboretum Kornickie 22: 5-26

Browicz Z, Yaltırık F (1982) *Populus* L. In: Davis P H (ed) Flora of Turkey and the East Aegean Islands Volume 7. Edinburgh, University Press, pp 716–720

Bruelheide H, Manegold M, Jandt U (2004) The genetical structure of *Populus euphratica* and *Alhagi sparsifolia* stands in the Taklimakan desert. In: Runge M,

Zhang X (ed) Ecophysiology and Habitat Requirements of Perennial Plant Species in the Taklimakan Desert. Shaker Verlag GmbH, Germany, pp 153-160

Calagari M, Mofidabadi AJ, Tabari M, HosseiniSM (2004) Intraspecific Hybridization of *Populus euphratica* Oliv . using in vitro Technique. Journal of Sciences, Islamic Republic of Iran 15(2): 109-112

Carlsson J (2008) Effect of microsatellite null alleles on assignment testing. Journal of Heredity 99: 616-623

Chakraborty R, De Andrade M, Daiger SP, Budowle B (1992) Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Annals of Human Genetics 56: 45-57

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution 24(3): 621-631

Chase MW, Zmarzty S, Lledó D, Wurdack K, Swensen SM, Fay MF (2002) When in doubt, put it in Flacourtiaceae: a molecular phylogenetic analysis based on plastid rbcL DNA sequences. Kew Bull. 57:141-181

Chase MR, Kesseli R, Bawa K (1996) Microsatellite Markers for Population and Conservation Genetics of Tropical Trees. American Journal of Botany 83(1):51-57

Chen S, Li J, Wang S, Huttermann A, Altman A (2001) Salt, nutrient uptake and transport, and ABA of *Populus euphratica*; a hybrid in response to increasing soil NaCl. Trees 15:186-194

Coğrafya Harita Website. Available at http://cografyaharita.com/turkiye_hidrografya_haritalari.html Accessed January 2018

Dabrowski MJ, Pilot M, Kruczyk M, Zmihorski M, Umer HM, Gliwicz J (2014) Reliability assessment of null allele detection: Inconsistencies between and within different methods. Molecular Ecology Resources 14(2): 361-373

Dai X, Hu Q, Cai Q et al (2014) The willow genome and divergent evolution from poplar after the common genome duplication. Cell Res 24(10):1274-1277

Dayanandan S, Rajora OP, Bawa KS (1998) Isolation and characterization of microsatellites in trembling aspen (*Populus tremuloides*). Theor Appl Genet 96:950–956

De Sousa SN, Finkeldey R, Gailing O (2005) Experimental verification of microsatellite null alleles in Norway spruce (*Picea abies* [L.] Karst.): implications for population genetics studies. Plant Molecular Biology Reporter 23:113-119

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society Series B (Methodological) 39:1-38

Dorn RD (1976) A synopsis of American Salix. Canadian Journal of Botany 54: 2769-2789

Doyle J (1991) DNA Protocols for Plants. In: Hewitt GM, Johnston AWB, Young JPW (ed) Molecular Techniques in Taxonomy, NATO ASI Series (Series H: Cell Biology), vol 57. Springer, Berlin, Heidelberg, pp 283-293

Du Q, Wang B, Wei Z, Zhang D, Li B (2012) Genetic diversity and population structure of Chinese white poplar (*Populus tomentosa*) revealed by SSR markers. J Hered 103:853-862

Du FK, Xu F, Qu H, Feng S, Tang J, Wu R (2013) Exploiting the Transcriptome of Euphrates Poplar, *Populus euphratica* (Salicaceae) to Develop and Characterize New EST-SSR Markers and Construct an EST-SSR Database. PLoS ONE 8(4): e61337

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4(2): 359-361

Eckenwalder JE (1996) Systematics and evolution of *Populus*. In: Stettler RF, Bradshaw HD Jr, Heilman PE, Hinckley TM (ed) Biology of *Populus* and Its Implications for Management and Conservation. National Research Council of Canada Research Press, Ottawa, pp 7-32

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. Trends in genetics 16(12):551-558

Ellegren H (2002) Microsatellite evolution: a battle between replication slippage and point mutation. Trends Genet. 18 (2):70

Eusemann P (2010) Population genetics and reproduction biology of *Populus euphratica* OLIV. (Salicaceae) at the Tarim River, Xinjiang Province, NW China. Dissertation, University of Greifswald

Eusemann P, Fehrenz S, Schnittler M (2009) Development of two microsatellite multiplex PCR systems for high throughput genotyping in *Populus euphratica*. Journal of Forestry Research 20(3): 195–198

Eusemann P, Petzold A, Thevs N, Schnittler M (2013) Growth patterns and genetic structure of *Populus euphratica* Oliv. (Salicaceae) forests in NW China-Implications for conservation and management. For. Ecol. Manag. 297: 27-36

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

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564-567

Fay MF, Lledo MD, Kornblum MM, Crespo MB (1999) From the waters of Babylon? *Populus euphratica* in Spain is clonal and probably introduced. Biodiversity and Conservation 8: 769-778.

Field D, Wills C (1998) Abundant microsatellite polymorphism in *Saccharomyces cerevisiae*, and the different distributions of microsatellites in eight prokaryotes and *S. cerevisiae*, result from strong mutation pressures and a variety of selective forces. Proc Natl Acad Sci. USA 95:1647–1652

Fossati T, Grassi F, Sala F, Castiglione S (2003) Molecular analysis of natural populations of *Populus nigra* L. intermingled with cultivated hybrids. Mol Ecol 12: 2033–2043

Francis RM (2016) pophelper: An r package and web app to analyse and visualize population structure. Molecular Ecology Resources 17(1): 27–32

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. Molecular Ecology 10(2): 305–318

Gaudet M, Jorge V, Paolucci I, Beritognolo I, Scarascia Mugnozza G, Sabatti M (2008) Genetic linkage maps of *Populus nigra* L. including AFLPs, SSRs, SNPs, and sex trait. Tree Genetics and Genomes 4: 25-36

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-Statistics. The Journal of Heredity 86(6): 485–486

Greenaway W, Gümüsdere I, Whatley FR (1991) Analysis of phenolics of bud exudate of *Populus euphratica* by GC-MS. Phytochemistry 30 (6):1883–1885

Gries D, Foetzki A, Arndt SK, Bruelheide H, Thomas FM, Zhang X, Runge M (2005) Production of perennial vegetation in an oasis-desert transition zone in NW China-allometric estimation and assessment of flooding and use effects. Plant Ecol 181:23-43

Gu R, Fonseca S, Puskás LG, Hackler L Jr, Zvara A, Dudits D, Pais MS (2004) Transcript identification and profiling during salt stress and recovery of *Populus euphratica*. Tree Physiology 24:265-276

Guo S, Thompson E (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48: 361-372

Gülbaba AG (1991) Fırat kavağı (*Populus euphratica* Oliv.)'nın vejetatif yoldan üretilmesi. Kavak ve Hızlı Gelişen Yabancı Tür Orman Ağaçları Araştırma Enstitüsü Dergisi, 1191: 2 (In Turkish)

Han ZG, Guo WZ, Song XL, Zhang TZ (2004) Genetic mapping of EST-derived microsatellites from the diploid Gossypium arboreum in allotetraploid cotton. Mol Genet Genomics 272: 308-327.

Harrison S, Taylor AD (1997) Empirical evidence for metapopulation dynamics: a critical review. In: Hanski I, Gilpin AD (ed) Metapopulation Dynamics: Ecology, Genetics and Evolution. Academic Press, San Diego, CA, pp 27-42

Hukin D, Cochard H, Dreyer E, Le Thiec D, Bogeat-Triboulot MB (2005) Cavitation vulnerability in roots and shoots: does *Populus euphratica* Oliv., a poplar from arid areas of Central Asia, differ from other poplar species? Journal of Experimental Botany 56: 2003–2010

Isebrands JG, Richardson J (2014) Introduction. In : Isebrands JD, Richardson J (ed) Poplars and Willows: Trees for Society and the Environment. The Food and Agriculture Organization of the United Nations and CAB International, Rome, pp 1-7

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

Kalia RK, Rai MK, Kalia S, Singh R, Dhawan AK (2011) Microsatellite markers: an overview of the recent progress in plants. Euphytica 177: 309-334

Kalinowski ST, Taper ML (2006) Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. Conservation Genetics 7: 991–995

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16(5):1099–1106

Karacadağ Development Agency Website (2014) Fırat Kavağı Doğal Sit Alanı Çalıştay Raporu: Şanlıurfa Birecik Stratejik Bölge Gelişim Planı İçerisinde Fırat Kavağı ve Doğal Sit Alanı. (In Turkish) Available at https://www.karacadag.gov.tr/dosyalar2014/FIRAT%20KAVA%C4%9EI%20DO%C4%9EAL%20S%C4%B0T%20ALANI%20%C3%87ALI%C5%9ETAYI.pdf Accessed December 2017

Karatay H (2003) Fırat Nehri Boyunca Yayılış Gösteren Fırat Kavağı (*Populus euphratica* Oliv.) Populasyonunun Bazı Morfolojik Karakterleri Üzerine Bir Araştırma. Dissertation, Graduate School of Natural and Applied Sciences, Karadeniz Technical University, Trabzon (In Turkish)

Kruglyak S, Durrett RT, Schug MD, Aquadro CF (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutation. Proc Natl Acad Sci USA 95:10774-10778

Legionnet A, Faivre-Rampant P, Villar M, Lefevre F (1997) Sexual and asexual reproduction in natural stands of *Populus nigra*. Botanica Acta 110(3): 257–263

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Lewis PO, Zaykin D (2001) Genetic Data Analysis V1.1. Available at https://phylogeny.uconn.edu/software/ Accessed December 2017

Lexer C, Fay MF, Joseph A, Nica MS, Heinze B (2005) Barrier to gene fow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen) the role of ecology and life history in gene introgression. Mol Ecol 14:1045–1057

Li Y-C, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: structure, function, and evolution. Mol Biol Evol 21:991-1007

Ling H, Zhang P, Xu H, Zhao X (2015) How to regenerate and protect desert riparian *Populus euphratica* forest in arid areas. Sci Rep 5:15418.

Luikart G, Cornuet J-M (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12: 228-237

Mamıkoğlu NG (2007). Türkiye'nin ağaçları ve çalıları. NTV Press, İstanbul (In Turkish)

Marshall TC, Slate J, Kruuk LE, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7(5): 639-655

Mofidabadi AJ, Modir-Rahmati AR (2000) Production of *Populus euphratica* Oliv. x *P. alba* L. hybrid poplars through ovary and ovule cultures. Plant Genetic Resources Newsletter 122:13-15

Mofidabadi AJ, Modir-Rahmati AR, Tavesoli A (1998) Application of ovary and ovule culture in *Populus alba* L. × *P. euphratica* Oliv. hybridization. Silvae Genetica 47: 332-334

Nadir E, Hargalit H, Gallily T, Ben-Sasson SA (1996) Microsatellite spreading in the human genome: Evolutionary mechanisms and structural implications. Proceedings of the National Academy of Sciences USA 93:6470-6475

Namroud MC, Park A, Tremblay F, Bergeron Y (2005) Clonal and spatial genetic structures of aspen (*Populus tremuloides* Michx). Mol Ecol 14:2969-2980

Parida SK, Kalia SK, Sunita K, Dalal V, Hemaprabha G, Selvi A, Pandit A, Singh A, Gaikwad K, Sharma TR, Srivastava PS, Singh NK, Mohapatra T (2009) Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. Theor Appl Genet 118:327-338

Peakall R, Smouse PE (2012) GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19): 2537-2539

Pemberton JMS, Slate J, Bancroft DR, Barrett JA (1995) Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. Molecular Ecology 4: 249-252

Petit R, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12: 844-855

Petzold A, Pfeiffer J, Jansen F, Eusemann P, Schnittler M (2013) Sex ratios and clonal growth in dioecious *Populus euphratica* Oliv., Xinjiang Prov., Western China. Trees 27: 729-744

Powell W, Morgante M, McDevitt R, Vendramin GG, Rafalski JJ (1995) Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. Proc Natl Acad Sci 92: 7759-7763

Prasad M, Varshney RK, Roy JK, Balyan HS, Gupta PK (2000) The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theor Applied Genet 100: 584-592

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and systematic. Trends Ecol Evol 16:142–147

Qiu Q, Ma T, Hu Q et al (2011) Genome-scale transcriptome analysis of the desert poplar, *Populus euphratica*. Tree Physiology 31:452-461

Rajendrakumar P, Biswal AK, Balachandran SM, Srinivasarao K, Sundaram RM (2007) Simple sequence repeats in organellar genomes of rice: frequency and distribution in genic and intergenic regions. Bioinformatics 23: 1-4

Rajora OP, Rahman MH, Buchert GP, Dancik BP (2000) Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (*Pinus strobes*) Ontario, Canada. Mol Ecol 9: 339-348

Rathmacher G, Niggemann M, Kohnen M, Ziegenhagen B, Bialozyt R (2010) Short-distance gene flow in *Populus nigra* L. accounts for small-scale spatial genetic structures: implications for in situ conservation measures. Conserv Genet 11:1327-1338.

Raymond M, Rousset F (1995) Genepop (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism. Journal of Heredity 86: 248-249

Rottenberg A, Nevo E, Zohary D (2000) Genetic variability in sexually dimorphic and monomorphic populations of *Populus euphratica* (Salicaceae). Canadian Journal of Forest Research 30: 482-486

Rousset F (2008) GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8(1):103-106

Saito Y, Shiraishi S, Tanimoto T, Yin L, Watanabe S, Ide Y (2002) Genetic diversity of *Populus euphratica* populations in northwestern China determined by RAPD DNA analysis. New Forests 23(2): 97-103

Shinde D, Lai Y, Sun F,Arnheim N (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. Nucleic Acids Research 31(3):974-980

Skvortsov AK (1999) Willows of Russia and Adjacent Countries. Taxonomical and Geographical Revision (English translation of 1968 Russian edition). University of Joensuu, Joensuu, Finland

Smulders MJM, van Der Schoot J, Arens P, Vosman B (2001) Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). Molec Ecol Notes 1(3): 188-90

Stanton BJ, Villar M (1996) Controlled reproduction of *Populus*. In: Stettler RF, Bradshaw Jr HD, Heilman PE, Hinckley TM (ed) Biology of *Populus* and its Implications for Management and Conservation. NRC Research Press,Ottawa, ON, Canada, pp 113-138

Strand M, Prolla TA, Liskay RM, Petes TD (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature 365(6443):274-276

Tabbener HE, Cottrell JE (2003) The use of PCR based DNA markers to study the paternity of popular seedlings. Forest Ecology and Management 179:363-376

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryote genomes. Nucleic Acids Res 12: 4127-4138

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome 14:1812–1819

The Angiosperm Phylogeny Group (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Botanical Journal of the Linnean Society 141: 399-436

The Forest Tree Genetic Resources Panel of the Food and Agriculture Organization of the United Nations and the Canadian Forest Service. Report of the International Boreal Forest Genetic Resources Workshop. 3–21 (1995) Available at http://www.cfs.nrcan.gc.ca/pubwarehouse/pdfs/10268_e.pdf. Accessed February 2018

The International *Populus* Genome Consortium website Available at: http://www.ornl.gov/sci/ipgc/ssr_resource.htm Accessed 25 December 2017

Thevs N, Zerbe S, Schnittler M, Abdusalih N, Succow M (2008) Structure, reproduction, and flood-induced dynamics of riparian Tugai forests at the Tarim River in Xinjiang, NW China. Forestry 81:45-57

TUBIVES website. Available at: http://www.tubives.com/index.php?sayfa=1&tax_id=8512 Accessed 24 December 2017

Tunçtaner, K (2008) Kavaklarda Genetik Islah ve Seleksiyon. Çevre ve Orman Bakanlığı Kavak ve Hızlı Gelişen Orman Ağaçları Araştırma Enstitüsü, İzmit

Tuskan GA, DiFazio S, Jansson S et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313(5793):1596–1604

UNESCO World Heritage Center Website Available at: http://whc.unesco.org/en/tentativelists/5532/ Accessed 24 December 2017

van de Ven WTJ, McNicol RJ (1996) Microsatellites as DNA markers in Sitka spruce. Theor Appl Genet 93: 613-617

van der Schoot J, Pospiskova M, Vosman B, Smulders MJM (2000) Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.). Theor Appl Genet 101(1-2): 317-22

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535-538

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Mol Ecol 10:249–256

Wang S, Chen B, Li H (1996) Euphrates poplar forest, 1st edn. China Environmental Science Press, Beijing.

Wang C, Liu LX, Liu YG (2015) Development and characterization of 20 polymorphic microsatellite markers from RAPD product in *Populus euphratica*. Conservation Genetics Resources 7(3): 669-671

Wang J, Li Z, Guo Q, Ren G, Wu Y (2011) Genetic variation within and between populations of a desert poplar (*Populus euphratica*) revealed by SSR markers. Annals of Forest Science 68: 1143-1149

Wang J, Wu Y, Ren G, Guo Q, Liu J, Lascoux M (2011) Genetic differentiation and delimitation between ecologically diverged *Populus euphratica* and *P. pruinosa*. PLoS ONE 6(10): 1-10

Weir BS (1996) Genetic Data Analysis II: Methods for Discrete Population Genetic Data. Sinauer Associates, Inc., Sunderland, MA

Westermann J, Zerbe S, Eckstein D (2008) Age structure and growth of degraded *Populus euphratica* floodplain forests in north-west China and perspectives for their recovery. J Integr Plant Biol 50:536-546

Wiehle M, Eusemann P, Thevs N, Schnittleret M (2009) Root suckering patterns in *Populus euphratica* (Euphrates poplar, Salicaceae). Trees 23: 991-1001

Wu Y, Wang J, Liu J (2008) Development and characterization of microsatellite markers in *Populus euphratica* (Populaceae). Molecular Ecology Resources 8(5): 1142-1144

Xu H, Ye M, Li J (2009) The ecological characteristics of the riparian vegetation affected by river overflowing disturbance in the lower Tarim River. Environ Geol 58:1749-1755

Xu F, Feng S, Wu R, Du FK (2013) Two highly validated SSR multiplexes (8-plex) for Euphrates' poplar, *Populus euphratica* (Salicaceae). Molecular Ecology Resources 13: 144-153

Yimit H, Ayup M, Wang GZ, Luo H, Ebeidulla D (2006) P. euphratica ecosystem fragility and protecting strategy on Tarim *P. euphratica* nature reserve in Xinjiang. In: Gao W, Ustin SL (ed) Remote Sensing and Modeling of Ecosystems for Sustainability III. Proceedings of SPIE, 6298

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11:1-16

Zsuffa L, Giordano E, Pryor LD, Stettler RF (1996) Trends in poplar culture: some global and regional perspectives. In: Stettler RF, Bradshaw HD Jr, Heilman PR, Hinckley TM (ed) Biology of *Populus* and Its Implications for Management and Conservation. National Research Council of Canada Research Press, Ottawa, Ontario, pp 515-539

APPENDIX A

CTAB DNA EXTRACTION PROTOCOL

1. 2% (w/v) polyvinylpyrollidine (PVP-40) was added to 2xCTAB extraction buffer (2% (w/v)CTAB, 5M NaCI, 0,5M EDTA, 1M Tris-HCI, pH 8.0) freshly and heat at 65°C in a water bath at least half an hour.
2. 0.1g leaf tissue was weighed, placed in a mortar together with 2000µL preheated CTAB solution and ground until a homogeneous green liquid was obtained.
3. ~1800µL of the liquid was transferred into a 2ml Eppendorf tube. 100µL β-mercaptoethanol and 5µL Proteinase K were added to each tube and incubated in water bath at 65°C for 1 hour with occasional gentle swirling.
4. The tubes were centrifuged at 4°C for 20 minutes at 15000 rpm. Aqueous phase (~800µL) was collected to a new 2 ml Eppendorf tube.
5. 0.8V Chloroform/Isoamylalcohol(24:1) was added, inverted gently a few times and centrifuged at 4°C for 15 minutes at 14000rpm.
6. ~500µL supernatant was taken in a new 1.5mL Eppendorf tube and ice cold Isopropanol was added with 1:1 ratio. Tubes were gently inverted a couple of times, then incubated at -80°C for 1 hour (may be o/n).
7. Samples were centrifuged at 14000rpm for 15 minutes at 4 °C. The supernatant was poured off and the pellet was washed with cold 70% EtOH, twice. (Pellet was spinned down if necessary)
8. The pellet was air dryed and resuspended in 60µL TE buffer o/n.

APPENDIX B

POPULATION GENETICS STATISTICS

Number of different alleles (Na)

Determined by direct count.

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.

Shannon's Information Index (I)

Calculated on a single-locus basis, where ln = the natural logarithm and pi is the frequency of the i^{th} allele. Equivalent to the Shannon-Weaver Index of ecology. Unlike He, not bounded by 1 and may therefore be a better measure of allelic and genetic diversity, though largely overlooked in genetic studies (Sherwin *et al*., 2006). The formula is as follows;

$$I = \sum pi \mathrm{Ln}\, pi$$

pi is the allele frequency of the i^{th} allele at the locus in question for the specified population.

Heterozygosity

Ho is the observed heterozygosity, i.e. the proportion of N samples that are heterozygous at a given locus.

$$\text{Ho} = \frac{Number\ of\ heterozygotes}{N}$$

He is the expected heterozygosity, i.e. the proportion of heterozygosity expected under random mating and *pi* is the allele frequency of the *i*-th allele.

$$\text{He} = 1 - \sum pi^2$$

The Fixation Index F (also called the Inbreeding Coefficient) exhibits values ranging from -1 to +1. 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 selection for heterozygotes.

$$\text{F} = \frac{He - Ho}{He}$$

For codominant genetic data at a single locus, the total genetic diversity (heterozygosity) can be divided into within and among populations as follows:

$\overline{Ho}$ = Observed heterozygosity averaged across subpopulations.

$\overline{He}$= Expected heterozygosity averaged across subpopulations.

H_T = Total expected heterozygosity (calculated as if all the subpopulations were pooled).

$$\overline{Ho} = \sum_{i=1}^{k} Ho/k$$

Where Ho= observed heterozygosity in subpopulation *i*, and *k* is the number of subpopulations

$$He = 1 - \sum_{i=1}^{h} p_{i,s}^2$$

$$\overline{He} = \sum_{i=1}^{k} He/k$$

Where He is the expected heterozygosity within subpopulation *s*, and $p_{i,s}$ is the frequency of the *i*-th allele in subpopulation *s.* The summation of the allele frequency squared is over all *i*-th alleles to *h* the max number of alleles.

$$H_T = 1 - \sum_{i=1}^{h} p_{Ti}^2$$

Where H_T is the total expected heterozygosity, and p_{Ti} is the frequency of allele *i* over the total population. If subpopulation sample sizes are equal then $p_{Ti} = pi$, where p_i is the frequency of allele *i* averaged over the subpopulations of equal size.

Wright's F-statistics (Wright 1946, 1951, 1965)

F_{IS} = The inbreeding coefficient within individuals relative to the subpopulation. It measures the reduction in heterozygosity of an individual due to non random mating within its subpopulation.

$$F_{IS} = \frac{\overline{He} - \overline{Ho}}{\overline{He}}$$

F_{IT} = the inbreeding coefficient within individuals relative to the total. This statistic takes into account the effects of both non random mating within subpopulations and genetic differentiation among the subpopulations.

$$F_{IT} = \frac{H_T - \overline{Ho}}{H_T}$$

F_{ST} = the inbreeding coefficient within subpopulations relative to the total. This statistic provides a measure of the genetic differentiation between subpopulations. That is, the proportion of the total genetic diversity (heterozygosity) that is distributed among the subpopulations. F_{ST} is almost always greater than or equal to zero. If all

subpopulations are in Hardy-Weinberg equilibrium with the same allele frequencies, $F_{ST} = 0$.

$$F_{ST} = \frac{H_T - \overline{He}}{H_T}$$

F-statistics are related according to the following equation:

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

Probability of Identity (PI)

The Probability of Identity PI provides an estimate of the average probability that two unrelated individuals, drawn from the same randomly mating population, will by chance have the same multilocus genotype. Also called Population Match Probability. PI is a indication of the statistical power of a specific set of marker loci.

PI is the frequency of the *i*th allele at a locus. For multiple loci calculated as the product of individual locus *PI*'s. *PI* represents the average probability of a match for any genotype, rather than for a specific genotype, as in Genotype Probability (Taberlet and Luikart, 1999; Waits *et al.*, 2001). The formula is as follows;

$$PI = 2\left(\sum p_i^2\right)^2 - \sum p_i^4$$

Number of Migrants (Nm)

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

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

Polymorphic Information Content (PIC)

Polymorphic information content (PIC) is a measure of informativeness related to expected heterozygosity and likewise is calculated from allele frequencies (Botstein *et al.* 1980; Hearne *et al.* 1992). It is commonly used in linkage mapping.

Cervus calculates an average PIC across all loci, the arithmetic average of the PIC values at each locus. PIC value is calculated for each primer not for each allele as follows:

Therefore PIC will be calculated by the formula:

PIC= 1-∑(pi2)

Pi will be calculated for each allele. pi is the frequency of the i^{th} allele

pi=no. of alleles/no. of genotype

PIC Values range between 0 and 1. Primers whose pic value is zero or less than zero (i.e negative value) should not be used for analysis.

Percentage of Polymorphic Loci (%P)

P = mean proportion of loci polymorphic in population

$$P = \sum \frac{P_i}{N}$$

P_i = proportion of loci polymorphic in a population and N = number of populations.

Garza-Williamsion Index

Following Garza and Wlliamson (2001), the G-W statistic is given as

$$G - W = \frac{k}{R + 1}$$

where k is the number of alleles at a given loci in a population sample, and R is the allelic range. Originally, the denominator was defined as just R in Garza and Wlliamson (2001), but this could lead to a division by zero if a sample is monomorphic. This adjustment was introduced in Excoffier *et al*. (2005).

This statistic was shown to be sensitive to population bottleneck, because the number of alleles is usually more reduced than the range by a recent reduction in population size, such that the distribution of allele length will show "vacant" positions.Therefore the G-W statistic is supposed to be very small in population having been through a bottleneck and close to one in stationary populations.

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);

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

APPENDIX C

FILE FORMATS OF SOFTWARES

Arlequin File Format-arp file

```
#================================
#Project file created by Arlequin
#================================

#    (symbol # is used to comment what follows on a line)

[Profile]
        Title="Populus euphratica populations in Goksu-Firat rivers"

        NbSamples=2
        DataType=MICROSAT
        GenotypicData=1
        LocusSeparator=WHITESPACE
        GameticPhase=0
        RecessiveData=0
        MissingData="?"

[Data]
        [[Samples]]
          SampleName="Gokcay"
          SampleSize=46
          SampleData= {
95      1   287  219  220  178  213  182  226  197  195  179  348  105
            289  219  220  178  213  182  232  197  201  185  357  107
96      1   287  219  220  178  213  182  232  197  201  179  351  103
            289  219  220  178  237  182  232  197  201  185  351  103
          }
          SampleName="Ermenek"
          SampleSize=23
          SampleData= {
72      1   289  219  220  178  213  182  226  197  195  181  348  101
            289  219  220  178  237  182  232  197  201  185  351  107
73      1   289  219  220  178  213  182  229  197  201  185  351  103
            289  219  220  182  237  182  232  197  201  189  351  105
```

FSTAT File Format-dat file

```
4 5 289 3
WPMS5
WPMS7
WPMS10
WPMS12
WPMS14
1    287289  219219  220220  178178  213213
1    287289  219219  220220  178178  213237
1    289289  219219  220220  178178  213237
1    289289  219219  220220  178178  213237
2    289289  219219  220220  178178  213237
2    289289  219219  220220  178182  213237
2    289289  219219  220220  178182  213237
2    289289  219219  220220  178182  213213
3    287289  219219  220220  178178  213219
3    291291  219219  220220  178178  213237
3    291291  219219  220220  178182  213213
3    287289  219219  220220  178182  213213
4    287289  219219  220220  178178  213237
4    289289  219219  220220  178178  213219
4    289289  219219  220220  178178  213213
4    289289  219219  220220  178182  213237
```

GDA File Format-txt file

```
#nexus

[!Data from Populus euphratica
]

begin gdadata;
     dimensions nloci=3 npops=3;
     format tokens missing=? datapoint=standard;
                locusallelelabels

                1 'WPMS5'   [/ 287 289 291 ],
                2 'WPMS7'   [/ 215 219 ],
                3 'WPMS10'   [/ 220 222 246 248 ];

matrix
GOKCAY:
95      287/289 219/219 220/220
96      287/289 219/219 220/220
97      289/289 219/219 220/220
98      289/289 219/219 220/220,
ERMENEK:
72      289/289 219/219 220/220
73      289/289 219/219 220/220
74      289/289 219/219 220/220
75      289/289 219/219 220/220,
Mut:
1       287/289 219/219 220/220
2       291/291 219/219 220/220
3       291/291 219/219 220/220
4       287/289 219/219 220/220;

end;
```

GenAlEx format- excel file

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	5	16	4	4	4	4	4	4	5	16			
2	goksu-firat												
3	Sample #	Pop	WPMS5	WPMS5	WPMS7	WPMS7	WPMS10	WPMS10	WPMS12	WPMS12	WPMS14	WPMS14	
4	95	Gokcay	287	289	219	219	220	220	178	178	213	213	
5	96	Gokcay	287	289	219	219	220	220	178	178	213	237	
6	97	Gokcay	289	289	219	219	220	220	178	178	213	237	
7	98	Gokcay	289	289	219	219	220	220	178	178	213	237	
8	72	Ermenek	289	289	219	219	220	220	178	178	213	237	
9	73	Ermenek	289	289	219	219	220	220	178	182	213	237	
10	74	Ermenek	289	289	219	219	220	220	178	182	213	237	
11	75	Ermenek	289	289	219	219	220	220	178	182	213	213	
12	1	Mut	287	289	219	219	220	220	178	178	213	219	
13	2	Mut	291	291	219	219	220	220	178	178	213	237	
14	3	Mut	291	291	219	219	220	220	178	182	213	213	
15	4	Mut	287	289	219	219	220	220	178	182	213	213	
16	40	Silifke	287	289	219	219	220	220	178	182	213	237	
17	41	Silifke	289	289	219	219	220	220	178	178	219	237	
18	42	Silifke	287	289	219	219	220	220	178	182	213	219	
19	43	Silifke	287	289	219	219	220	220	178	182	213	231	
20													

GenClone 2.0 File Format-txt file

160	999	999	22	2	WPMS5	WPMS7	WPMS10	WPMS12	WPMS14	WPMS15
1	999	999	287289	219219	220220	178178	213213	182182	226232	214214
2	999	999	287289	219219	220220	178178	213237	182182	232232	214214
3	999	999	289289	219219	220220	178178	213237	182182	226232	214214
4	999	999	289289	219219	220220	178178	213237	182182	226232	214214
5	999	999	287289	219219	220220	178182	213213	182182	226229	214214
6	999	999	287289	219219	220220	178178	213237	182182	229232	214214
7	999	999	287291	219219	220220	000000	213213	182182	229229	214214
8	999	999	287289	219219	220220	178178	213231	182182	229232	214214
9	999	999	287291	215215	220220	000000	213213	182182	229229	214214
10	999	999	289289	219219	220220	178178	213219	182182	226232	214214
11	999	999	287287	215215	220220	178178	213219	182182	226229	214214
12	999	999	289289	219219	222222	178178	219219	182182	232232	214214
13	999	999	287287	219219	220220	178178	213237	182182	229229	214214
14	999	999	287291	219219	220220	178182	213219	182182	232232	214214
15	999	999	287289	219219	220220	178178	213219	182182	226226	214214
16	999	999	289289	219219	220220	178178	213213	182182	229232	214214
17	999	999	291291	219219	220220	178178	213213	182185	229229	214214
18	999	999	291291	219219	220220	000000	213213	182182	229229	214214
19	999	999	287289	219219	220220	178178	213219	182182	226232	214214
20	999	999	291291	215219	220220	178178	237237	182182	229229	214214
21	999	999	287289	215215	220220	178178	213219	182182	226232	214214
22	999	999	287289	219219	220220	178182	213219	182182	226232	214214
23	999	999	287291	219219	220220	178178	237237	182182	226226	214214
24	999	999	289289	215215	220220	000000	213213	182182	232232	214214
25	999	999	289289	215215	220220	182186	213213	182182	232232	214214
26	999	999	291291	215215	220220	178178	219237	182182	226232	214214
27	999	999	289289	219219	220220	000000	219237	182182	229229	214214
28	999	999	287289	219219	220220	178178	213219	182185	232232	214214
29	999	999	287287	215215	220220	178178	213213	182182	232232	214214
30	999	999	287291	215215	220220	178178	213219	182182	226226	214214

Genepop File Format-txt file

```
"Populus euphratica populations in Goksu-Firat rivers"
WPMS5
WPMS7
WPMS10
WPMS12
POP
Gokcay,         287289  219219  220220  178178  213213  182182
Gokcay,         287289  219219  220220  178178  213237  182182
Gokcay,         289289  219219  220220  178178  213237  182182
Gokcay,         289289  219219  220220  178178  213237  182182
POP
Ermenek,        289289  219219  220220  178178  213237  182182
Ermenek,        289289  219219  220220  178182  213237  182182
Ermenek,        289289  219219  220220  178182  213237  182182
Ermenek,        289289  219219  220220  178182  213213  182182
Ermenek,        287289  219219  220220  178178  213237  182182
POP
Mut,            287289  219219  220220  178178  213219  179182
Mut,            291291  219219  220220  178178  213237  182182
Mut,            291291  219219  220220  178182  213213  182182
Mut,            287289  219219  220220  178182  213213  182182
POP
Silifke,        287289  219219  220220  178182  213237  182185
Silifke,        289289  219219  220220  178178  219237  182182
Silifke,        287289  219219  220220  178182  213219  182182
Silifke,        287289  219219  220220  178182  213231  182182
```

Microchecker File Format-txt file

```
Title Line:"Populus euphratica populations in Goksu-Firat rivers"
WPMS5
WPMS7
WPMS10
WPMS12
POP
1_,287289 219219 220220 178178 213213 182182 226232 214214 197197 195201
1_,287289 219219 220220 178178 213237 182182 232232 214214 197197 201201
POP
2_,289289 219219 220220 178178 213237 182182 226232 214214 197197 195201
2_,289289 219219 220220 178182 213237 182182 229232 214214 197197 201201
POP
3_,287289 219219 220220 178178 213219 179182 226232 214214 197197 195201
3_,291291 219219 220220 178178 213237 182182 226229 214214 197197 201201
POP
4_,287289 219219 220220 178182 213237 182185 226226 214214 197197 195201
4_,289289 219219 220220 178178 219237 182182 226226 214214 197197 195201
```

Structure File Format-txt file

w5	w7	w10	w12		w14	w15		w18		w20
95	1	287	219	220	178	213	182	226	214	
95	1	289	219	220	178	213	182	232	214	
96	1	287	219	220	178	213	182	232	214	
96	1	289	219	220	178	237	182	232	214	
97	1	289	219	220	178	213	182	226	214	
97	1	289	219	220	178	237	182	232	214	
98	1	289	219	220	178	213	182	226	214	
98	1	289	219	220	178	237	182	232	214	
99	1	287	219	220	178	213	182	226	214	
99	1	289	219	220	182	213	182	229	214	
100	1	287	219	220	178	213	182	229	214	
100	1	289	219	220	178	237	182	232	214	
101	1	287	219	220	-9	213	182	229	214	
101	1	291	219	220	-9	213	182	229	214	
102	1	287	219	220	178	213	182	229	214	
102	1	289	219	220	178	231	182	232	214	
103	1	287	215	220	-9	213	182	229	214	
103	1	291	215	220	-9	213	182	229	214	
104	1	289	219	220	178	213	182	226	214	
104	1	289	219	220	178	219	182	232	214	
105	1	287	215	220	178	213	182	226	214	
105	1	287	215	220	178	219	182	229	214	
106	1	289	219	222	178	219	182	232	214	
106	1	289	219	222	178	219	182	232	214	
107	1	287	219	220	178	213	182	229	214	
107	1	287	219	220	178	237	182	229	214	
108	1	287	219	220	178	213	182	232	214	
108	1	291	219	220	182	219	182	232	214	
109	1	287	219	220	178	213	182	226	214	
109	1	289	219	220	178	219	182	226	214	
110	1	289	219	220	178	213	182	229	214	
110	1	289	219	220	178	213	182	232	214	

APPENDIX D

Cervus 3.0.7 NULL ALLELE FREQUENCY RESULTS

Table 1: Number of alleles, observed (H_{Obs}) and expected (H_{Exp}) heterozygosity, Hardy Weinberg (HW) deviation, Null allele frequency (F) for each locus (Cervus 3.0.7)

Locus	Number of Alleles	N	H_{Obs}	H_{Exp}	PIC	HW	F(Null)
WPMS5	3	160	0.431	0.600	0.530	***	0.1559
WPMS7	2	160	0.100	0.363	0.297	***	0.5668
WPMS10	4	160	0.038	0.049	0.049	ND	0.1235
WPMS12	5	146	0.205	0.285	0.259	ND	0.1666
WPMS14	5	160	0.650	0.664	0.601	NS	0.0098
WPMS15	4	160	0.175	0.184	0.172	ND	0.0161
WPMS18	3	160	0.569	0.599	0.514	NS	0.0108
WPMS20	1	159	0.000	0.000	0.000	ND	ND
PMGC14	3	160	0.013	0.025	0.025	ND	0.2086
PMGC2163	4	160	0.556	0.435	0.354	*	-0.1256
PMGC2889	6	159	0.742	0.767	0.728	NS	0.0156
PMGC93	5	160	0.613	0.665	0.614	NS	0.0423
Pe2	10	158	0.772	0.761	0.720	NS	-0.0110
Pe5	6	160	0.700	0.583	0.545	***	-0.1339
Pe6	3	160	0.700	0.649	0.572	*	-0.0412
Pe7	1	160	0.000	0.000	0.000	ND	ND
Pe8	4	159	0.409	0.441	0.409	NS	0.0216
Pe13	2	159	0.440	0.487	0.368	NS	0.0490
Pe14	7	160	0.700	0.732	0.685	NS	0.0175
Pe15	9	154	0.753	0.776	0.740	NS	0.0121
Pe17	10	154	0.292	0.793	0.766	***	0.4650
Popeu13	9	154	0.468	0.730	0.701	***	0.2225

APPENDIX E

DAMS AND HYDROELECTRIC POWER PLANTS ON THE GÖKSU RIVER

No:	Power plant name	Province / District	Maximum Operating Level	Minimum Tail Water Elevation	Installed Power	Construction Year
1	**Damlapınar HPP**	Karaman	692 m		16 MW	2008
2	**Kepezkaya HPP**	Karaman	503 m		28 MW	2009
3	**Bucakkışla HPP**	Karaman, Merkez	403 m		41 MW	-
4	**Kayraktepe Dam and HPP** (planned)	Mersin, Silifke	124 m	28 m	282 MW	-
5	**Silifke HPP**	Mersin			0 MW	-
6	**Ermenek Dam and HPP**	Karaman	660 m	694 m	302 MW	2002
7	**Gezende Dam**	Mersin, Mut	310 m	333 m	159 MW	1979

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Kansu, Çiğdem
Nationality: Turkish (TC)
Date and Place of Birth: 29 October 1985, Antalya
Marital Status: Married
Phone: +90 312 210 5160
Fax: +90 312 210 79 76
email: cigdemkansu@gmail.com

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Biology	2011
BS	METU Molecular Biology and Genetics	2008
High School	AKEV High School, Antalya	2003

WORK EXPERIENCE

Year	Place	Enrollment
2010- Present	METU Biological Sciences	Research Assistant
2009-2010	Namık Kemal University Biology	Research Assistant

FOREIGN LANGUAGES

Advanced English

PUBLICATIONS

Dizkirici A., Celep F., **Kansu C.**, Kahraman A., Dogan M., Kaya Z. 2015. A molecular phylogeny of *Salvia euphratica* sensu lato (*Salvia* L., Lamiaceae) and its closely related species with a focus on the section Hymenosphace. Plant Systematics and Evolution, 301 (10): 2313–2323

Dizkirici A., **Kansu C**., Onde S. 2016. Molecular Phylogeny of *Triticum* and *Aegilops* based on ITS and *mat*K sequence data. Pakistan J Botany, 48 (1) :143-153

Dizkırıcı, A., **Cetiner, C.**, Onde, S., Birsin, M., Kaya, Z., and Ozgen, M. 2013 , Phylogenetic relationships of *Triticum* and *Aegilops* genera based on DNA sequences of trnT-F chloroplast DNA non-coding region. Genetic Resources and Crop Evolution, 60 : 2227-2240