GENETIC STRUCTURES OF *SALIX ALBA* AND *SALIX EXCELSA* POPULATIONS FROM TWO MAJOR RIVER SYSTEMS IN TURKEY

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

BY

FUNDA ÖZDEMIR DEĞİRMENCİ

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

NOVEMBER 2017

Approval of the Thesis:

GENETIC STRUCTURES OF *SALIX ALBA* AND *SALIX EXCELSA* POPULATIONS FROM TWO MAJOR RIVER SYSTEMS IN TURKEY

Submitted by **FUNDA ÖZDEMİR DEĞİRMENCİ** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Biology Department, Middle East Technical University** by,

Prof. Dr. Gülbin Dural Ünver _____
Dean, Graduate School of **Natural and Applied Sciences**

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

Prof. Dr. Zeki Kaya _____
Supervisor, **Biological Sci., METU**

Examining Committee Members:

Prof.Dr. Musa Doğan* _____
Biological Sci., METU

Prof. Dr. Zeki Kaya _____
Biological Sci., METU

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

Prof. Dr. Sertaç Önde _____
Biological Sci., METU

Asst. Prof. Dr. Fatih Temel _____
Forest Engineering Dept., Düzce University

Date: 27/11/2017

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 : Funda Özdemir Değirmenci

Signature :

ABSTRACT

GENETIC STRUCTURES OF *SALIX ALBA* AND *SALIX EXCELSA* POPULATIONS FROM TWO MAJOR RIVER SYSTEMS IN TURKEY

Özdemir Değirmenci, Funda
Ph.D., Department of Biology
Supervisor : Prof. Dr. Zeki KAYA

November 2017, 131 Pages

Salix alba (white willow) is an indicator species and an important component of an healthy riparian ecosystem with great renewable energy potential in Turkey. It is used as one of the most effective phytoremediation tool in the world for river cleaning and ecosystem rehabilitation efforts. Genetic structure of *Salix alba* populations in two river systems (Göksu and Kızılırmak Rivers) in Turkey were studied with the use of 20 microsatellites markers to provide information for effective conservation and breeding programs.

Since *Salix excelsa* was used an a synonym of *Salix alba* and was not well differentiated from *S. alba*, 112 genotypes of *S. alba* belonging to the Göksu River from four populations and 147 genotypes belonging to the Kızılırmak River from five populations were sampled. Only some loci showed evidences of null alleles that are specific to a certain population. No significant linkage disequilibrium and clonal duplication were detected in populations.

Effectiveness of markers were evaluated by Allelic richness, Polymorphic Information Content, Probability of identity, Shannon Index and expected

heterozygosity measures which revealed that all loci were found to be highly polymorphic.

Genetic structure analysis clearly revealed that *S. alba* populations in two different river systems represent two different founder populations with very high membership values. All populations maintain moderate level of genetic diversity. The highest pairwise Fst values with the lowest number of migrants were obtained from the Kızılırmak downstream populations. Therefore, these populations were genetically most distant to the remaining populations. Cluster, Bayesian, PCoA analysis, and AMOVA clustered 259 genotypes from nine different locations of two rivers to two major groups because of high level of gene flow within river systems, but less gene flow between river systems.

The genetic structure data of the studied river systems will be useful for future breeding programmes, efficient conservation, management and utilization of genetic resources of *S. alba*. Under increasing habitat deterioration and fragmentation, efficient *in situ* and *ex situ* conservation studies should utilize and conserve the species‟ genetic resources.

Keywords: Genetic diversity, SSR, Population structure, *Salix*

ÖZ

TÜRKİYE' NİN İKİ BÜYÜK NEHİR SİSTEMİNDE BULUNAN *SALIX ALBA* VE *SALIX EXCELSA* TÜRLERİNİN GENETİK YAPILARI

Özdemir Değirmenci, Funda
Doktora, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. Zeki KAYA

Kasım 2017, 131 sayfa

Salix alba, yenilenebilir enerji potansiyeli ile, sağlıklı bir nehir ekosisteminin önemli bir bileşeni olarak, Türkiye'de bulunan gösterge türlerdendir. Dünyadaki nehir temizliği ve ekosistem rehabilitasyon uygulamalarında en etkili fitoremediasyon aracı olarak kullanılır. Türkiye'deki iki nehir sisteminde (Göksu ve Kızılırmak Nehirleri) bulunan *Salix alba* popülasyonlarında etkin koruma ve ıslah programlarına bilgi sağlamak için 20 mikrosatellit belirteci ile populasyonların genetik yapısı çalışıldı.

Salix excelsa, *Salix alba*'nın sinonimi olarak kullanıldığından ve *S. alba*'dan çok iyi ayrılmadığından, dört popülasyona ait 112 genotip *S. alba* Göksu Nehri'nden, beş popülasyona ait 147 genotip Kızılırmak Nehri'nden örneklenmiştir. Yalnızca bazı lokuslar, belirli bir populasyona ait null alellerin olduğunu göstermiştir. Alt popülasyonlarda önemli bir bağlantı dengesizliği ve klonal çoğalma gözlemlenmemiştir.

Belirteçlerin etkinliği, Allelik zenginlik, Polimorfizm Bilgi Çeşitliliği, Tanımlama Olasılığı, Shannon Çeşitlilik İndeksi ve beklenen heterozigotluk değerleri tüm lokusların oldukça polimorfik olduğunu gösterdi.

Genetik yapı analizi, iki farklı nehir havzasındaki *S. alba* popülasyonlarının iki farklı kurucu popülasyonuna yüksek ait olma değeri ile sahip olduklarını açık bir şekilde ortaya koymuştur. Tüm popülasyonlar orta seviyede genetik çeşitliliğe sahiplerdir. En yüksek Fst değerleri en düşük sayıdaki göç eden birey sayısı ile birlikte Kızılırmak alt popülasyonlarında görülmüştür. Dolayısıyla, bu popülasyonlar genetik olarak diğer popülasyonlara daha uzaktır. Gruplama, Bayes, PCoA analiz ve AMOVA, iki nehrin dokuz farklı lokasyonundan 259 genotipi nehir sistemi içinde yüksek düzeyde, nehirler arasında daha düşükdüzeyde gen akışı nedeniyle iki gruba ayırmıştır.

Elde edilen veriler gelecek ıslah programlarında, *S. alba* türünün genetik kaynaklarının etkili korunma, yönetim ve kullanımında faydalı olacaktır. Artan habitat bozulması ve parçalanması dikkate alındığında, etkin *in situ* ve *ex situ* koruma çalışmaları, türün genetik potansiyelini korumak için gerçekleştirilmelidir.

Anahtar Kelimeler: Genetik çeşitlilik, SSR, Popülasyon yapısı, Salix

To my son…

ACKNOWLEDGEMENTS

I would like to greatly thank to my supervisor Prof. Dr. Zeki Kaya for his guidance, supervision and endless patience throughout the study. I would like to also thank to Prof. Dr. Sertaç ÖNDE for his concern and kindness.

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

I specifically thank to TÜBİTAK (The Scientific and Technological Research Council of Turkey) because of the MS and PhD scholarship that I was honored.

This thesis was supported by the research fund: TÜBİTAK/TOVAG 213O154 Project "Molecular Phylogeny of Turkish *Salix* L. species and Genetic Characterization of two Economically Valuable Willow Species (*Salix alba* and *Salix excelsa*) for tree breeding purposes", and BAP-08-11-DPT2002K12510.

I am also deeply thankful to my lovely friends Asiye Çifti and Çiğdem Kansu for their endless help and productive critics in the writing part of the thesis. I am also really thankfull to my project and field trip colleague, Pelin Acar; my labmates; Bircan TaĞkıran, Abdulbaki Çoban, Burak Yelmen, Sıla Aygün, Yasin Tokdemir and Nurbahar Usta from Plant Genetics and Tissue Culture Laboratory.

I would like to express my special gratitude to my mother Mukadder Adaçay, my grandmother Zühre Özdemir, my sister Sevgi GüneĞand my lovely cousin Kübra Yılmaz for their endless support, encouragement and love all through my life. I wish also to express my appreciation to Prof. Dr. Orhan ARSLAN and my friends for their endless love and support.

Finally, but never least, my kindest thanks must be to my dear husband, Nuri Değirmenci, who has been a constant source of support and encouragement during the challenges of my studies and life. I am truly thankful for having you in my life.

I dedicate this thesis to my sweet son Orhan Arslan Değirmenci.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

%P	Percentage of Polymorphic Loci
AMOVA	Analysis of Molecular Variance
Ar	Allelic Richness
CLUMPP	CLUster Matching and Permutation Program
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribonucleic Acid
dNTP	Deoxy ribonucleotide triphosphate
EDTA	Ethylenediaminetetraaceticacid disodium salt
EU	European Union
EUFORGEN	European Forest Genetic Resources Programme
F	Inbreeding Coefficients
Fct	Difference Among Groups For The Total Population
Fis	Inbreeding Coefficient Within Individuals
Fit	Inbreeding Coefficient Within Total Population
Fsc	Differences Among Population Within Groups
Fst	Differences Among Subpopulation (AMOVA)
Fst	Inbreeding Coefficient Within Subpopulations
GDA	Genetic Data Analysis
GPS	Global Positioning System
G-W Index (M)	Modified Garza-Williamson Index
He	Expected Heterozygosity
Ho	Observed Heterozygosity
HWE	Hardy–Weinberg Equilibrium
I	Shannon‟s Information Index
IAA	Isoamyl Alcohol
LD	Linkage Disequilibrium
MCMC	Markov Chain Monte Carlo
MgCl2	Magnesium Chloride
MLGs	Multilocus Genotypes
Na	Number of Alleles
Ne	Number of Effective Alleles

Nm	Number of Migrants
OD	Optical Density
PCoA	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
PI	Probability of Identity
PIC	Polymorphic Information Content
rd	Standardized Index of Association
SSR	Simple Sequence Repeats
TBE	Tris-Borate-EDTA
TE	Tris EDTA
TOVAG	Agricultural Forestry and Veterinary Research Group
TÜBİTAK	The Scientific and Technological Research Council of Turkey
UPGMA	Unweighted Pair-Group Method with Arithmetic Mean

CHAPTER 1

INTRODUCTION

There is an expanding interest in developing approaches aimed to protect the genetic diversity of trees throughout the world. Genetic diversity is an important element in the dynamics of populations, because it is directly relevant to the evolutionary potential of the populations (Hughes *et al.,* 2008). The field of population genetics concerns genetic differences within and among populations. Measurement of diversity parameters enable ecological and evolutionary interpretation of the species under the forces of evolution which is caused by environmental changes, random changes from one generation to the next, migration or genetic mutations (Toro *et al.,* 2009; Caballero and Garcia-Dorado, 2013).

Conservation of populations has increasingly become a vital concern since human activity leads to the disappearance of many species. Moreover, an important aspect of conservation is the protection of genetic diversity (Newman and Pilson, 1997). Conservation schemes combine general rules derived from the F-Statistics framework with the particular ecological scenario in question (Allendorf *et al.,* 2012). Ecological and evolutionary consequences of genetic diversity measures should be taken into account in conservation and management studies for preserving genetic diversity (Greenbaum *et al.,* 2014).

In Turkey, the rapid human population growth and natural resource requirements threaten the biological diversity of natural ecosystems, including forests. Forest areas are limited with approximately 26% of the total land of the country. The preservation of healthy forests is vital to support sustainable development in Turkey (Kaya and

Raynal, 2001). Only *in situ* or *ex situ* conservation provide effective management of tree populations. In the face of climate change, these approaches are very important for long-term sustainability of forests.

Salix species are important and abundant member of riparian ecosystems in Turkey. In terms of river plantations, water quality, erosion control and biodiversity center, they are accepted as pioneer species. The contribution and scope of global *Salix* plantation, usage, sustainable forestry and rural development are increasing. Over 176,000 hectares (ha) of land worldwide have been cultivated with willow, of which 90,000 hectares are used for timber production. The rest is used for ecological goals such as ecosystem restoration, phytoremediation, and biological engineering (Ball *et al.,* 2005).

The European Union in 2007 targeted a 20% of total energy coming from renewable resources by 2020. Accordingly, the EU now places a high priority on biomass energy production. Genus *Salix* include various species in terms of biomass productivity, and capacity of heavy metal ions absorption. Compared with other plants, *Salix* is highly effective in terms of biomass productivity (Fischer *et al.,* 2005; Vande *et al.,* 2007)

Poplar and Fast Growing Forest Trees Research Institute (Gzmit, Turkey) of the General Directorate of Forestry, the Ministry of Forestry and Water Affairs works on projects related to poplar, willow, and other fast growing species of forest trees. The Institute also provides technical support to both private and public sectors to increase the production of wood derived from non-forest trees‟ resources. Poplars and willows species are excluded from Turkish forestry regulations. Besides, poplar is present at the "forest atlas" book created by General Directorate of Forestry of the Ministry of Forestry and Water Affairs, while willow was not referred as forest tree at all (Forest Atlas, 2013). In Europe, *Salix alba* L. has two *ex situ* genetic conservation areas that are found in EUFORGEN (European Forest Genetic Resources Programme) and one of them is core network. Both forest areas are located in Romania.

S. alba L. (white willow) is one of the largest and best known willows in the world. White willow is the most important species of Turkish river ecosystems and has economic and ecological importance. To date no studies have been conducted on the identification and conservation of the genetic resources of *S. alba*. Molecular data presented in this dissertation, will contribute to the conservation of *S. alba* genetic resources and will help breeding programmes to be effectively carried out.

1.1. The Salicaceae Family

Salicaceae family (the Willow family) is a family of flowering plants containing two genera; *Populus* and *Salix*, commonly known as poplars and willows (Cronquist, 1981) which comprise approximately 450 species (Argus, 2007). The Angiosperm Phylogeny Group (2003) suggested that some of the genera in Flacourtiaceae should be included in Salicaceae. Resettlement of Salicaceae places *Populus* and *Salix* in tribe of *Saliceae*. Few genera (*Flacourtieae, Samydeae, Homolieae, Scolopieae, Prockieae, Abatieae and Bembiciea*) in the family Flacourtiaceae are other tribes in the family Salicaceae sensu lato (Chase *et al*., 2002). Therefore, most of Flacourtiaceae members are placed in Salicaceae and contain 43 genera and about 1,000 species (Cronk *et al.,* 2015). They are noticeable in Northern Hemisphere, especially in temperate climates (Wang *et al*., 2014). Poplars and willows are used for producing potential timber and biofuel, reforestation programmes and removal of heavy metals from the soil (Pohjonen, 2008; Han *et al*. 2010; Hinchee *et al*. 2010).

The lineages of *Populus* and *Salix* diverged about 60-65 million years ago based on the fossil record. However, recent DNA sequencing of *Populus* genome has produced evidence that these lineages share the same large-scale genomic history. These two genera have many diverse species and small genome size. They grow rapidly and undergo easy vegetative propagation (Tuskan *et al.,* 2006). They are pioneer plants in many ecosystems, especially on river alluvial. They tolerate high environmental stress. Moreover, two genera are suitable for basic research in molecular biology and genetics, plant domestication and conservation as model organisms (Kuzovkina and Vietto, 2014).

All species of Salicaceae family are characterized with high demands for water and light. They are also well known because of the colonization of newly emerging substrates. Salicaceae can survive in high humidity habitats with dry climatic conditions, at the same time, they settle all type of favorable habitats in humid climates. The seeds are minute and can be scattered in large quantities by the wind. In addition, seeds of most species can germinate immediately when exposed to moist surfaces. Embryo of seeds have a soft pervious skin (Skvortsov, 1999). The coexisting *Populus* and *Salix* sex ratios change by height and type; 1: 1 equilibrium sex ratios are seen at middle altitude, but found in skewed at high altitudes, in which *Populus* exhibit male-biased and *Salix* female-biased sex ratios (Lei *et al.,* 2017). Several factors (ecological, demographic and genetic factors) may affect the sex ratio such as difference in reproductive cost between the sexes, early onset of flowering, more frequent flowering in one of the sexes, gender-specific differences in mortality and local competition, X-linked meiotic drive and selfish gene elements (Delph, 1999; Taylor and Ingvarsson, 2003; Barrett *et al.,* 2010).

The common points of poplars and willows in ecological features are numerous. They are used in similar environmental projects, however, the willow is superior in many ways. Willow can be better adapted when compared to poplars. The number of willow varieties is about 10 times higher than the number of poplars. As a result, their geographical distribution and physiognomic range show greater diversity for exploitation of environmental resource and hence wider environmental applications (Verwijst, 2001).

1.2. Genus *Salix*

According to several authorities, genus *Salix* (willows, sallows and osiers) includes different number of species such as 350 (Skvortsov, 1968), 526 (Fang, 1987), 450 (Argus, 1997) or 500 (Hardig, 2010). *Salix* L. is considered to be represented by 450–520 species (Wu *et al*., 2015). *Salix* species concentrated mainly in temperate and cold regions of the Northern Hemisphere (Argus, 1997; Figure 1.1). They show great morphological variation in the world (Hardig *et al*, 2010). The centers of abundance of the genus are China (275 species) and the former Soviet Union (125

species) (Liu, 1999). There are 100 species in North America and 65 species in Europe (Argus, 1997).

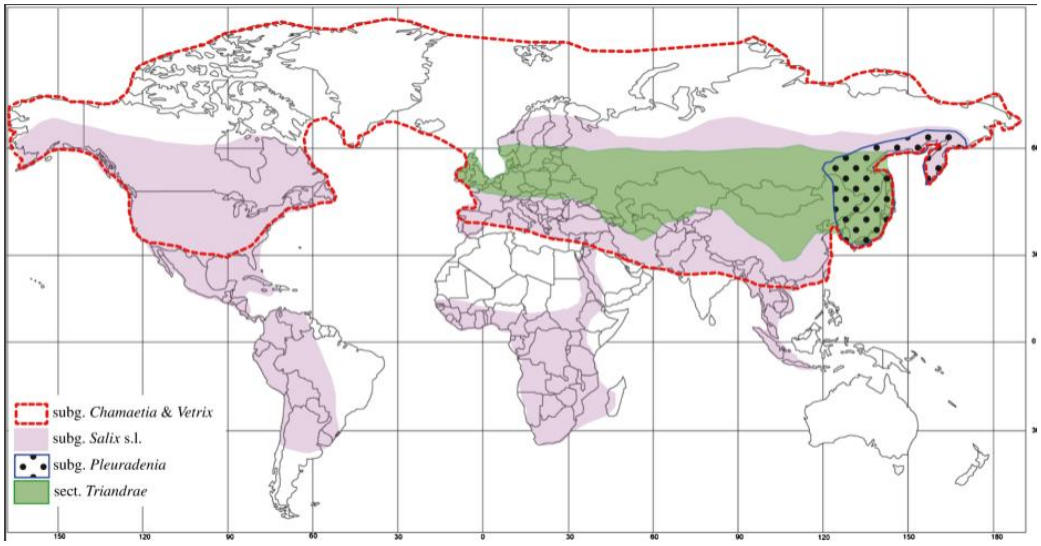


Figure 1.1. Geographic distributions of genus *Salix* (Wu *et al.*, 2015)

The willows are fast-growing woody plants of various habitats and their sizes range from huge upright trees to tiny dwarf shrubs, just a few centimeters long. The majority of species are shrubs.

Alluvial (riparian) and non-alluvial (wetland) rivers are the two major ecosystems for willows. Trees growing in alluvial rivers demand quite high aeration and fast water flow. Therefore, they inhabit on deposits in river bed that accumulate through river flow. These are mostly trees or stron and tall shrubs (Skvortsov, 1999). Considering the major adaptations and limitations of willow as pioneer species, they are able to colonize oligotrophic areas confined to nutrients such as marshes, sand hills, river sand and pebbles, to form symbiotic association with mycorrhizal fungi that provide additional nutrients (Heijden and Kuyper, 2003). Willows are known as early successional plants that colonize newly opened habitats, such as high land that include man-made habitats (Argus, 1986; Skvortsov, 1999).

Willow species also vary according to demands of moisture, humidity and temperature conditions. At different stages of development and growth, *Salix* species show different moisture requirements. Moisture availability is crucial during seed

germination because *Salix* seeds can only live for a few weeks (Maroder *et al.*, 2000). It may be a limiting factor in areas with dry conditions during the last months of spring. However, after seedling establishment, no fixed soil moisture is required for many willow species to survive (Skvortsov, 1999). Willows produce thousands of seeds per year, thus, this increases the probability of seed germination and seedling growth. They settle into vegetation space in which plants can stand for full sunlight and have relatively high growth rate with relatively short life expectancy (Raven, 1992). The size of surface acidity and mineralization are very important for some species (Skvortsov, 1999).

Most *Salix* species contain two sets of 19 chromosomes (2n =2x= 38) that are small and uniform in size. However, two sets of 22 chromosomes are also observed. The genus is very heterogeneous in terms of having different species. There are many inter-specific hybrids at every ploidy level. Approximately 40% of the willow species are identified as polyploid, and the polyploidy levels in *Salix* genus vary from diploid (2*n*) to dodecaploid (12*n*) (Suda and Argus, 1968; Argus, 1997; Skvortsov, 1999; Dobes and Vitek 2000).

Willows contribute to human societies from a social and economic point of view. *Salix* species characterized by specific physiological adaptations and ecological stability are used for application in conservation and environmental projects in many climatic and adverse micro-scale conditions. These involve ecosystem restoration (ecological restoration of wetlands and wildlife conservation, land reclamation, afforestation of industrial areas), phytoremediation (phytoextraction, phytodegradation, rizofiltration and phytostabilization), bioengineering (water and wind erosion and windbreaks, living walls and shelters), and biomass production for both fuel and fiber (Kuzovkina and Quigley, 2005).

Poplar, willow and *Robinia pseudoacacia*, which are short rotation forestry crops, were determined and compared to assess their propriety for energy devices. These evaluations were made in relation to energy potential and combustion related operating problems based on the physical and thermochemical feature of the trees. The obtained results showed that the most appropriate tree is willow (Monedero *et al.*, 2017).

Willows with an improved root system show strong resistance against heavy metals. Thus they are widely used for phytoremediation of land that is contaminated with heavy metals. Some *Salix* species represent the salinity tolerance that is a valuable adaptive trait for reclamation of mine sites (Major *et al.,* 2017).

Salix is a promising resource for reducing the effects of environmental degradation. The utility of the *Salix* genus for improvement in environmental projects is highlighted by degradation 10 of 15 kinds of soil chemicals classified by Logan (1992) by using the species. *Salix* proposes elimination of chemicals from erosion region, mine spill, industrial waste, dredge spoil, ore grills, sewage sludge, oil spills, destruction of oil shale waste, nuclear waste and landfills (Kuzovkina and Quigley, 2005).

Colonization of *Salix* species in degraded areas may accelerate rehabilitation and reestablishment of ecosystems. *Salix* species generally increase surface shade, annual foliar production, root movement and humus formation, thereby they reconstruct the soil structure and source of nourishment (Stott, 1992). With their large biomass, they play important role in food chains of different ecosystems (Kacalkova *et al.,* 2015; Yang *et al.,* 2015).

There is a wide ranging natural willow distribution in Anatolia. The number of willow species in Turkey is 23 according to Davis (1965), 28 according to Arıhan and Güvenç (2009), 33 according to Kantarcı (2012), 27 according to Terzioğlu (2014) and 24 according to Velioğlu and Akgül (2016). They are distributed in small groups and individually in the valleys of the rivers. The total area of willow stands is about 2090 ha (877,73 ha. pure, 1209,28 ha. mixed with other forest trees) (Velioğlu and Akgül, 2016). *S. trabzonica, S. purpurea* subsp. *leucodermis* (in Aegean region), *S. rizeensis* (in Soğanlı mountains in Black Sea region) and *S. anatolica* (in Adana) are endemic to Turkey (Skvortsov and Edmonson, 1970; Donner, 1990; Güner and Zielinski, 1993; Güner, 2000; Zielinski and Tomaszewski, 2007).

1.3. Taxonomy of *Salix alba* and *Salix excelsa*

Willows belong to kingdom Plantae and are under the Tracheophyta phylum within the class of Magnoliopsida. The order of the *Salix alba* and *Salix excelsa* is Salicales and they are considered members of the Salicaceae family. The genus, subgenus and section of the *S. alba* and *S. excelsa* is *Salix* (Table 1.1). The subgenus *Salix* mostly consists of tree species which share many common features with Populus (Kuzovkina *et al.,* 2008). When these two species are compared in terms of their morphological characters, only bud scale shapes and catkin lengths are different. If bud scales are lanceolate-oblong, it could be categorized as *S. alba*, whereas if bud scales are ovoid-lanceolate, it could be classified as *S. excelsa.* Catkins length of *S. excelsa* (40-50 x 10 mm) is longer than *S. alba* (25-30 x 3-4 mm). It is obvious that these two species are morphologically similar. These taxa are difficult to discriminate using morphological characters in natural populations.

S. alba, white willow, is taxonomic synonym f *S. excelsa* Gmelin (Isebrands and Richardson, 2014). Accepted scientific name *S. excelsa* is synonyms of *S. alba* var. *australior* (Anderss.). A cross-checking in various databases (Catalogue of Life, International Plant Names Index, Global Biodiversity Information Facility, The IUCN Red List of Threatened Species) confirms the same result. *S. alba* and *S. excelsa,* were evaluated based on chloroplast (*trn*T-F, *mat*K and *rbc*L) and nuclear genome (*ITS*) regions to investigate the evolutionary relationships. It was seen that there was only one nucleotide substitutions at 1654^{th} bp position of *mat*K region (Acar, 2017). Mutation-based changes have occurred, but speciation is still not clear because there is no complete geographic isolation and natural selection. This is due to chloroplast hybridization and is referred to as incomplete lineage sorting in *S. alba* and *S. excelsa* (Maddison and Knowles, 2006). It is anticipated that these two species may be merged and treated as a single species. Molecular studies also support this idea.

Table 1.1. Taxonomy S*. alba* L. and *S. excelsa* S.G. Gmel.

Kingdom	*Plantae*	
Phylum	*Tracheophyta*	
Class	*Magnoliopsida*	
Order	*Salicales*	
Family	*Salicaceae*	
Genus	*Salix*	
Subgenus	*Salix*	
Section	*Salix*	
Scientific Name	*Salix alba*	*Salix excelsa*
Species Authority	L.	S.G. Gmel.
Common Name(s)	White Willow	-
Red List Category&Criteria	Least Concern	Least Concern
Year Published	2014	2017

http://www.iucnredlist.org/ (Khele, 2013; Kavak, 2017)

1.4. Biology of *Salix alba* L.

Salix alba L. is (white willow) a simple deciduous broadleaved and dioecious (unisexual) species with male (staminate) and female (pistillate) flowers (inflorescences) taking place on separate plants and thus mates by outcrossing. White willow is one of the largest, fastest-growing, and best-known willow speices and can reach up to 30 m in height and 1 m or more in diameter. Its bark is reddish green or brown, smooth in young trees, and there is longitudinal fissure with age (Luna, 1995). They are short lived tree species (20-30 years old) and their seeds remain viable only a few days (Praciak *et al.,* 2013). The leaves are lanceolate-oblong and there are silver-gray on the upper side, and intense silky white feathers on the lower side, which gives the tree a unique pale appearance (Mitchell *et al.,* 1974; Johnson and More, 2006).

Species flower bloom early spring to early summer. The pollens are sticky and flowers are mainly pollinated by insects, especially by bees. Pollination also happens by wind. Imperfect flowers are mostly found on catkins (aments) which occur before or together with the leaves. The male catkins are yellow. They can grow up to 5 cm long, while, the female catkins are shorter and greenish-yellow that turn into a fluffy white (Figure 1.2). The seeds are so small, yellow, and ovoid with silky hair. They are dispersed by wind (Mitchell *et al.,* 1974; Skvortsov, 1999). The fruit is located in ovoid-conic capsule that contain many brownish or black seeds with feathery (Arıhan and Güvenç, 2009).

Figure 1.2. *Salix alba* L. male (on the left) and female (on the right) trees on the banks of Kızılırmak River (Photographed by Funda Ö. Değirmenci)

The flower of *S. alba* carries many stamens. *S. alba* is a European-West Siberian species extending to the Mediterranean region whereas *S. excelsa* is an Iranian species. Both are cultivated in eastern Asia, Transcaucasia, and Middle Asia, where is one of the oldest regions of plant domestication. The natural boundaries of plant communities in these regions were destroyed and devastated years ago by human activities. As a consequence, it should be handled by numerous intermediate forms, likewise, it is seen appropriate that the idea of merging both species and being treated as a single species (Skvortsov, 1999).

Willows have cross fertilization ability with other *Salix* species. *S. fragilis* and *S. alba* are crossing everywhere in Central Europe, Western and temperate Europe, and Russia. Barcaccia *et al.,* (2014) suggested that *S. alba* and *S. fragilis* are allopolyploid and share a common diploid ancestor. However, hybridization between them is not common in Asia. In general, *S. alba* show a very uniform characters. A major departure from uniformity may be caused by hybridization of *S. alba* and *S. excelsa* in Asia, Syria, and Caucasus (Skvortsov, 1999).

Chromosomes were counted as n = 2x = 38 and 2n = 76 in *S. alba* based on cytological investigations (Druskovic, 1995; Khalilia *et al.,* 2012). White willows indicate an allotetraploid genome with disomic inheritance (Barcaccia *et al.,* 2003, Barcaccia *et al.,* 2014). Presence of tetraploidy in *S. alba* makes genetic analysis difficult (Håkansson, 1955; Barcaccia *et al.,* 2003).

1.5. Distribution of *Salix alba* L.

S. alba is a native tree species in Asia and Europe and also naturalised in North America (Uotila, 2011). Since it is extensive dispersion from Western Siberia to North Africa as well as its wide cultivation since ancient times, it is difficult to determine the natural range of *S. alba* (Isebrands and Richardson, 2014) (Figure 1.3).

Figure 1.3. Natural range of *Salix alba* L. Orange triangles show commonly believed introduced and/or naturalized populations (Houston *et al.,* 2016).

White willow is the most important willow species in Turkey. Various forms ranging from large trees to shrub are well known (Velioğlu and Akgül, 2016). *Salix alba* with poplar and other willow species are widely distributed in almost all river basins in geographical regions of Turkey, which have very different climate characteristics (Davis, 1965; Avcı, 1999) (Figure 1.4).

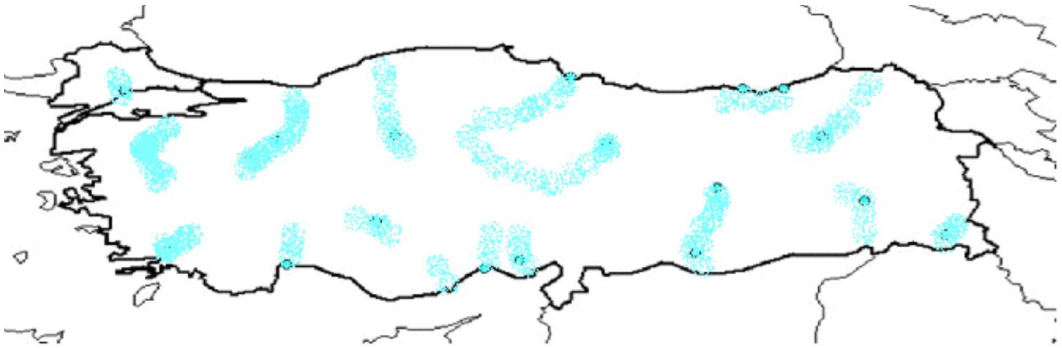


Figure 1.4. Natural distribution of *S. alba* in Turkey (Edited from Turkish Plant Database; http://www.tubives.com/index.php?sayfa=karsilastir)

1.6. Habitat and Ecology of *Salix alba* L.

Salix alba is a kind of riparian species that grows in temperate climates with moderate winters and hot summers (relatively short drought periods) (Isebrands and Richardson, 2014).

S. alba typically grows on banks of rivers and lakes (Figure 1.5). It also grows in ponds, streams, wet hollows and ditches, riverine and riparian woodland (Preston *et al*., 2002). It can be found at 2400 m above sea level and reach at the highest elevations in southernmost parts of its range (Isebrands and Richardson, 2014).

White willow is alluvial species with preference fine sandy or sandy-muddy deposits. This species comparing with non-alluvial ones, it has generally wider latitudinal and altitudinal ranges. It is fully dependent on light without tolerating shadow (Skvortsov, 1999).

S. alba forms small colonies by help of root sucker or adventitious roots and reproduces vegetatively by broken branches which are fragile on the ground (Kuzovkina *et al.*, 2008).

Figure 1.5. View of *Salix alba* population along the Kızılırmak river bank (Photographed by Funda Ö. Değirmenci)

1.7. Significance, and Usage of *Salix alba* L. and Threats to Its Genetic Resources

Willows have a long history of cultivation for many historical and traditional uses such as cricket bats, treatment of disorders (many skin diseases) and aspirin production in earlier times (Karp, 2013; Hanley and Karp, 2014; Madan and Levitt, 2014). Salicylic acid (precursor of aspirin) is naturally found in the bark of the *S. alba* (Miner and Hoffhines, 2007). Due to having a high salicylate content, *S. alba* has economic potential for the production of herbal medicines by the pharmaceutical industry (Sochor *et al.*, 2013). Zabihi *et al.* (2017) reported that *S. alba* extract improved redox homeostasis in heart and kidney tissues of hypercholesterolemic rabbits owing to antioxidant property of its phenolic content. Salicylic acid is also a plant growth regulator and useful in the growth and development of full plant (Raskin, 1992). However, tannin is also obtained from the bark, there is little

commercial significance of these two non-wood products (Praciak *et al.,* 2013). The dried inner bark is pulverized and can be used in bread-making. The boiling of the tender branches is served as a beverage used as an impurity for tea (Sastry and Kanathekar, 1990).

White willow usually grown in southern areas of almost all continents for the production of wicker and wood (Kuzovkina *et al.,* 2008). Wood is used in a number of commercial areas e.g. wooden kitchenware, archery springs, circles, wicker baskets, canola and some construction use (Isebrands and Richardson, 2014). *S. alba* is used as a fence pole, it is as durable as oak (Luna, 1995). The cricket bats and artificial limbs industry is only dependent on *S. alba* wood. There is an available market for the willow timber-based production (Saini and Sharma, 2001).

White willow is ideal for pulp industry among fast growing trees. More recently, it has been investigated as a source for biomass production in different power stations. *S. alba* seems to be promising biomass source for energy purposes and phytoremediation (Mleczek *et al.,* 2010). Planting of tree species with the integration of agricultural plant, willow is the most environmentally friendly, and farmers‟ preference (Paray *et al.,* 2017). Variability of wood density and fibre length study on 36 genotypes of *Salix alba L*. from seven different European countries such as Italy, Hungary, U.K, Netherlands, Turkey, Yugoslavia and Croatia, indicated that Turkish white willow source gave the best performance in terms of both character (wood density and fiber length) (Gupta *et al.,* 2014). Based on studied morphological and physiological indicators, *S. alba* shows more resistance to pollutants. It is recommended for reconstruction of the sanitary and protective plantations. *S. alba* has been recommended for heavy metal-accumulating plants by many authors (Vashegyi *et al.,* 2005; Giachetti and Sebastiani, 2007; Vassilev *et al.,* 2007; Migeon *et al.,* 2008; Pajević *et al.,* 2009; Vamerali *et al.,* 2009; Pourrut *et al.,* 2011). It has capacity to detoksify harmful technological substances and a more effective ability to accumulate them in leaves and stems (Rostunov *et al*. 2017). It was found that phytoextraction of cadmium from contaminated environment was possible using *S. alba* (Arsenov *et al.,* 2017). Esbri *et al.* (2017) recently reported that *S. alba* can be used as biomonitors of gaseous mercury pollution assessment. It was proven to be a useful, reliable and cost-effective methodology.

Everywhere within and beyond the natural range of the species, *S. alba* is preferred to be cultivated in residential areas, on roadsides, edges of reservoirs, parks, and sometimes in forest areas. It is cultivated extremely as ornamental plants with various varieties with attractive stem colors for decorative purposes (Skvortsov, 1999).

Distribution range of white willow overlaps with many areas in Europe where erosion rates are high, especially on moist slopes with high drainage areas within European mountain systems (Bosco *et al.,* 2015). In these crucial places, white willow contributes to erosion mitigation. It is generally useful for controlling erosion preventing climate change and stabilizing water channels, as well as for ecosystem restoration such as degraded lands and forest landscape, and phytoremediation (Kuzovkina and Quigley, 2005; Ball *et al.,* 2005; Stettler, 2011; Savill, 2013).

In Turkey, willows are traditionally planted in rows around margins of moist meadows and fields, and regularly are pruned to produce small sized fuel-wood. Rural communities use willow timber for various agricultural and domestic purposes. The hedged material provide raw material for non-wood products for example livestock, medicinal extracts (Velioğlu and Akgül, 2016).

Many animals use white willow as living space. There are many insects, amphibian, and birds that depend on the species; especially, predaceous diving beetle, American toad, tree swallows (Bosi, 2001; Koutsos *et al.,* 2010; Dubiec, 2013).

There are different factors that threat the *Salix alba* as well as other tree species. White willow is categorized as “Least Concern” in Europe since it is widespread with stable populations. However, river management activities change the river flow and sediment ratios and reduce floodplain areas, which cause a decrease in the regeneration ability of *S. alba* (Barsoum 2001). Furthermore, as a result of agricultural and urban drainage, wetlands are significantly disappeared. Hence, the overall habitat and distribution of native willow have decreased. It is currently highlighted that conservation of natural willow populations and related ecosystems are very important (Vries, 2001).

S. alba trees are under attack by various pests (Yasaman *et al.,* 2017). White willow is specifically vulnerable to filigran disease, which is induced by *Erwinia salisis* bacteria (Turner, 1992). Diseased trees are exposed to wilting, fading and leaf browning. The wood of affected trees becomes fragile and useless. *S. alba* is also attacked by Asian beetles (*Anoplophora glabripennis*) and flying insect (*Phytobia barnesi)* that disrupt the quality of the wood (Faccoli *et al.,* 2015; Rigo, 2016). Moth defoliators (*Lymantria dispar, Porthetria obfuscate* and *Yponomeuta rorella)* damage trees by destroying leaves (Tillesse *et al.,* 2007; Camerini, 2009; Rigo, 2016).

1.8. Population genetics

Population genetics evaluates the genetic diversity in populations and includes the study and modeling of changes in gene and allele frequencies in gene pool of populations (Klug *et al.,* 2006; Hartl and Clark, 2007). Therefore, population genetic approaches such as Mendel's laws of inheritance and mathematical logic, Hardy-Weinberg law are used to clarify ecological, evolutionary, conservation questions about the contemporary dynamics of natural populations (Manel, 2005; Klug *et al.,* 2006).

The Hardy-Weinberg principle states that individuals mate entirely at random, furthermore the allele frequencies and genotype frequencies do not change from one generation to next under the certain assumptions. The most important of these assumptions are the following: (i) Mating is random (panmictic population); (ii) The population is infinitely large; (iii) Genes are not added from outside to population (there is no gene flow or migration); (iv) Genes do not mutate from one allelic state to another (There is no mutation); (v) All individuals have equal probabilities of survival and of reproduction (There is no natural selection affecting the locus of interest) (Futuyma, 2013).

Nonrandom mating, chance, gene flow, mutation and selection can change alleles and genotypes frequencies. These are the major factors that cause evolutionary change within population.

Inbreeding is a form of nonrandom mating, that occur when individual mates with relatives rather than with nonrelatives. It increases the proportion of homozygotes. For reason, if species harbor recessive alleles that have deleterious effects, these cause a decline in components of fitness, such as survival and fecundity. Such decline is called inbreeding depression. The probability that a random pair of gene copies are identical with descent is denoted by *F,* the inbreeding coefficient and can be estimated by the deficiency of heterozygotes relative to the H-W equilibrium value (Wright, 1951). Fis estimates reduction of heterozygosity of an individual due to non-random mating within subpopulations. Fit (total inbreeding in all subpopulations) is a measure of the reduction of heterozygosity of an individual in relation to the total population. The reduction in heterozygosity in a subpopulation due to the genetic drift is estimated by Fst, that is a measure of the relative differentiation in allele frequencies between subpopulations. In the case of random mating, values close to zero are predicted, while significant positive values indicate inbreeding or null alleles (heterozygote deficiency). Negative Fis and Fit values indicate an excess of heterozygotes, due to negative assortative mating (tree mating with tree not similar to themselves) or selection result in favor of heterozygotes (Hartl and Clark 1997).

Natural populations are finite in size, therefore genetic drift occur in all natural populations. It is one of the most important random processes in evolution. It occurs when there is a random fluctuation in alleles frequencies. Eventually, one or more alleles become fixed. The rate decline in heterozygosity is often used as a measure of the rate of the genetic drift within populations. Genetic drift also occurs as a result of bottleneck or founder effect. Restriction in size of populations are called bottleneck that occurs when a new population is established by a small number of founder. Because the founders are a small sample from the source populations, allele frequencies in the new populations may differ by chance from the source populations. The resulting random genetic drift is often called a founder effect (Futuyma, 2013).

Gene flow is gene exchange of mating individuals from different populations, also called migration. Immigrants from other populations may carry different allele combination. If they interbreed with residents, this will change allele and genotype

frequencies (Futuyma, 2013). It should be noted that in most cases gene flow has a greater effect on population structure compared to mutations. A single gamete change between two populations prevents the neutral alleles from being fixed in the recipient population. From an evolutionary point of view, estimating the magnitude of gene flow between populations is of great importance (Gösta *et al.,* 2013).

Mutation is an alteration of allele, so it would cause the formation of new alleles. Initially, mutations are seen in a very small percentage of individuals in the populations. As a consequence, the process of mutation is not responsible to ultimate source of variation (Klug *et al.,* 2006; Futuyma, 2013).

The natural selection is the centerpiece of the evolutionary theory. Natural selection shows consistent differences in fitness among phenotypically different classes of biological entities (Endler, 1986). Fitness is a term that refers to changes in survival or reproduction rates of genotypes or individuals in a population over time against natural selection (Klug *et al.,* 2006; Futuyma, 2013).

1.8.1. Genetic Markers

To appreciate history and evolution of populations, it is often essential to examine a large number of polymorphisms (Cavalli-Sforza, 1998). A genetic marker is any visible character or otherwise detectable phenotype, that alleles in individual loci are separated in the Mendelian inheritance. Different applications of genetic markers and whether they are desirable for particular applications should be considered. Some valuable applications of genetic markers in forest trees can be classified as follows (1) Defining mating systems, inbreeding levels and temporal and spatial patterns of genetic variation in stands, (2) Identifying the geographic patterns of genetic diversity, (3) Assuming of taxonomic and phylogenetic relationships between species, (4) Assessing the effects of domestication practices on genetic diversity, including forestry management and tree improvement, (5) Identification of fingerprints and germplasm in breeding and reproduction populations, (6) Establishment of genetic linkage maps, and (7) Marker assisted breeding (White *et al.,* 2007).

The characteristics of the desired genetic markers can be listed as follows : (i) Revealing high level of genetic polymorphism; (ii) Being codominant (distinguishable from heterozygous and homozygous genotypes); (iii) Offering clear distinctive allele features (different alleles can be identified easily); (iv) Representing equal distribution over entire genome; (v) Neutral selection (without pleiotropic effects); (vi) Easy detection (the entire process can be automated); (vii) Having low cost of marker development and genotyping; and (viii) Having high reproducibility (the data can be collected and shared between laboratories) (ix) Not being affected by environmental and developmental variation (White *et al.,* 2007; Xu, 2010).

1.8.2. Microsatellites or SSRs (Simple Sequence Repeats)

Recent evidence with the help of whole genome sequencing reveals the important role of repeats. Tandem repeats (TRs) are repeating DNA sequence that exist contiguously or contemporaneously with each other. Unit or motif is the repeated DNA sequence, and based on the unit length TRs which are classified into two main categories; microsatellite and minisatellites. Microsatellite, or simple sequence repeats (SSRs), are short repeats unit length (repeated units of one to six nucleotides) having one to ten nucleotides, while minisatellites are repeats that have unit length are greater than ten nucleotides (Gemayel *et al.*, 2012). Tandem repeats are evolutionarily suitable due to instability. Mutation rates are 10 to 100,000 times higher than average mutation rates in other parts of genome. The mutation rate is situated between 10^{-3} to 10^{-6} per cell production (i.e., 1 to 10 position of magnitude than point mutations) (Verstrepen *et al.,* 2005). They are commonly found in whole genome, especially in euchromatin of eukaryotes, and in coding and non-coding regions of nuclear and organellar genomes (Phumichai *et al.*, 2015). Repeat polymorphisms usually occur when there is addition or deletion of whole repeat units. Polymerase strand-slippage in DNA replication or recombination errors such as unequal crossing over and gene conversion are causes of these polymorphisms (Vieira *et al.,* 2016).

SSRs are more informative than other molecular markers, including SNPs (Vieira *et al.,* 2016). Microsatellites are categorized based on the type of repetition sequence;

(i) perfect that completely consisting of a single motif; (ii) imperfect if a non-motif repeats between base pairs are formed; (iii) interrupted when a sequence of a few base pairs is added to the motif, or composite created by multiple, contiguous, repeating motifs (Oliveira *et al.*, 2006).

Microsatellites contain all of the properties of the desired genetic markers listed earlier. They are the most well-known and adjustable genetic marker that has numerous applications in population genetics, conservation and evolutionary biology. They have many advantages such as short length ranges, extensions of continuous identical repeats, abundance in plant genomes, high rates of polymorphisms, and presence of Mendelian inheritance. Information obtained from these repeat regions helps to develop statistical procedures for interpopulation comparisons (Schlotterer, 2004). Nearly all microsatellites are found in non-coding region of genome. For that reason variations are independent of natural selection. These features make them ideal genetic marker for conservation and management of plant genetic resource (Mason, 2015). Functional significance of microsatellites is that they may have a neutral effect on genome or may perform important functions in certain species (Oliveira *et al.,* 2006). Therefore, SSRs have been one of the most preferred molecular markers for genotyping plants over the past 20 years (Mason, 2015).

For wild species, SSRs are especially used (i) to determine genetic variation based on genetic distance; (*ii*) to predict gene flow and migration rate, effective population size; and (*iii*) to perform evolutionary studies to clarify intraspecific genetic associations. Besides, SSRs are generally appropriate for (*i*) generating linkage maps; (*ii*) mapping loci included in quantitative traits (QTL); (*iii*) assessing degree of family relationship between genotypes; (*iv*) utilizing marker assisted selection; and (*v*) identifying cultivar DNA fingerprints. SSRs are specifically useful for producing integrated maps for plant species, which is used to create maps of full-sib family connections, and for joining genetic, physical, and sequence-based maps, providing a tool for linking phenotypic and genotypic variation for breeders and geneticists. In parentage analysis, SSRs that have nuclear repeats 3 to 5 nucleotides in length are preferred (Vieira *et al.,* 2016).

1.8.3. Literature Review of *S. alba* and its relative species

Molecular genetic studies on genus *Salix* have been revealed by using microsatellites. Genetic linkage maps have been developed for *Salix*, and many microsatellite loci have been identified in their genomes (Hanley *et al.,* 2002, 2006). Many of them presented cross-species amplification in other *Salix* species, and genus *Populus*. SSR markers have been characterized in several *Salix* species (Lian *et al.* 2001; Stamati *et al.* 2003; Barker e*t al.* 2003; Tuskan *et al.,* 2004; Kikuchi *et al.* 2005; Hoshikawa *et al.,* 2009; King *et al.,* 2010; Lauron-Moreau *et al.,* 2013; Bozzi *et al.,* 2015).

Many scientists used SSR markers to assess the genetic structure and diversity of the *S. reinii*, *S. caprea, S. herbacea, S. eriocephala, S. purpurea, S. viminalis, S. daphnoides* (Lian *et al.,* 2003; Palme *et al.*, 2003; Reisch *et al.,* 2007; Lin *et al.,* 2009; Puschenreiter *et al.,* 2010; Trybush *et al.,* 2012; Singh *et al.,* 2013a; Singh *et al.,* 2013b, Sochor *et al.,* 2013; Perdereau *et al.,* 2014; Berlin *et al.,* 2014; Ukwubile *et al.,* 2014; Singh *et al.,* 2014). Microsatellites were also used to determine clonal diversity and conservation management implications of *S. lanata, S. lapponum, S. herbacea*, *S. hukaoana* and arctic *Salix* spp. (*S. pulchra, S. glauca, S. lanata, S. arbusculoides* and *S. barclayi*) (Stamati *et al.,* 2007; Douhovniko *et al.,* 2010; Kikuchi *et al.,* 2011). Hanley and Karp (2014) developed linkage map in shrubby willows by using SSR markers (sub-genus Vetrix).

On the other hand, there is only one study about genetic diversity and structure of *S. alba* using SSR markers. Genetic analysis of Latvian *S. alba* L. and hybrid populations were evaluated using nuclear and chloroplast DNA markers. Low level of genetic differentiation was found between populations. The diversity level within each population was found to be similar (Rungis *et al.,* 2017).

In Turkey, only *Populus nigra* L has been studied among riparian salicaceous species based on genetic diversity and structure using microsatellite markers (Çiftçi *et al.,* 2017). Until today, studies related to willow have been made with clone bank material of the species. Mainly, the growth performance of clones were assessed

using different criteria. Willow studies conducted in Turkey are specific to genetic variations of willow clones and their adaptability to different regions in terms of rooting percentage, growth performance, vegetation period, stem form, basic density, dry matter, holocellulose production, calorie and protein values (Tunçtaner, 1990). It was found that *S. alba* has good advantage in terms of height and diameter growth. Besides, Tunçtaner (1993) expressed that clone collection of willow should be enriched with native and introduced clones and germplasm collection of indigenous willow. Growth performances and particleboard properties of poplar and willow clones in lake regions were studied based on diameter, height, survival and index values of 24 poplar and 11 willow clones (Tunçtaner *et al.,* 2001). Clones of *salisetum* (*Salix* clone trial) established in Kırsehir-Kocabey in Central Anatolia were compared in terms of growth performance (diameter and height), bole straightness and survival. *S. excelsa* was found the best performer (Toplu *et al.,* 2004). In another study, it was attempted to breed high quality willow sapling for biomass plantations with *S. excelsa* clones identified before as the most suitable clones for biomass production (Uludağ *et al.,* 2004). Genetic diversity of trees is a key component in biodiversity and it has a central role in the resilience and adaptation of ecosystems to climate change. Therefore, evaluation of previous studies is not adequate to develop a program for understanding the genetic structure and conservation of genetic variability of *Salix alba*. Genetic monitoring should be implemented to make better evaluation with regard to efficient breeding and conservation for sustainable *Salix* plantations.

1.9. Justification of the Study

The total area of willow in Turkey is about 2087.28 hectares (877.73 ha pure stand, 1209.28 ha mixed stand with other forest trees) (Velioğlu and Akgül, 2016). *Salix alba* trees grow naturally in a wide range of rivers habitat as forming groups or individually in all geographic regions of Turkey. It has been cultivated for many centuries and transported far beyond of its natural range. They are among the most common plants grown in parks, estates, settlements, and riverbanks. Various varieties have been selected and grown for decorative features. These variants of white willow are generally regarded as subspecies, or varieties without any genetic

data which may support these classification. Correct identification of different *Salix* species sampled from their natural location is necessary for the implementation of ecological and evolutionary researches. Understanding phylogenetic relationships among Turkish *Salix* species can provide useful information when combining with genetic diversity studies for future breeding studies.

Many rivers rising from Turkey flow into different seas. The Göksu and Kızılırmak Rivers two of the major river systems of Turkey were selected for determining the genetic diversity structure of *Salix alba* populations in their natural range. These two rivers flowing to two different seas (Mediterranean and Blacksea, respectively) has no common river basins. The Göksu River arises from two main sources, the Geyik Mountain and the Haydar Mountains which are parts of Taurus Mountains. These two branches merge in Mut and shape Göksu River. Göksu river which is 260 km long flows into Mediterranean Sea southeast of Silifke town (in Mersin province). Kızılırmak River is the longest river (1,355 km long) and starts from Sivas, Imranlı and Kızıldağ at altitudes of 2000 meters. The river passes through Sivas, Kayseri, NevŞehir, KırŞehir, Kırıkkale, Ankara, Aksaray, Çankırı, Çorum and Samsun province and flows into Black Sea. The natural river ecosystems and habitat of *Salix* have been either disappeared or highly fragmented due to various environmental, industrial factors such as building of dams, hydroelectric power stations, sandbars, industrial and urban wastes. These activities have seriously affected the natural population of the species in two river systems. There are 7 and 15 dams and/or hydroelectric power plants on Göksu and Kızılırmak Rivers, respectively (Appendix G).

S. alba is an important commercial lumber species. It is used for wood, cricket bat, archery springs, circles, canoes, wooden kitchenware, sandals, coal and log cabins. White willow bark is also used for medical purposes; tannin to silk, wool and leather; for the production of paint varnish for paint; and for rope construction in Turkey.

Willow is one of the most important species of river ecosystems. It is used commonly as effective phytoremediation tool in the world for river cleaning and ecosystem rehabilitation efforts. There is an increasing interest in using *S. alba* as a short-rotation plantation species in the world. In Turkey, these potential of the

species are not recognized or fully exploited. For this reason, it is necessary to set up programs for the restoration of river ecosystems with *S. alba* and protection of genetic resources of the species. Nevertheless, little information about the genetic diversity and population structure of native *S. alba* in Turkey is available in literature. With the current study, by using microsatellite markers, genetic diversity pattern of the *Salix alba* populations in Turkey has been determined for the first time. Obtained information can be used to initiate effective new conservation strategies and breeding programmes for future generations.

1.10. Aim of the Study

The main objective of the study was to determine genetic diversity structure of *Salix alba* populations in two river systems by using nuclear microsatellite markers.

The specific objectives of the study were:

- To test the capabilities of existing nuclear SSR markers
- To characterize existing genetic resources for *S. alba* in two river systems
- To describe genetic diversity pattern and population differentiation among populations within and between two river systems
- To understand the effect of different river basins (Göksu and Kızılırmak Rivers) on the genetic structure of *S. alba* and how gene flow changes depending on distance among and within rivers
- To provide useful genetic information for breeding, efficient conservation, management and utilization of genetic resources of economically valuable *S. alba*.

As the study outcomes, it was expected that the evolutionary relationships would be put forward, the genetic resources would be better conserved and the lacking of genetic information about *S. alba* species in literature would be provided.

CHAPTER 2

MATERIALS AND METHODS

2.1. Plant Materials

The study was conducted in two river basins, namely Kızılırmak and Göksu rivers. Two hundred and fifty nine *Salix alba* genotypes were sampled from nine populations, 112 genotypes belonging to the Göksu River (four subpopulations) and 147 individuals belonging to the Kızılırmak River (five subpopulations). The sampled populations were determined to represent the "upstream," "middle, "and "downstream" parts of the rivers. Willows are usually clonally reproduced, because of this, a minimum distance (200m) among the genotypes (trees) within a population was taken into consideration. Global positioning system (GPS) was used to obtain location of sampled trees during field work. All sampled leaves of trees were stored in silica gel to avoid distortion of the leaves until DNA extraction. Detailed information was given in Table 2.1 and in Figures 2.1 and 2.2

Table 2.1. Detailed information on studied *S. alba* populations

River	*Population Code*	*Pop No*	*N*	*Location*	*Lattitude (N)*	*Longtitude (E)*	*Altitude (m)*
Göksu	GKSUPSTREAMPOP	1	31	Mut	36°84'03" 36°27'48"	33°17'75" 33°07'24"	246-284
	GKSERMENEKPOP	2	23	Ermenek	36°57'43" 36° 34'25"	33°47'27" 33°03'12"	333-342
	GKSMIDDLEPOP	3	28	Mut	36°56'99" 36°22'21"	33°47'50" 33°25'19"	91-104
	GKSDOWNSTREAMPOP	4	30	Silifke	36°43'53" 36°25'38"	33°76'42" 33°44'47"	27-58
Sub Total			112				
Kızılırmak	KIZUPSTREAMPOP	5	33	Kayseri/Ürgüp	38°83'20" 38°71'19"	35°22'52" 34°67'47"	789-1113
	KIZMIDDLEPOPKIR	6	32	Kırşehir	40°08'78" 39°61'52"	33°48'62" 33°43'52"	640-816
	KIZMIDDLEPOPKRK	7	52	Kırıkkale	39°69'17" 38°22'94"	34°98'85" 32°97'79"	730-1269
	KIZDOWNSTREAMPOPCOR	8	15	Çorum	41°09'71" 41°00'10"	35°75'73" 34°41'80"	358-424
	KIZDOWNSTREAMPOPBAF	9	15	Bafra	41°69'01" 41°60'02"	35°93'65" 35°90'44"	0-2
Sub total			147				
Total			259				

N=Number of trees

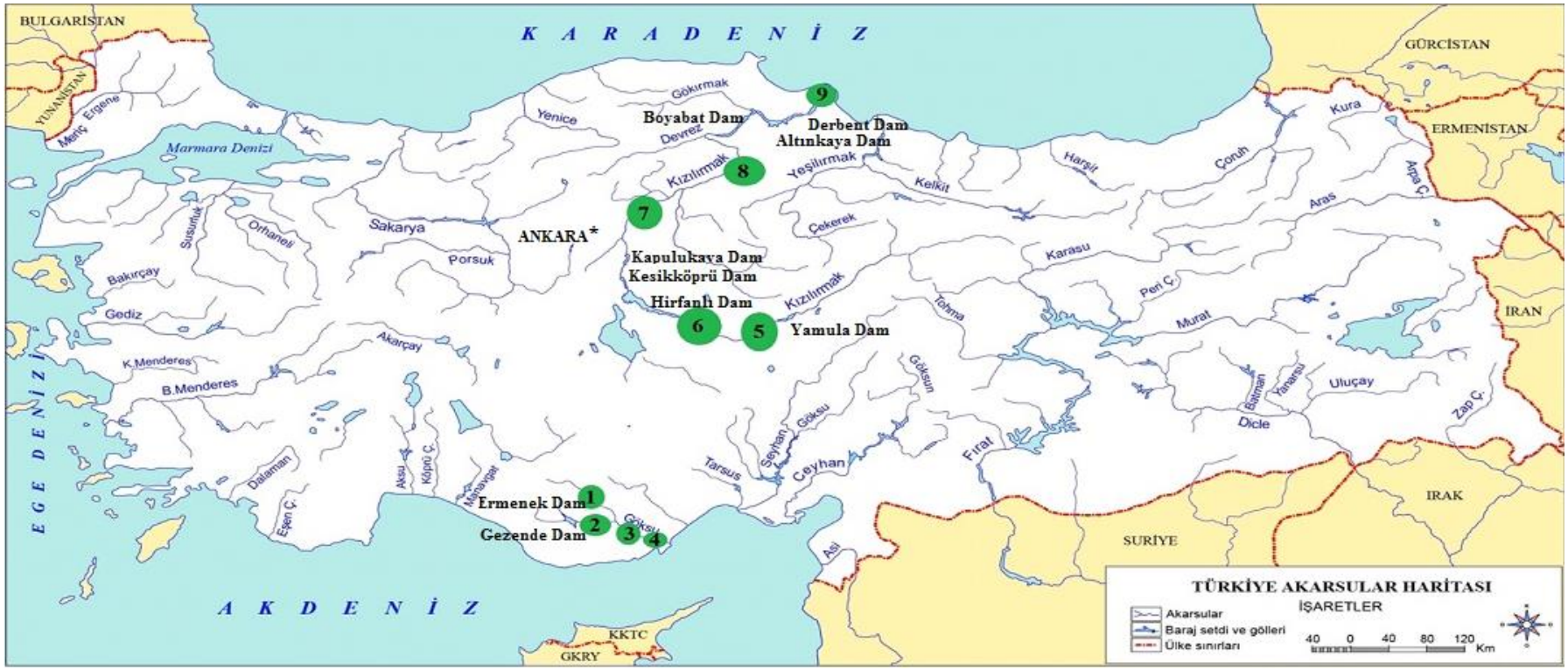


Figure 2.1. Map showing the locations of studied populations (cografyaharita.com). The population numbers 1-4 and 5-9 represent Göksu and Kızılırmak River System, respectively.

Figure 2.2. Views of *Salix alba* habitats from Göksu (top) and Kızılırmak Rivers (bottom), respectively.

2.2. DNA Extraction and Quantification

Silica-gel dried leaves were crushed and powdered with liquid nitrogen in mortar by pestle. The samples were stored in a deep freezer at -80°C until DNA isolation. Genomic DNA was extracted with a modified cetyltrimethylammonium bromide (CTAB) protocol (Doyle and Doyle, 1987). The details information about DNA isolation method and buffers, solutions were provided in Appendix A and B, respectively.

The NanoDrop spectrophotometer (NanoDrop 2000, Thermo Scientific, USA) was used to quantify concentration of DNA. The purity of DNA was evaluated by the optical density (OD) ratios above 1.8 for A260/280 and above 1.5 for A260/230. For each sample, DNA isolation was repeated until getting satisfactory

quantity and quality of DNA to be used in further polymerase chain reactions (PCR). The template DNA concentration was diluted to 20 ng/µl before it was used.

2.3. Methods

2.3.1. SSR Markers

Forty-three microsatellite loci that were developed for different *Salix* species (Lian *et al.,* 2001, Barker *et al.,* 2003, Stamati *et al*., 2003, Lin *et al.,* 2009, Lauron *et al.,* 2013) and 21 microsatellite loci that were developed for various *Populus* species (Van Der Schoot *et al.,* 2000; Smulders *et al.,* 2001; Web Site of International *Populus* Genome Consortium 2014) were firstly tested in *S. alba*. 20 microsatellite primers which well-amplified and showed polymorphic bands were selected. Some primers were tested first time for *S. alba* in this study (Table 2.2).

Different reaction components and their overall effects on a PCR were checked by using 5x HOT FIREPol® Blend Master Mix Ready to Load (Solis BioDyne, Tartu, Estonia) which included 15 mM $MgCl_2$. The details about reaction mixture were provided in Table 2.3. Several PCR cycles and annealing temperatures were also tested in thermocycler (Eppendorf-Master cycler, Eppendorf, Canada). The optimal PCR cycles for the SSR primers were given in detail in Table 2.4.

Amplification products (5µl) were loaded in a 3% agarose gel prepared by 1x TBE (Tris-Borate-EDTA) solution and run in 1XTBE (0.4 M Tris Boric acid EDTA) buffer at ~120 mA for at least 30 minutes. The amplified products were visualized under UV light (Vilber Lourmat, France). Thermo Scientific GeneRuler Low Range DNA Ladder was preferred to determine the product size. Finally, the gels were

Table 2.2. Detailed information about the used microsatellite markers

No	SSR loci	Motif Units	Primers (5'-3')	Allele size range (bp)	Saliceace species	Reference
1	Sare03	$(AC)_{28}$	TCATCATCCCATTCTCATTTG AAATGGTAAGCCCTTGTTGTC	76-139	*Salix reinii*	Lian *et al*., 2001
2	Sare04	$(AC)_{16}C(AC)_{10}$	GACTTCTAGTATTTCTACCCCCTC TATAATTGAGAAAGAAAAAAGACG	89-143		
3	Sare08	$(GT)_{27}$	TATACAATAGGCCCTGGTACC ACCCAGCAAAGCAATAACTGG	134-173		
4	SB24	$[TG]_{21}AG[TG]_3AG$ $[TG]_3$ $AG[TG]_3AGTGAG$ $[TG]_3$	ACTTCAATCTCTCTGTATTCT CTATTTATGGGTTGGTCGATC	114- 281	*Salix burjatica*	Barker *et al*., 2003
5	SB80	-	TAATGGAGTTCACAGTCCTCC ATACAGAGCCCATTTCATCAC	115-143		
6	SB194	-	TGTGAGATAAGATTTGTCGGT CCATAAATAAAAAACGTGAAC	-		
7	SB196	$[GCC]_9$	CTGTTTCCTGCCACTATTACC TATAATCTGTCTCCTTTTGGC	169-184		
8	SB233	$[TA]_2[TGTGCG]_4$ $[TG]_9$	AAATTACCGTCCAACTAAAGA CATTAGCCATGAACAAGTAAA	185-235		
9	SB243	$[GCC]_3ATCATTCCCC[GCC]_4$	ATTCCTTTCTTCATCAGTAGC GACAACGCCATTCACATGACC	102-113		
10	SB265	-	ATTAGGGTTTGTTGCTTGGTG AACATACGTTTCAACGAGAAG	-	-	
11	SB493	-	TTTCTGGATCAATGGAGCTTG CATCTTTCCTTCTTTTACTCC	-	-	
12	W293	-	TGATTGGGCTAAAGATGAAGC AACTCAGCAACCACCAGAAAC	-	-	
13	W784	-	GCACAGATAAAAAATTGGTTG ATATGACTAGGAGGATGTGTT	-	-	
14	gSlMCT011	(CT)11	TTCATCTCCCCGTTCACTTC ACCGTTAGGATGGCATCTCG	-	*Salix lanata*	Stamati *et al*., 2003
15	gSlMCT024	$(CT)_{10}$	TCATTTGCTCGATGAGGTTG GTGGTAGTTGCAAAAGGGGA	-		
16	gSlMCT052	(CT)15/(AG)20	ATTCTTTTTCCACTCGCCAC GGATTGACCCATCTCGATTC	-		

17	WPMS18*	$(GTG)_{13}$	CTTCACATAGGACATAGCAGCATC CACCAGAGTCATCACCAGTTATTG	245-248	*Salix eriocephala*	Lauron *et al.*, 2013
18	PMGC2709*	GA	ATTGTAATTATTGAACACATGCC GTGCAGTTCAGAGTATTGTTG	-	*Salix eriocephala* *Salix purpurea*	Lin *et al.*, 2009
19	PMGC2889*	GA	CCCAAGATCCGATTTTTGGG CACAATGTACAAATCGCTGTC	-		
20	PMGC2163*	(GA)	CAATCGAAGGTAAGGTTAGTG CGTTGGACATAGATCACACG	186-270	*Populus nigra*	IPGC SSR Resource

* Microsatellite loci that were originally derived from The International Populus Genome Consortium

photographed by Alpha Imager Gel Documentation System (Alpha Innotech, San Leandro, CA, USA) (APPENDIX I).

Table 2.3. Optimized PCR mixtures for studied SSR primers

SSR loci	Master Mix (1X)	Primer pairs (10µM)	Water	DNA (20ng/µl)	Total (µl)
Sare03 **Sare04** **Sare08** **SB80** **gSlMCT011** **gSlMCT052** **PMGC2163***	5 µl	0.8+0.8 µl	7.4 µl	6 µl	20 µl
SB194 **SB196** **SB233** **SB265** **W293** **W784**	4 µl	0.5+0.5 µl	10 µl	5 µl	
SB24 **SB243** **SB493** **gSlMCT024** **WPMS18*** **PMGC2709*** **PMGC2889***	5 µl	0.5+0.5 µl	9 µl		

* Microsatellite loci that were originally derived from The International Populus Genome Consortium

Table 2.4. PCR amplification conditions for studied SSR primers

Primer Name	Step	Temperature	Time	Number of Cycles	Description
Sare03 Sare04 SB80	1	94°C	3 min.	1	Denaturation
	2	94°C	1 min.	30	Denaturation
		Ta	30 sec.		Annealing
		72°C	1 min.		Extension
	3	72 °C	5 min.	1	Final Extension
Sare08 SB24 SB243	1	94°C	3 min.	1	Denaturation
	2	94°C	1 min.	30	Denaturation
		Ta	45 sec.		Annealing
		72°C	1 min.		Extension
	3	72 °C	10 min.	1	Final Extension
SB194 SB493 W293 W784	1	94°C	3 min.	1	Denaturation
	2	94°C	30 sec.	30	Denaturation
		Ta	45 sec.		Annealing
		72°C	45 sec.		Extension
	3	72 °C	10 min.	1	Final Extension
SB196 SB233 SB265	1	94°C	3 min.	1	Denaturation
	2	94°C	1 min.	30	Denaturation
		Ta	30 sec.		Annealing
		72°C	1 min.		Extension
	3	72 °C	5 min.	1	Final Extension
gSlMCT011 gSlMCT024 gSlMCT052	1	94°C	3 min.	1	Denaturation
	2	94°C	40 sec.	30	Denaturation
		Ta	30 sec.		Annealing
		72°C	1 min.		Extension
	3	72 °C	10 min.	1	Final Extension
WPMS18* PMGC2709* PMGC2889* PMGC2163*	1	94°C	3 min.	1	Denaturation
	2	94°C	40 sec.	30	Denaturation
		Ta	30 sec.		Annealing
		72°C	30 sec.		Extension
	3	72 °C	10 min.	1	Final Extension

* Microsatellite loci that were originally derived from The International *Populus* Genome Consortium

2.3.2. Data Collection and Formatting

After optimization of reaction components and PCR cycles, forward primers were resynthesized fluorescently by SACEM Company (Ankara) in order to discriminate PCR (polymerase chain reaction) products while verifying fragment analysis results. Fam, Hex and Tamra colored dyes were selected to label the 5' end-of forward primers (Table 2.5). Genomic DNAs of 259 genotypes from two river system were amplified by using fluorescently labeled and unlabeled primers for each of 20 microsatellite loci. Annealing temperatures of 20 SSR primers were given in Table 2.5. After amplification, the products were merged based on different colors and sizes of primers. Therefore, at least three different SSR primers‟ products were analyzed in one tube at a time (Figure 2.4).

Table 2.5 Fluorescent dye and annealing temperature (Ta) of 20 SSR primers

SSR loci		Fluorescent Dye			Ta (^{o}C)
		FAM	HEX	TAMRA	
Group 1	SB24	✓			55
	W784	✓			50
	SB265		✓		53
Group 2	W293	✓			55
	gSlMCT024	✓			57
	SB196		✓		57
	Sare08			✓	58
Group 3	SB493	✓			55
	PMGC2709*		✓		52
	WPMS18*			✓	55
Group 4	Sare04	✓			50
	PMGC2889*		✓		52
	SB243			✓	55
Group 5	SB233	✓			57
	PMGC2163*		✓		55
	gSlMCT052		✓		58
	SB194			✓	50
Group 6	Sare03		✓		55
	gSlMCT011		✓		58
	SB80			✓	58

Assay procedure was made by the BM Labosis Company (Çankaya, Ankara). Analysis of the samples were performed with the Applied Biosystems 3730 XL DNA Analyzer (Applied Biosystems, Foster City, CA, USA), using an internal standard size marker The GeneScan ROX labeled 400HD (APPENDIX C). Allele sizes were checked and scored manually from electropherograms using the Peak Scanner Software 2.0 (Applied Biosystems Inc., Foster City, CA). Sizing Default NPP (NPP: No Primer Peaks) was chosen for the analyzing method in the software. In Figure 2.5, electrophenograms of two different genotypes were provided as an example. After determining the actual allele of each peak, raw data were formed in an Excel file to be used for further statistical analysis (APPENDIX D).

Data from 259 genotypes sampled from Göksu and Kızılırmak Rivers that included nine populations screened with 20 SSR primers were transformed into various formats, so that the data could be properly analysed. GENEPOP (Raymond and Rousset, 1995; Rousset, 2008), GenAlEx (Peakall and Smouse 2006, 2012), GDA (Lewis and Zaykin, 2001) and GenClone (Arnaud-Haond and Belkhir 2007) were the data formats for the analysis. GENEPOP software was used in order to convert an input data file from Genepop to Arlequin (Excoffier and Lischer, 2010), FSTAT (Goudet, 1995), MICRO-CHECKER (Oosterhout *et al*., 2004) and CERVUS (Marshall *et al*. 1998; Kalinowski *et al.* 2007) input data files format. The STRUCTURE (Pritchard *et al*., 2000) data file format was prepared using TextPad 8 in Arlequin input file format. Examples of data file formats and the statstics used in thesis were provided in Appendicies E and F, respectively.

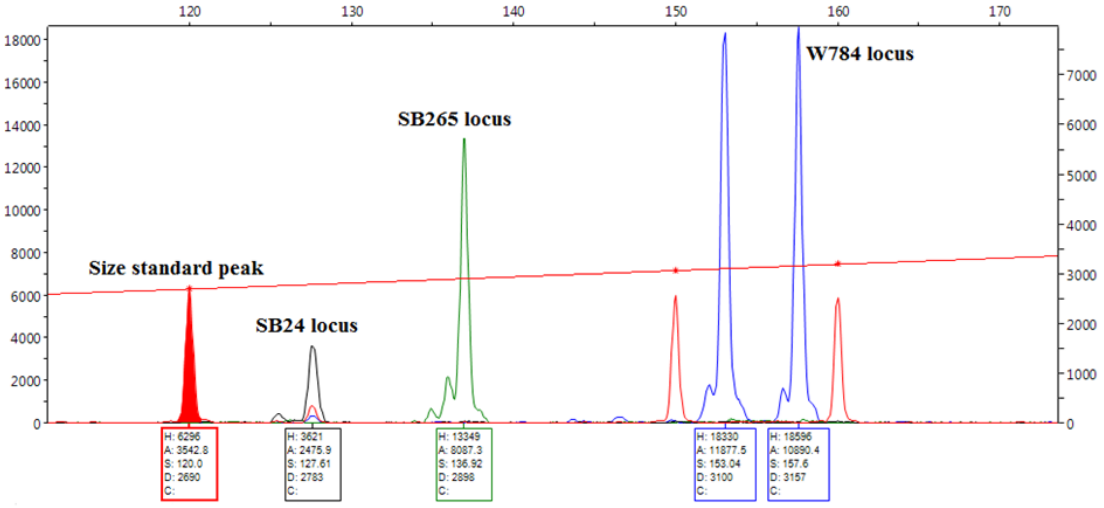


Figure 2.3. Electropherogram showing the profiles of SSR primers (SB24, SB265 and W784) for three different loci, amplified separately using three different dyes.

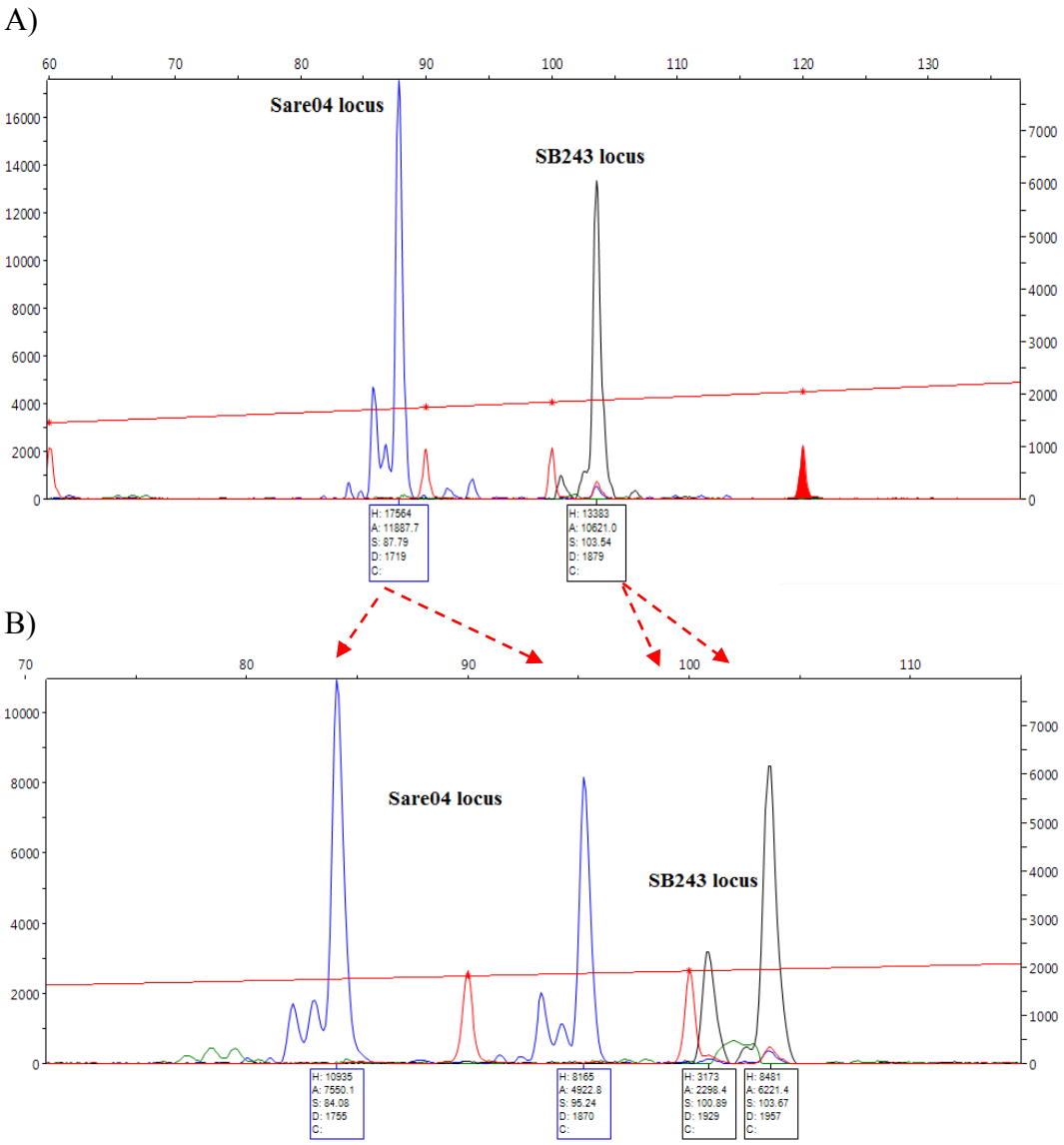


Figure 2.4. Electropherogram of the homozygote (A) and heterozygote (B) genotype profiles for Sare04 and SB243 loci.

2.4. Analysis of Data

2.4.1. Microsatellite Markers Source Quality

The MICRO-CHECKER software with Bonferroni-adjusted 95% confidence interval was used to detect the existence of genotyping failure such as null alleles, stuttering, large allele dropout and typographic errors by cause of DNA degradation, low DNA

concentrations and primer-site mutations. The program also tested the frequency of null alleles. According to Brookfield (1996), null allele estimator1 was preferred in order to calculate the null allele frequency (*r*), since, it was thought that the non-amplified samples are 'artifacts' and could be either null allele homozygotes, degraded DNA , or problems with the PCR.

Probability test (Guo and Thompson, 1992) option for each locus in the populations was used to test Hardy–Weinberg equilibrium (HWE) via GENEPOP software. Exact P-Values were estimated by the Markov chain method (parameters for all tests; 1000 dememorizations, 100 batches, 1000 iterations each). Statistical significance was estimated using a sequential Bonferroni correction for multiple comparisons (Holm, 1979).

Linkage disequilibrium (LD) implies that alleles are not randomly associated in different loci. LD occurs as a result of differentiation among populations and isolation by distance, asexual reproduction, linkage between the alleles, selection and genetic drift (Agapow and Burt, 2001). LD was estimated implemented in R *poppr* package (Kamvar *et al.,* 2014) based on index of association (Brown *et al.* 1980) proposed by Agapow and Burt (2001) and varies from -1 to +1.

It was important that clone correction was necessary for correctly interpreting population genetic analysis of species which reproduce asexually (Kamvar *et al*., 2014). GenClone 2.0 software was used to determine the distinct and identical multilocus genotypes (MLGs) among all genotypes. These statistics are useful for finding suitability of loci for further analysis.

2.4.2. Assessment of the Genetic Diversity Analysis in Loci and Populations

Genetic diversity parameters such as number of alleles (Na), number of effective alleles (Ne), Shannon‟s information index (I), observed heterozygosity (Ho), expected heterozygosity (He), fixation index (F), the probability of identity (PI), percentage of polymorphic loci (%P), F-Statistics (Fis, Fit, Fst) and the number of migrants (Nm) of loci were estimated using GenAlEx software (Peakall and Smouse

2006, 2012). The software enriched understanding of the data by calculating standard error.

The number of alleles stands for observed alleles which were found in the studied subpopulation. The effective number of alleles is used to indicate the number of alleles that are equally relevant to the expected heterozygosity as in the studied population. Private alleles are specific to a single set of population and occur at any frequency.

Additional parameters were calculated such as Allelic Richness (Ar), Polymorphic Information Content (PIC) for deciding effectiveness of a primer together with Probability of Identity (PI) and Shannon‟s information index (I). These measures vary in formulations, ecological and evolutionary interpretations, and mathematical frameworks in which they are applied (Allendorf, 1986; Toro *et al.,* 2009, Caballero and Garci´a-Dorado, 2013). Allelic richness (Ar) (Petit *et al*., 1998) was evaluated using FSTAT software. Observed heterozygosity (Ho) and expected heterozygosity (He) are measured considering allele frequency rather than the number of alleles and can provide information about fitness of individuals (Szulkin *et al.,* 2010). The average number of alleles per locus shows allelic richness (allelic diversity), but it also gives information on how population will respond to environmental changes. Therefore, Ar is important indicator for showing the evolutionary potential of a population (Allendorf *et al.,* 2012; Caballero and Garci´a-Dorado, 2013). It has been proposed that this measure has key value in the conservation and management of populations (Foulley and Ollivier, 2006). Polymorphic information content (PIC) values (Botstein *et al.* 1980), and Hardy-Weinberg chi-square statistic for all loci were calculated using Cervus software (Marshall *et al.* 1998; Kalinowski *et al.* 2007). PIC measures usefulness of markers and is most often used to describe genotypic variation. Furthermore, is the best measure to calculate the efficiency of a marker linked to the observed heterozygosity. This is mainly for assessing genetic variation in population that will shed light on the evolutionary pressure on allele. The PIC value will be almost zero, if there are no allelic variations. It will only reach a maximum of 1.0 if there are new alleles in genotype, which is a rare phenomenon.

The probability of identity (PI) predicts that two different individuals in the dataset, may have the same multilocus genotype. This information concerns the primers applicability and informativeness of primers in genetic analysis due to the assessment the resolution power of the selected combination of primers (Taberlet and Luikart, 1999; Waits *et al.,* 2001). Shannon‟s information index is equal to the Shannon-Weaver Index of ecology. It is thought that Shannon‟s information index may be a better measure of allelic and genetic diversity since it is not bounded by 1 unlike He (Sherwin *et al*., 2006).

The observed heterozygosity is defined as the number of individuals that are heterozygous per loci. The expected heterozygosity (gene diversity) is calculated from the individual allele frequencies (Nei, 1987). Fixation index (F) is calculated on a per locus basis and ranges from -1 to +1. Positive values indicate inbreeding, while negative values indicate excess of heterozygosity. Chi-square statistic is used to test whether subpopulation is in HWE at a locus.

F-statistics (Fis, Fit, and Fst) are parameters used to explain three levels of inbreeding (Nei, 1977). These values define the degree of reduction in heterozygosity when compared to the Hardy-Weinberg proportions (Allendorf and Luikart, 2007). The number of migrants per generation (*Nm*) determines an evaluation of gene flow among subpopulations. It is obtained through F_{ST} values.

Percentage of polymorphic loci per population (%P) is a measure used to quantify genetic diversity in a population. Subpopulations that have experienced bottleneck will show an excess of heterozygosity over the expected values in many loci, since alleles usually disappear faster than heterozygosity during bottlenecks. Garza-Williamson index (Garza and Williamson, 2001) (called M-ratio, the ratio of the number of alleles to range in allele size), was applied to test if there is an indication of genetic bottlenecks in the past (>100 generations). If seven or more loci are examined and the Garza-Williamsion index is lower than the critical value (Mc) of 0.68. Mc obtained from simulations based on the empirical data in bottlenecked populations, it indicates a reduction in population size (Garza and Williamson, 2001; Williamson-Natesan, 2005). It is more effective to detect genetic bottleneck, if the bottleneck lasted for several generations or if the population provides a rapid

demographic improvement (Williamson-Natesan, 2005). The index was estimated by ARLEQUIN.

To be able to apply numerical analysis to latitude and longitude values, first they were converted from "Degrees/Minutes/Seconds (DMS)" unit to "Decimal Degrees (DD)" unit. A linear model is an equation which shows the linear relationships between a dependent and independent variable. Both latitude, longitude and altitude evaluated as independent variables for He and Ho. Analysis of variance (ANOVA) was conducted to understand if they have a significant effect on He and Ho.

2.4.3. Assessment of Populations Genetic Structure Analysis

Four methods were tested for determining genetic structure of the studied *S. alba* populations. Firstly, to evaluate a measure of genetic differentiation among the nine populations, the population pairwise Fst values (Slatkin, 1995), their statistical significance and number of migrants (Nm) were estimated by using ARLEQUIN with number of different alleles distance method (number of permutations, 1000). Gene flow (the number of migrants) is migration of a recipient population from another population with a different allele frequency. It is inversely related to Fst. This means that migration strongly prevents population differentiation.

Besides, GDA software was used to construct phenogram based on Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster analysis using the similarity matrix derived from coancestry identity (Nei, 1978).

Furthermore, the STRUCTURE software (Pritchard *et al.,* 2000; Falush *et al.* 2003, 2007; Hubisz *et al.,* 2009) was utilized to determine population structure. The software is used for assigning individuals to populations and estimating the most likely number of different genetic groups based on allelic frequencies. Bayesian clustering methods were applied in two different assumptions; with a priory identification associated with population location and without a priory identification of specific grouping. Admixture ancestry model and correlated allele frequency model (with $\lambda = 1$) were used in all runs. Run parameters consisted of 10 replicates

each with 250,000 Markov Chain Monte Carlo (MCMC) replications after 50000 burning length for K = 1 to 9 clusters. The mean log-likelihood and the Δ K statistics (true number of clusters) (Evanno *et al.,* 2005) were calculated and graphic representation of these statistics was obtained with the help of the STRUCTURE HARVESTER which is a web-based software (Earl and vonHodt, 2012). The CLUMPP (CLUster Matching and Permutation Program) software was then used to find the average membership coefficient matrices of 10 replicate cluster analysis of the same K (Jakobsson and Rosenberg 2007). Output data of the CLUMPP was directly used as input data into the POPHELPER program (Francis, 2017) to display graphical representation of population clusters. The actual number of different genetic groups were determined by selecting a higher number of clusters, where the majority of individuals placed into these clusters with larger than 0.5 membership value.

Lastly, Principal Coordinate Analysis (PCoA= Classical Multidimensional Scaling, CMDS) was used for getting further confirmation regarding the cluster analysis results. The software helps to discover and visualize not only the similarities, but also dissimilarities of the data. PCoA was composed using the covariance matrix with data standardization obtained from the pairwise Fst values based on Nei Genetic Distance in GenAlEx software. Moreover, the program presented a graphical representation of genetic distance among populations.

2.5. Analyses of Molecular Variance (AMOVA)

Analyses of Molecular Variance was performed both to evaluate the amount of genetic variation within populations, among populations within a river basin, and among river basins of *S. alba* accessions in a hierarchical form. Besides, it was used to estimate of fixation indices (Fct, Fsc and Fst) based on the Fst values (number of different alleles) with 1,000 permutations using ARLEQUIN. It was recommended that Fst should be preferred for microsatellites, since basic assumptions of the simple step wise mutation are seldom held in natural populations. The locus-by-locus AMOVA (significance levels were tested with 1000 permutations) was preferred

because firstly it allowed the analysis to be performed separately for each locus, and secondly a population with missing data was handled.

To confirm the population subdivisions that were assumed by STRUCTURE, AMOVA was again performed separately with ARLEQUIN for Göksu and Kızılırmak Rivers.

CHAPTER 3

RESULTS

3.1. SSR Primers Discovery and Data Assessment

A total of 64 SSR primers were screened. Of these, 20 primers were found to be useful for detecting polymorphism in 259 genotypes. Percentage of missing data is found to be 1.3%. As reported by results of MICRO-CHECKER software, there was no confirmation on mistyped allele sizes and typographic errors, deviations from a regular repeat motif. Likewise, scoring errors due to large allele drop-out were not observed in any locus across the nine populations. The loci SB80 in the Ermenek population- Göksu, and the loci Sare04 in the upstream subpopulation- Kızılırmak had stuttering that might have led to in scoring errors, as is indicated by the highly significant shortage of heterozygote genotypes that was evident from alleles of one repeat unit difference. Taking into account these corrections, allele size range and detected number of alleles were given in the Table 3.1.

Table 3.2. SSR loci and allele size ranges of them

SSR Locus	Allele Size Range	Number of Alleles	Alleles
Sare03	77-117	20	77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 111 ,113, 115, 117
Sare04	76-180	24	76, 82, 84, 86, 88, 90, 92, 94, 96 ,98, 100, 102, 104, 106, 108, 110, 112, 114, 120, 122, 124, 126, 130, 180
Sare08	121-191	21	121, 127, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 153, 155, 157, 159, 161, 163, 167, 169, 191
SB24	114-160	16	114, 116, 118, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 158, 160
SB80	115-121	4	115, 117, 119, 121
SB194	109-115	3	109, 113, 115
SB196	167-179	4	167, 170, 173, 179
SB233	170-220	20	170, 172, 176, 178, 180, 184, 186, 190, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220
SB243	101-113	5	101, 104, 107, 110, 113
SB265	109-139	5	109, 133, 135, 137, 139
SB493	208-244	9	208, 210, 212, 214, 216, 226, 238, 242, 244
W293	108-136	11	108, 112, 114, 116, 118, 120, 122, 124, 126, 130, 136
W784	148-168	6	148, 150, 152, 158, 162, 168
gSlMCT011	274-400	7	274, 276, 280, 284, 286, 290, 400
gSlMCT024	288-304	8	288, 290, 294, 296, 298, 300, 302, 304
gSlMCT052	185-281	24	185, 189, 191, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 225, 227, 231, 243, 263, 271, 281
WPMS18	211-232	7	211, 214, 217, 220, 223, 226, 232
PMGC2709	160-224	22	160, 166, 172, 180, 184, 186, 188, 190, 192, 194, 196, 198, 200 202, 204, 208, 210, 212, 214, 216, 218, 224
PMGC2889	174-222	20	174, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 222
PMGC2163	186-224	14	186, 188, 194, 198, 200, 204, 206, 208, 210, 212, 214, 216, 218, 224

Table 3.2. Estimated null allele frequencies of SSR loci for the studied *S.alba* populations

SSR Locus	GKSUP STREAMPOP	GKS ERMENEKPOP	GKS MIDDLEPOP	GKSDOWN STREAM	KIZUP STREAMPOP	KIZMIDDLE POPKIR	KIZMIDDLE POPKRK	KIZDOWNSTREAM POPCOR	KIZDOWNSTREAM POPBAF
Sare03	0.00	-0.01	0.16*	0.08*	0.11*	0.00	-0.03	-0.05	0.08
Sare04	0.01	-0.04	-0.05	-0.09	0.15*	0.15	0.01	0.02	0.08
Sare08	0.03	0.07	-0.11	-0.01	0.02	0.00	0.00	0.06	0.23*
SB24	-0.06	-0.18	-0.08	-0.07	0.07	0.05*	0.14*	0.21*	0.08
SB80	0.01	0.21*	-0.05	0.08*	0.00	-0.05	0.00	-0.08	-0.04
SB194	0.04	-0.06	-0.09	0.02	0.04	-0.03	-0.19	-0.09	0.12*
SB196	0.00	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	0.00
SB233	0.02	0.00	0.08	0.08	-0.04	-0.03	-0.05	-0.05	-0.06
SB243	-0.13	-0.17	-0.08	-0.16	0.01	-0.15	-0.13	-0.12	-0.10
SB265	0.03	0.04	0.01	0.03	-0.03	0.02	-0.03	0.02	0.05
SB493	0.04	0.03	-0.06	-0.04	-0.06	-0.04	-0.10	-0.02	0.00
W293	-0.10	-0.24	-0.17	-0.22	-0.17	-0.16	-0.18	-0.28	-0.15
W784	-0.28	-0.25	-0.28	-0.32	-0.31	-0.26	-0.30	-0.27	-0.23
gSlMCT011	0.05*	0.00	0.18*	0.00	0.04	-0.02	0.04	0.04	0.05
gSlMCT024	0.07	0.13*	0.06	0.00	-0.08	0.04	-0.02	0.17*	0.17*
gSlMCT052	-0.02	0.04	0.16*	0.03	0.08*	0.03	0.00	-0.01	0.12*
WPMS18	0.08*	0.02	-0.04	0.07	0.17*	0.10*	0.08*	0.00	0.10*
PMGC2709	0.00	0.01	0.05	0.05	0.04	0.00	0.01	-0.04	-0.06
PMGC2889	0.10*	-0.06	0.00	0.00	-0.03	-0.11	-0.06	-0.10	-0.13
PMGC2163	-0.17	-0.19	-0.06	-0.12	-0.15	-0.19	-0.05	-0.24	-0.20

* Null alleles may be present at these loci based on Brookfield1 equation (Brookfield, 1996).

Based on the studied 20 SSR primers, it is appeared that there were some loci with slightly high null allele frequencies which were specific to a certain populations (Table 3.2). Because of this, all analyses were carried out by using the loci with and without null alleles in order to assess the genetic diversity and genetic differentiation in the current study. Significant change in the statistical results in two analysis were not seen. Therefore, the results were given assuming no null alleles.

Linkage disequilibrium (LD) analysis were assessed for SSR studied loci. The result revealed that all loci were in weak LD, with an r_d of 0.0219 ($p = 0.001$). The r_d value does not fall into the expected range from permutations (Figures 3.1 and 3.2). In order to understand whether there is any physical linkage, locations of the markers on the chromosomes were confirmed by checking the genetic linkage maps. The loci SB194, SB243, SB24, SB80, and W784 were appeared to be on the different chromosome of *S. viminalis* (Hanley *et al.,* 2002). PMGC2709, PMGC2889, PMGC2163 and WPMS18 were located on different chromosome of *Populus nigra* (Gaudet *et al.,* 2008).

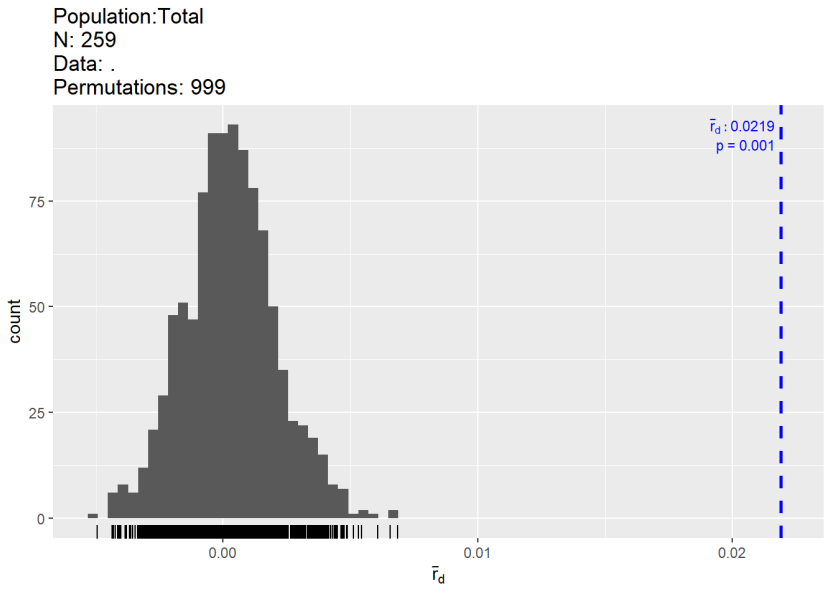


Figure 3.1. Estimation of overall linkage disequilibrium. The r_d is standardized index of association. Blue dashed line indicates observed value.

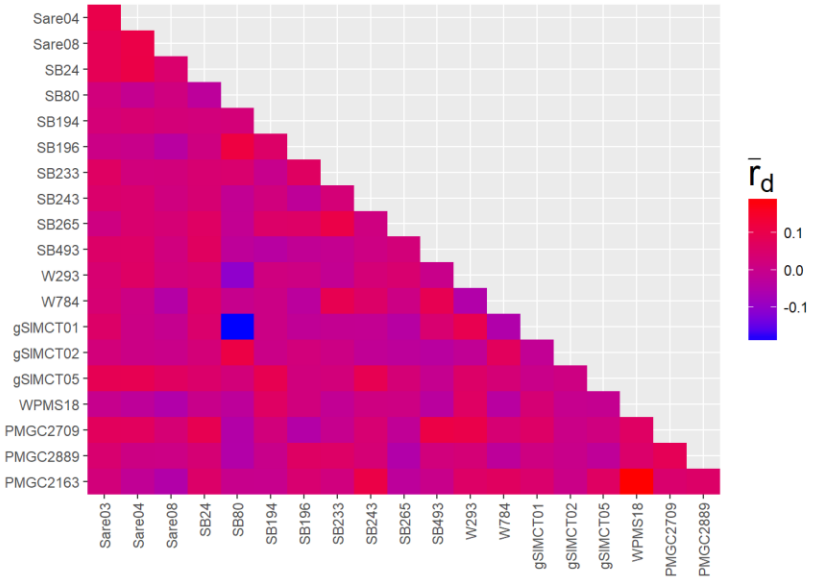


Figure 3.2. Heatmap of pairwise standardized index of association over the studied 20 loci.

According to GenClone 2.0 software, there is no duplicated genotype among studied 259 individuals. Therefore, distinct multilocus genotypes (MLGs) were discriminated. This result suggest that *S. alba* predominantly reproduce sexually.

3.2. Genetic Diversity of Loci

When descriptive statistics by loci were examined, it was observed that all loci were polymorphic (Table 3.3). On the other hand, the locus SB196 was monomorphic in the upstream population, the middle population- Kırıkkale and the downstream population- Çorum. In addition, the locus WPMS18 was also monomorphic in the downstream population- Çorum of the Kızılırmak River. The average of mean number of different alleles was estimated as 6.98 alleles per locus (range 1.67-14.00). Nevertheless, mean number of effective alleles was found 3.61 alleles per locus (range 1.05-7.83). It was found that the loci Sare03 and gSlMCT052 had the highest value of observed and effective number of alleles.

Table 3.3 Descriptive statistics by loci

SSR Locus	N	Na±Se	Ne±Se	Ar	PIC	PI	I±Se	Ho±Se	He±Se	F±Se	HWE
Sare03	28.78 ±3.70	14.00 ±0.62	7.83 ±0.49	10.59	0.92	0.03	2.29 ±0.04	0.80 ±0.04	0.87 ±0.01	0.08 ±0.05	ND
Sare04	28.78 ±3.70	11.78 ±1.01	6.19 ±0.47	9.92	0.90	0.05	2.07 ±0.06	0.78 ±0.05	0.83 ±0.01	0.06 ±0.06	***
Sare08	28.78 ±3.70	11.33 ±0.96	6.51 ±0.43	9.03	0.88	0.05	2.07 ±0.08	0.78 ±0.06	0.84 ±0.01	0.07 ±0.07	NS
SB24	28.78 ±3.70	8.56 ±0.47	4.65 ±0.37	7.72	0.85	0.08	1.75 ±0.07	0.74 ±0.08	0.77 ±0.02	0.04 ±0.10	***
SB80	23.33 ±3.74	3.11 ±0.26	1.74 ±0.18	3.16	0.35	0.48	0.66 ±0.12	0.35 ±0.08	0.37 ±0.07	0.01 ±0.09	NS
SB194	28.78 ±3.70	3.00 ±0.00	2.23 ±0.06	2.82	0.47	0.29	0.89 ±0.02	0.59 ±0.05	0.55 ±0.01	-0.08 ±0.09	NS
SB196	28.78 ±3.70	1.67 ±0.17	1.05 ±0.02	1.37	0.05	0.91	0.11 ±0.03	0.05 ±0.02	0.05 ±0.02	-0.04 ±0.01	ND
SB233	28.67 ±3.70	10.78 ±0.97	3.79 ±0.32	7.15	0.74	0.11	1.70 ±0.08	0.73 ±0.03	0.72 ±0.02	-0.02 ±0.04	NS
SB243	28.67 ±3.70	4.56 ±0.18	3.07 ±0.12	3.90	0.64	0.17	1.23 ±0.04	0.86 ±0.03	0.67 ±0.01	-0.29 ±0.05	***
SB265	28.78 ±3.70	3.11 ±0.26	1.47 ±0.08	2.66	0.30	0.53	0.54 ±0.04	0.28 ±0.04	0.30 ±0.04	0.08 ±0.05	NS
SB493	28.78 ±3.70	4.22 ±0.46	2.38 ±0.19	3.70	0.55	0.29	0.96 ±0.10	0.59 ±0.06	0.55 ±0.05	-0.08 ±0.04	NS
W293	28.67 ±3.70	6.44 ±0.58	2.88 ±0.14	5.39	0.66	0.18	1.29 ±0.06	0.95 ±0.02	0.65 ±0.02	-0.48 ±0.05	***
W784	28.78 ±3.70	3.56 ±0.24	2.30 ±0.06	2.88	0.46	0.29	0.91 ±0.03	1.00 ±0.00	0.56 ±0.01	-0.78 ±0.04	***
gSlMCT011	28.78 ±3.70	3.89 ±0.68	1.44 ±0.11	3.19	0.28	0.58	0.56 ±0.12	0.21 ±0.06	0.27 ±0.06	0.23 ±0.11	***
gSlMCT024	27.56 ±3.59	4.67 ±0.17	3.48 ±0.20	4.39	0.71	0.14	1.33 ±0.05	0.60 ±0.06	0.70 ±0.02	0.15 ±0.07	**
gSlMCT052	28.22 ±3.42	13.67 ±1.08	6.80 ±0.56	9.57	0.88	0.04	2.16 ±0.08	0.75 ±0.03	0.84 ±0.02	0.10 ±0.04	NS
WPMS18	28.78 ±3.70	4.00 ±0.53	1.54 ±0.14	3.30	0.34	0.53	0.63 ±0.11	0.23 ±0.06	0.31 ±0.06	0.32 ±0.08	***
PMGC2709	28.78 ±3.70	10.89 ±0.9	4.97 ±0.59	8.90	0.86	0.08	1.87 ±0.09	0.76 ±0.03	0.77 ±0.03	0.02 ±0.03	NS
PMGC2889	28.78 ±3.70	9.89 ±0.70	5.00 ±0.30	7.45	0.81	0.07	1.85 ±0.05	0.87 ±0.04	0.80 ±0.01	-0.10 ±0.05	**
PMGC2163	28.78 ±3.70	6.56 ±0.88	2.82 ±0.19	5.45	0.63	0.18	1.29 ±0.06	0.88 ±0.04	0.63 ±0.02	-0.40 ±0.06	***
Mean	28.40 ±0.78	6.98 ±0.32	3.61 ±0.16		0.61	0.25	1.31 ±0.05	0.64 ±0.02	0.60 ±0.02	-0.06 ±0.02	

N=mean number of individuals with amplification, Na=mean number of different alleles, Ne=mean number of effective alleles, Ar=allelic richness, PIC=polymorphic information content, PI= The probability of identity, I=Shannon Index, Ho=observed heterozygosity, He=expected heterozygosity, F=fixation index, HWE=Hardy Weinberg equilibrium (ND: non-deviating, NS: non-significant,***:p<0,001,**:p<0,01,*:p<0,05)

Allelic richness (Ar), Polymorphic Information Content (PIC), Probability of identity (PI) and Shannon Index (I) indicated the informativeness of each loci. Allelic richness varied from 1.37 (SB196) to 10.59 (Sare03). PIC values for allelic diversity varied from 0.05 to 0.92 with a mean value of 0.61. Primers with a PIC value greater than 0.5 are the most useful markers in the genetic diversity studies irrespective of the species specificity. It was seen that there was only one locus (SB196) classified as low informative markers with the value of PIC < 0.25. Six of the 19 markers were evaluated moderately informative and the remaining markers had the PIC values exceeding 0.5. For all loci, the probability of identity (PI) varied from 0.03 (for Sare03 to 0.91 (for SB196). The lowest PI value indicated that the locus Sare03 showed adequate power and resolution for the analyses of populations. The 12 of 20 microsatellites were informative, with a probability of identification (PI) of less than 0.20. The Shannon Index values ranged from 0.11 to 2.29 with highest values were found Sare03, Sare04, Sare08, and gSlMCT052. The changes in allelic richness values, Shannon Index, Polymorphic Information Content (PIC), and Probability of identity (PI) among different SSR loci are parallel to change of heterozygosity values. As a result, it was clearly seen that the most diverse loci were Sare03, Sare04, Sare08, and gSlMCT052 (Table 3.3).

The observed heterozygosity for single locus across subpopulations varied from 0.05 for SB196 to 1.00 in W784 with an average of 0.64, while expected heterozygosities ranged from 0.05 to 0.87 with an average of 0.60. It was found that the loci Sare04, SB24, SB243, W293, W784, gSlMCT011, gSlMCT024, WPMS18, PMGC2889, and PMGC2163 showed significant deviations from Hardy-Weinberg equilibrium (Table 3.3). The excess of heterozygosity was generally observed for 9 of 20 loci; whereas, rest of loci had positive fixation index values (Table 3.3). The highest Fst values were found for loci SB24, SB80, SB265, SB493 and PMGC2709. Accordingly, these loci contributed greatly to the differentiation of populations. The mean number of migrants (Nm) was 5.00. Locus W784 had highest number of migrants, whereas, the loci SB265 had the lowest (Table 3.4).

Table 3.4. The estimated F-Statistics for each loci

Locus	Fis	Fit	Fst	Nm
Sare03	0.08*	0.14*	0.06	3.67
Sare04	0.06*	0.14*	0.09	2.55
Sare08	0.07*	0.12*	0.06	3.99
SB24	0.04*	0.15*	0.12	1.90
SB80	0.03	0.17	0.14	1.57
SB194	-0.08*	-0.06*	0.02	14.27
SB196	-0.05	-0.02	0.03	7.96
SB233	-0.01	0.06	0.08	3.04
SB243	-0.28*	-0.24*	0.03	6.96
SB265	0.06	0.20	0.15	1.42
SB493	-0.08	0.03	0.10	2.24
W293	-0.47*	-0.39*	0.05	4.49
W784	-0.77*	-0.75*	0.01	17.82
gSlMCT011	0.21	0.26	0.06	3.72
gSlMCT024	0.14	0.20	0.07	3.43
gSlMCT052	0.11*	0.16*	0.06	4.05
WPMS18	0.27	0.33	0.08	3.00
PMGC2709	0.01	0.13	0.11	1.97
PMGC2889	-0.10*	-0.06*	0.03	7.57
PMGC2163	-0.40*	-0.32*	0.05	4.32
Mean (SE)	-0.06 ±0.06	0.01 ±0.06	0.07 ±0.01	5.00 ±0.95

Fis=the inbreeding coefficient within individuals, relative to the population, Fit=the inbreeding coefficient within total population, Fst=the inbreeding coefficient within subpopulations, Nm= the number of migrants (***:p<0, 001,**:p<0,01,*:p<0,05).

3.3. Genetic Diversity of Populations

As reported by genetic diversity parameters for each population of the Göksu and Kızılırmak River system, all populations had high polymorphism percentage (100% for Göksu and 96% for Kızılırmak). The mean number of different alleles (Na) were 7.08 in the Göksu populations and 6.91 in the Kızılırmak populations. The Na of the Göksu and the Kızılırmak populations varied from 6.40 (the Ermenek population) to 7.75 (the downstream population- Silifke) and from 5.35 (the downstream population- Çorum) to 8.25 (the middle population- Kırıkkale (Table 3.5). The mean number of effective alleles was 3.69. It ranged from 3.50 to 3.85 in the Göksu

populations. On the other hand, Ne was 3.54 for the Kızılırmak populations and varied from 3.06 to 3.88. The number of highest private alleles were found in the upstream population of Göksu and the middle population- Kırıkkale of the Kızılırmak River; whereas, the lowest value was seen in the middle population- Mut and the downstream population- Çorum (Figure 3.3). A summary of private alleles by population was given in APPENDIX H.

In general, all studied populations had moderately genetic diversities. The mean observed heterozygosity value of the Göksu River populations was 0.65, and varied from 0.61 in the upstream population to 0.67 in the middle population- Mut. Likewise, the mean observed heterozygosity value was 0.64 for all the Kızılırmak River populations. It ranged from 0.62 (upstream and downstream populations) to 0.68 (middle population- Kırıkkale). The mean expected heterozygosity ranged from 0.59 to 0.63 with the average of 0.61 in the Göksu River populations and from 0.55 to 0.63 with the average of 0.60 in the Kızılırmak River populations. All studied populations showed significant departure from Hardy-Weinberg equilibrium frequencies. The excess of heterozygosities were observed across all the populations except for the downstream population- Bafra of the Kızılırmak River and the upstream population of the Göksu River. The mean value of inbreeding coefficient (Fis) was found to be low (-0.05, in the Göksu River populations and -0.07, in the Kızılırmak populations) in the studied populations (Table 3.5). The genetic differentiation (Fst) of all subpopulations consisting of the Göksu and Kızılırmak River was estimated as 0.07. To put it another way, 93% of genetic variation occurred within populations. It was found that very low level of differentiation exists among these populations. Regarding only the Göksu River populations, the Fst value was very low (0.02). The Kızılırmak River populations had moderate level of genetic differentiation(Fst=0.05).

Table 3.5. Estimated genetic diversity parameters for studied populations in the Göksu and the Kızılırmak River System

Population	N	Na±Se	Ne±Se	Pa	P (%)	G-W index (M)	Ho±Se	He±Se	F±Se	Fst	
GKSUPSTREAMPOP	31	7.75 ±1.12	3.85 ±0.58	15	100.00	0.32	0.61 ±0.07	0.59 ±0.06	0.00 ±0.07*		
GKSERMENEKPOP	23	6.40 ±0.81	3.50 ±0.45	8	100.00	0.31	0.66 ±0.07	0.60 ±0.05	-0.09 ±0.07*		
GKSMIDDLEPOP	28	7.00±0.94	3.67 ±0.43	6	100.00	0.34	0.67 ±0.06	0.63 ±0.05	-0.04 ±0.07*	0.02	
GKSDOWNSTREAMPOP	30	7.15 ±0.97	3.74 ±0.55	9	100.00	0.32	0.65 ±0.07	0.60 ±0.06	-0.08 ±0.07*		
Mean	28	7.08 ±0.96	3.69 ±0.50	-	100.00	0.32	0.65 ±0.07	0.61 ±0.06	-0.05 ±0.07*		
											0.07
KIZUPSTREAMPOP	33	7.85 ±1.10	3.88 ±0.55	11	95.00	0.33	0.62 ±0.06	0.61 ±0.06	-0.03 ±0.07*		
KIZMIDDLEPOPKIR	32	7.50 ±1.07	3.85 ±0.55	9	100.00	0.31	0.68 ±0.06	0.63 ±0.05	-0.08 ±0.06*		
KIZMIDDLEPOPKRK	52	8.25 ±1.09	3.32 ±0.43	19	95.00	0.30	0.65 ±0.07	0.58 ±0.06	-0.12 ±0.06*	0.05	
KIZDOWNSTREAMPOPCOR	15	5.35 ±0.77	3.06 ±0.41	1	90.00	0.34	0.62 ±0.07	0.55 ±0.06	-0.13 ±0.08*		
KIZDOWNSTREAMPOPBAF	15	5.60 ±0.60	3.59 ±0.42	6	100.00	0.32	0.62 ±0.07	0.63 ±0.05	0.03 ±0.08*		
Mean	29.40	6.91 ±0.93	3.54 ±0.47	-	96.00	0.32	0.64 ±0.07	0.60 ±0.06	-0.07 ±0.07*		
Total	28.70	6.98 ±0.32	3.61 ±0.16		97.78% ±1.21%	0.32	0.64 ±0.02	0.60 ±0.02	-0.06 ±0.02		

N= number of individuals, Na=mean number of different alleles, Ne=mean number of effective alleles, Pa=private alleles,%P= percentage of polymorphic loci**,** G-W index(M)= Garza-Williamson index, Ho=observed heterozygosity, He=expected heterozygosity, F=fixation index, Fst=the inbreeding coefficient within subpopulations. Values following the estimated parameters are standart error. *Significant at P<0.01.

The Garza-Williamson indices show whether a population experienced a genetic bottleneck or not. All the Garza-Williamson indices (Table 3.5) of the nine populations were lower than <0.68 which is the critical value. It was clear that there was a past reduction of effective population sizes of the species in the Göksu and Kızılırmak Rivers.

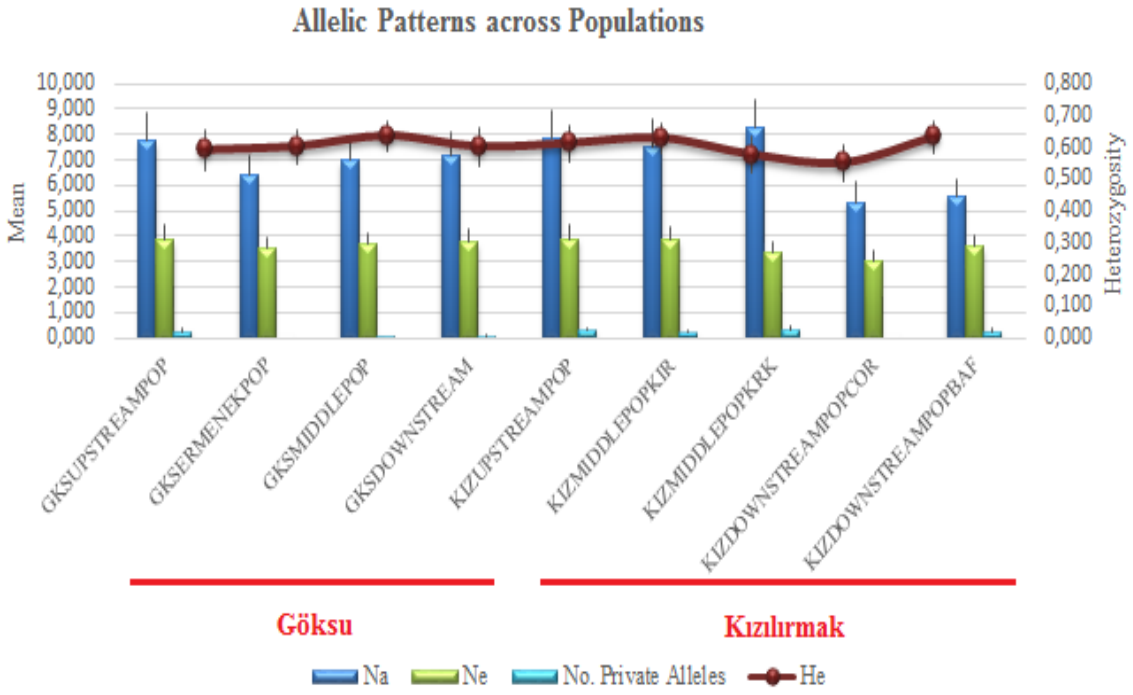
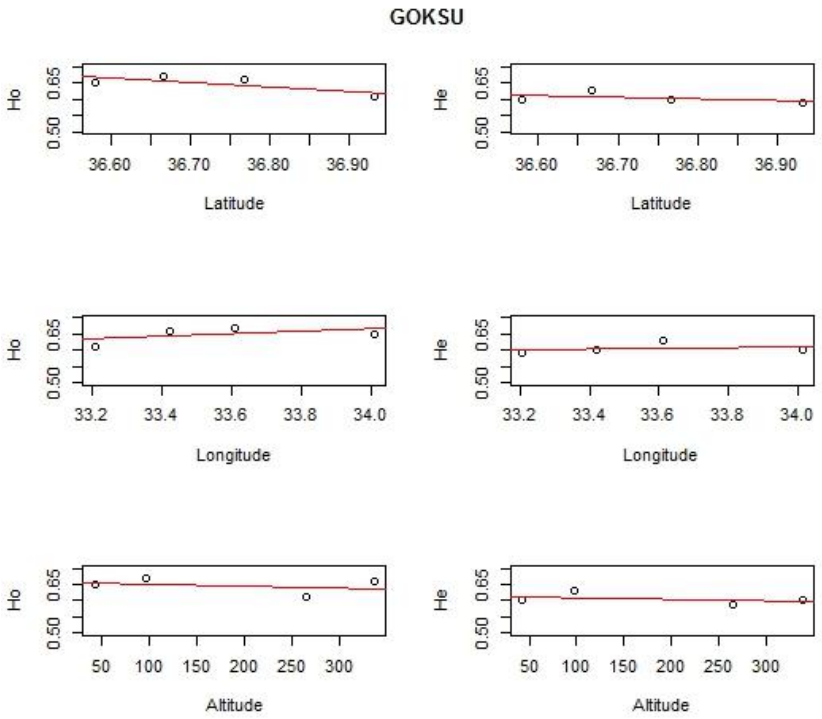


Figure 3.3 Allelic and Genetic Diversity Patterns Across populations of the Göksu and Kızılırmak Rivers

Effects of geographic factors such as longitude, latitude and altitude on genetic diversity of *S. alba* populations were represented in Figure 3.4. ANOVA analysis of all variables showed that variables had no significant effect on He or Ho in studied populations ($p>0.05$).

A)



B)

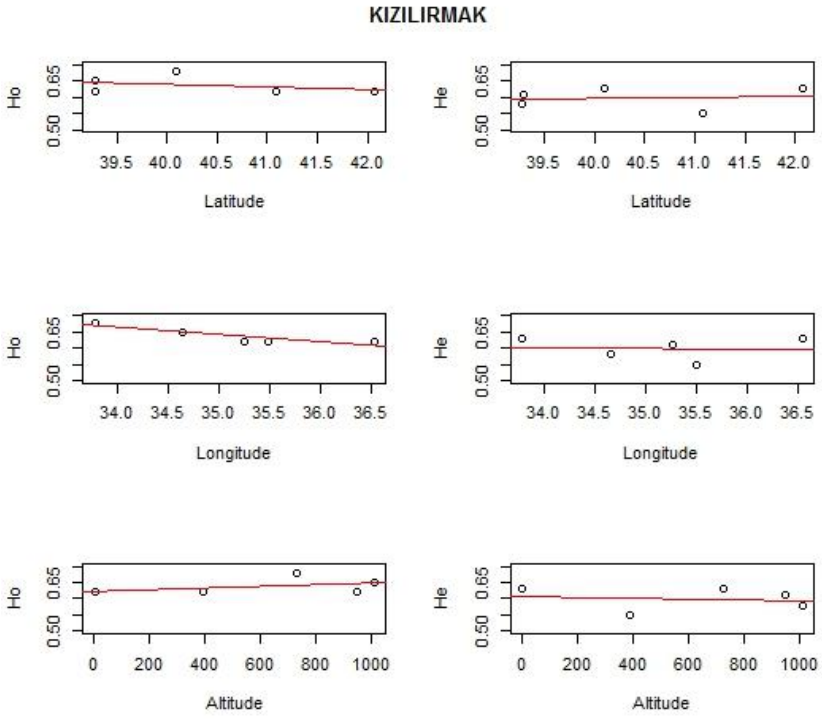


Figure 3.4. The scatter plots and best fit lines (Linear Regression Lines) represent Ho and He vs. latitude, longitude and altitude of studied population in the Göksu (A) and Kızılırmak Rivers (B).

3.4. Genetic Differentiation of Populations

Estimations of Pairwise Fst values and number of migrants (Nm), construction of genetic tree, determination of population structure were performed in order to get confirmation about nature of genetic substructuring of populations.

3.4.1. Pairwise Fst values, the UPGMA phenogram

Pairwise Fst values between populations varied from 0.010 (between middle populations- Kırıkkale and Kırşehir of Kızılırmak River) to 0.11 (between downstream populations- Çorum and Bafra of Kızılırmak River) (Table 3.6, Figure 3.5).

Table 3.6. Estimated pairwise Fst values (below diagonal) and number of migrants (Nm) (above diagonal) of *Salix alba* populations

POPULATIONS	1	2	3	4	5	6	7	8	9
GKSUPSTREAMPOP (1)	-	15.620	13.165	20.620	4.628	4.589	3.945	2.644	2.834
GKSERMENEKPOP (2)	0.016*	-	10.645	14.295	3.940	5.146	4.137	2.433	3.068
GKSMIDDLEPOP (3)	0.019*	0.023*	-	15.593	4.267	4.195	3.656	2.372	2.953
GKSDOWNSTREAMPOP (4)	0.012*	0.017*	0.016*	-	3.772	4.392	3.658	2.443	2.847
KIZUPSTREAMPOP (5)	0.051*	0.060*	0.055*	0.062*	-	11.530	9.337	3.244	3.298
KIZMIDDLEPOPKIR (6)	0.052*	0.046*	0.056*	0.054*	0.021*	-	24.384	3.823	4.941
KIZMIDDLEPOPKRK (7)	0.060*	0.057*	0.064*	0.064*	0.026*	0.010*	-	2.786	3.101
KIZDOWNSTREAMPOPCOR (8)	0.086*	0.093*	0.095*	0.093*	0.072*	0.061*	0.082*	-	2.082
KIZDOWNSTREAMPOPBAF(9)	0.081*	0.075*	0.078*	0.081*	0.070*	0.048*	0.075*	0.107*	-

* : $p < 0.05$

The Kızılırmak downstream populations- Çorum and Bafra were genetically most distant to other populations. The highest number of migrant (Nm) were found between upstream and downstream populations of Göksu as 20.620 and between KırÇehir and Kırıkkale subpopulations of Kızılırmak as 24.384. According to the constructed genetic tree of *Salix alba* populations; similarly the Kızılırmak downstream populations- Çorum and Bafra were genetically most distant to the other populations. One of two sub-clusters consisted of the Göksu River populations; whereas the remaining sub-cluster had the Kızılırmak upstream population, middle populations- KırÇehir and Kırıkkale (Figure 3.6).

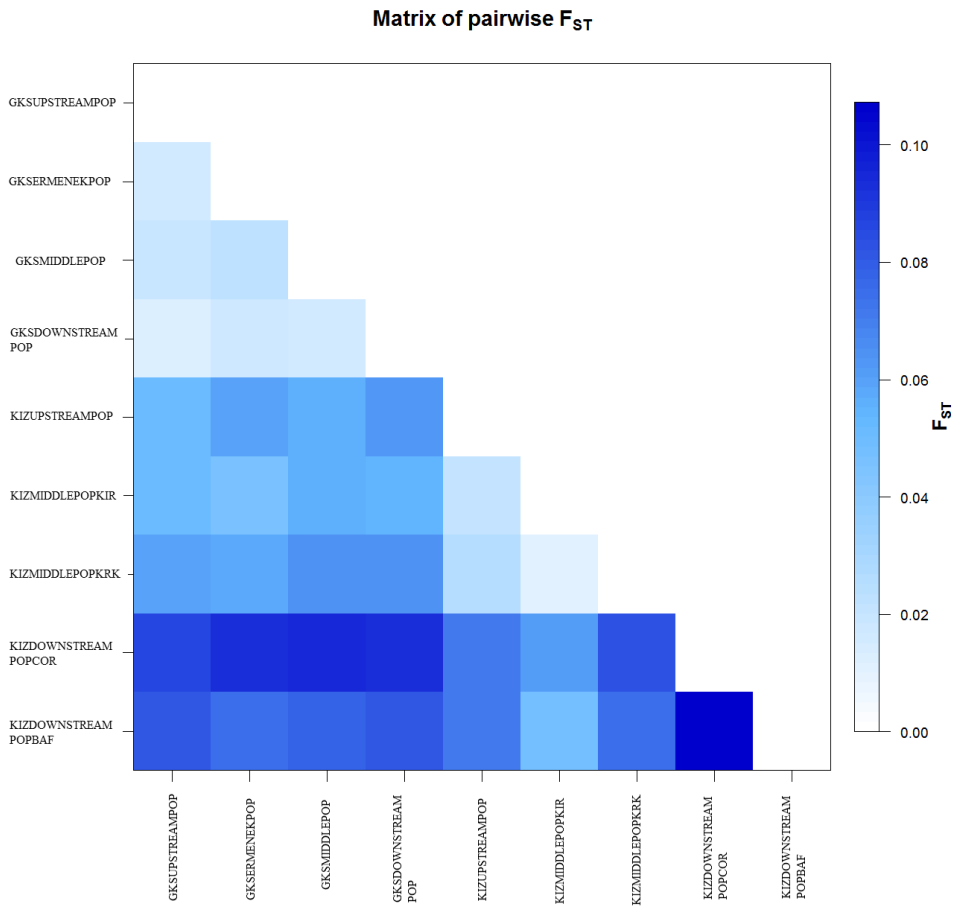


Figure 3.5. Diagrammatic representation of pariwise Fst values among populations

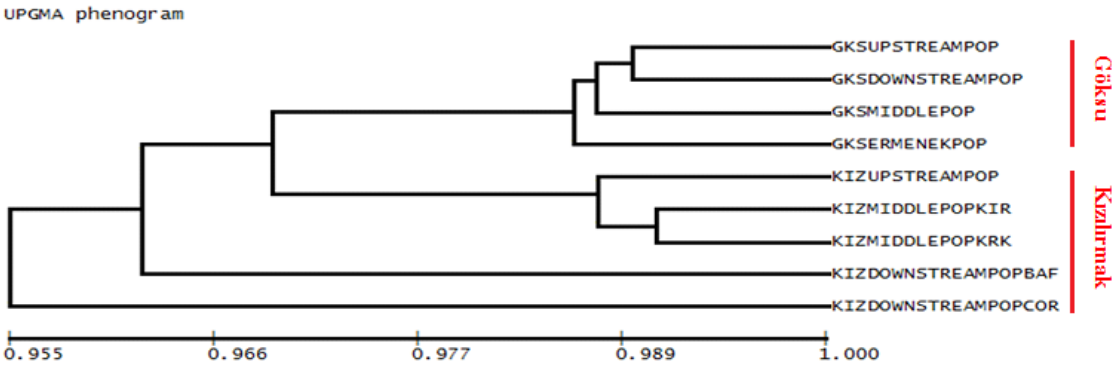


Figure 3.6. UPGMA phenogram based on coancestry identity of *Salix alba* populations from the Göksu and Kızılırmak Rivers.

3.4.2. The Population Genetic Structuring

Population structure analysis (STRUCTURE) was performed twice. Firstly, it was done without prior information on the geographic distribution of individuals. Secondly, the same analysis were carried out with prior information on the populations. There were no differences in the estimation of delta K values between the two methods. These analysis revealed that there are two groups, namely the Göksu and Kızılırmak rivers. The Göksu and Kızılırmak rivers were analyzed also separately based on sampled populations from each river. Pritchard *et al*. (2000) explained that "*We may not always be able to know the TRUE value of K, but we should aim for the smallest value of K that captures the major structure in the data*". The result of this study confirm this explanation. Finding of the study pointed out two main clusters based on the delta K value of 9 populations from the Göksu and the Kızılırmak Rivers. The graphical representation of delta K and Evanno method using delta K were shown in Figure 3.7 and Table 3.7, respectively. The first cluster consisted of all the Göksu populations except 3 genotypes (97.47 %) while, the all genotypes of the Kızılırmak Rivers placed in to the second cluster with a 100% proportion (Table 3.8). The detailed inferred membership values of the 259 *Salix alba* genotypes sampled from the Göksu and Kızılırmak River were given in Table 3.9.

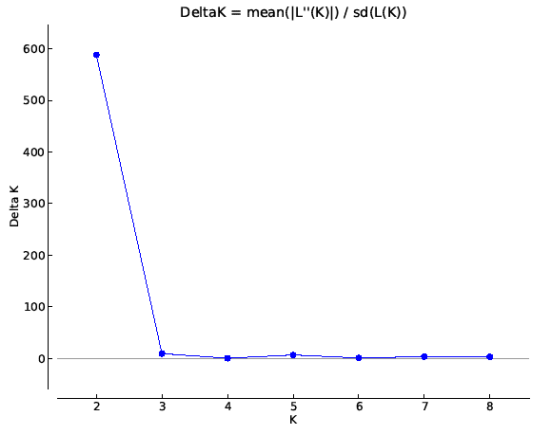


Figure 3.7. Graph of delta K values of 9 subpopulations from the Göksu and the Kızılırmak Rivers without prior information

Table 3.7. Evanno method using delta K

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	10	-16146.7	0.49	NA	NA	NA
2	**10**	**-15382.9**	**0.69**	**763.8**	**408.8**	**588.42**
3	10	-15027.9	20.79	355	187.53	9.02
4	10	-14860.4	311.44	167.47	80.54	0.26
5	10	-14612.4	71.28	248.01	459.43	6.45
6	10	-14823.8	707.47	-211.42	551.16	0.78
7	10	-14484.1	83.63	339.74	260.33	3.11
8	10	-14404.7	70.25	79.41	198.07	2.82
9	10	-14523.3	432.41	-118.66	NA	NA

Table 3.8. The number and proportions of 259 *Salix alba* genotypes assigned to two clusters (genetic groups)

Population	N	Cluster 1	Cluster 2
GKSUPSTREAMPOP	31	30 (96.77%)	1 (3.23%)
GKSERMENEKPOP	30	29 (96.67%)	1 (3.33%)
GKSMIDDLEPOP	23	23 (100%)	-
GKSDOWNSTREAM	28	27 (96.43%)	1 (3.57%)
KIZUPSTREAMPOP	33	-	33 (100%)
KIZMIDDLEPOPKIR	32	-	32 (100%)
KIZMIDDLEPOPKRK	52	-	52 (100%)
KIZDOWNSTREAMPOPCOR	15	-	15 (100%)
KIZDOWNSTREAMPOPBAF	15	-	15 (100%)

Table 3.9. The inferred membership values of the 259 *Salix alba* genotypes sampled from the Göksu and Kızılırmak Rivers

Population	N	Cluster 1			Cluster 2		
		0.8-0.9	0.7-0.8	<0.7	0.8-0.9	0.7-0.8	<0.7
GKSUPSTREAMPOP	31	30	-	-	-	-	1
GKSERMENEKPOP	30	25	2	2	-	-	1
GKSMIDDLEPOP	23	23	-	-	-	-	-
GKSDOWNSTREAM	28	27	-	-	1	-	-
KIZUPSTREAMPOP	33	-	-	-	31	1	1
KIZMIDDLEPOPKIR	32	-	-	-	31	1	-
KIZMIDDLEPOPKRK	52	-	-	-	50	2	-
KIZDOWNSTREAMPOPCOR	15	-	-	-	14	1	-
KIZDOWNSTREAMPOPBAF	15	-	-	-	14	1	-

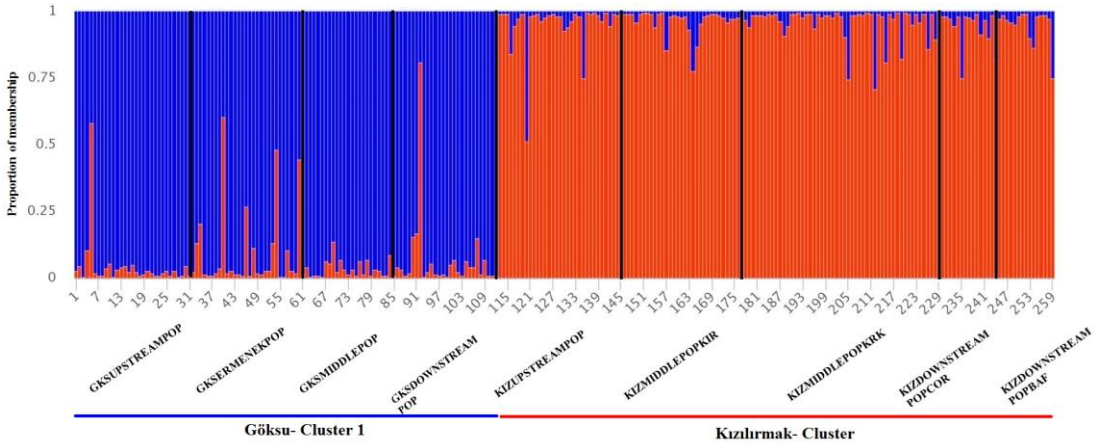


Figure 3.8. Graphical representation of the cluster membership of each 259 *S. alba* genotypes from 9 populations sampled from the Göksu and the Kızılırmak Rivers. Figure indicated the two inferred cluster and their corporation with populations. The different clusters were represented with different colors. Vertical lines separate the assumed subpopulations (Figure 3.8).

It was clear that there was a significant correlation between the genetic and geographic distances among the studied populations. The Göksu Rivers‟ populations were in the same cluster (cluster 1) whereas the Kızılırmak Rivers‟ populations were placed in another cluster (cluster 2).

One hundred and twelve genotypes from four different populations were sampled in the Göksu River basin. The results revealed that 57 of 112 sampled genotypes were

in the first cluster and 55 in the second cluster. Graphical representation of delta K and Evanno method using delta K were displayed in Figure 3.9 and Table 3.10, respectively. The highest estimated membership values were seen in the upstream population with 61.29% in the first cluster, and the middle population- Mut with 60.87% in the second cluster (Table 3.11). It was evident from these results that two clusters did not correlate with the geographic origins of the genotypes with regard to the proportion of membership (Figure 3.10). The detailed inferred membership values were presented in Table 3.12.

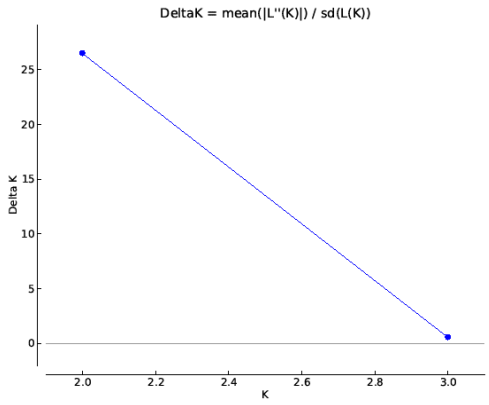


Figure 3.9. Graph of delta K values of 4 subpopulations from the Göksu River without prior information

Table3.10. Evanno method using delta K

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	10	-6568.50	0.69	NA	NA	NA
2	**10**	**-6546.30**	**15.53**	**22.20**	**412.14**	**26.53**
3	10	-6936.24	993.39	-389.94	572.51	0.58
4	10	-6753.67	31.56	182.57	NA	NA

Stdev: Standart Deviation

Table 3.11. The number and proportions of 112 *Salix alba* genotypes sampled from the Göksu River assigned to two genetic groups

Population	N	Cluster 1	Cluster 2
GKSUPSTREAMPOP	31	19 (61.29%)	12 (38.71%)
GKSERMENEKPOP	30	15 (50%)	15 (50%)
GKSMIDDLEPOP	23	9 (39.13%)	14(60.87%)
GKSDOWNSTREAM	28	14 (50%)	14 (50%)

Table 3.12. The inferred membership values of the 112 *Salix alba* genotypes sampled from the Göksu River

Population	N	Cluster 1			Cluster 2		
		0.8-0.9	0.7-0.8	<0.7	0.8-0.9	0.7-0.8	<0.7
GKSUPSTREAMPOP	31	14	4	1	4	2	6
GKSERMENEKPOP	30	10	1	4	10	-	5
GKSMIDDLEPOP	23	4	2	3	6	5	3
GKSDOWNSTREAM	28	10	1	3	7	4	3

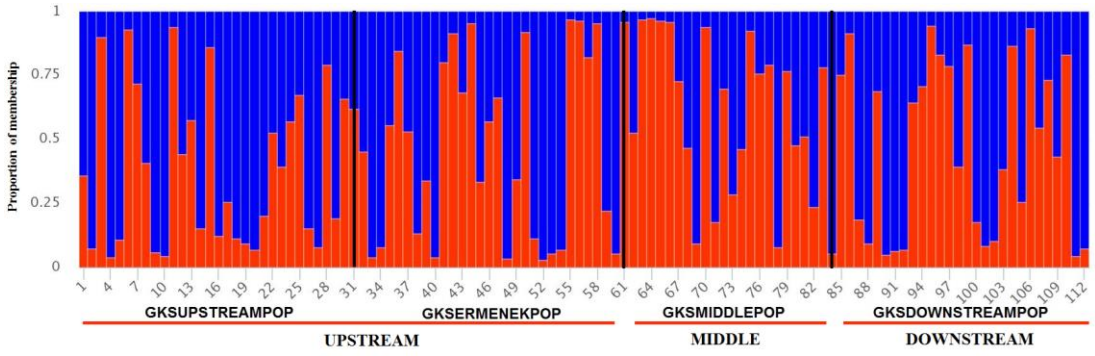


Figure 3.10. Graphical representation of the cluster membership of each 112 *Salix alba* genotypes from each of 4 populations sampled from the Göksu River.

One hundred and forty seven genotypes from five different populations were sampled around the Kızılırmak River. The maximum delta K was detected as K = 2 according to the Evanno‟s method (Figure 3.11, Table 3.13). The remaining 39 of 112 sampled genotypes formed the first cluster whereas 108 genotypes were in the second cluster. All genotypes of the downstream populations- Bafra and Çorum belonged to the second cluster (Table 3.14). Moreover, the upstream and the middle

population- Kırşehir existed in the second cluster with 87.88% and 78.12% membership values, respectively. The middle population- Kırıkkale had genotypes allocated to both the first and second clusters with 53.85% and 46.15% membership values (Table 3.14, Figure 3.12). The detailed inferred membership values were given in Table 3.15.

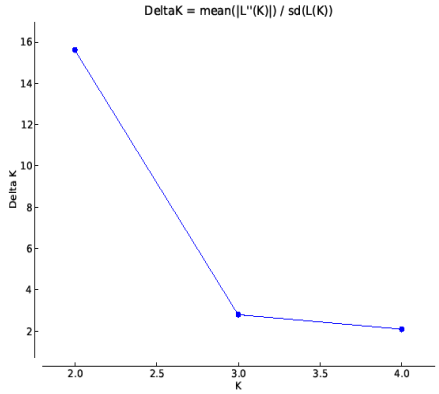


Figure 3.11. Graph of delta K values of 5 subpopulations from the Kızılırmak River without prior information

Table 3.13 Evanno method using delta K

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	10	-8748.79	0.86	NA	NA	NA
2	**10**	**-8390.82**	**6.57**	**357.97**	**102.59**	**15.62**
3	10	-8135.44	24.04	255.38	67.06	2.79
4	10	-7947.12	62.44	188.32	130.35	2.09
5	10	-7889.15	104.12	57.97	NA	NA

Table 3.14. The number and proportions of 147 *Salix alba* genotypes from the Kızılırmak River assigned to two genetic groups

Population	N	Cluster 1	Cluster 2
KIZUPSTREAMPOP	33	4(12.12%)	29 (87.88%)
KIZMIDDLEPOPKIR	32	7(21.88%)	25 (78.12%)
KIZMIDDLEPOPKRK	52	28(53.85%)	24(46.15%)
KIZDOWNSTREAMPOPCOR	15	-	15 (100%)
KIZDOWNSTREAMPOPBAF	15	-	15(100%)

Table 3.15. The inferred membership values of the 147 *Salix alba* genotypes sampled from the Kızılırmak River

Population	N	Cluster 1			Cluster 2		
		0.8-0.9	0.7-0.8	<0.7	0.8-0.9	0.7-0.8	<0.7
KIZUPSTREAMPOP	33	3	-	1	17	4	8
KIZMIDDLEPOPKIR	32	4	1	2	17	4	4
KIZMIDDLEPOPKRK	52	23	3	2	16	3	5
KIZDOWNSTREAMPOPCOR	15	-	-	-	15	-	-
KIZDOWNSTREAMPOPBAF	15	-	-	-	15	-	-

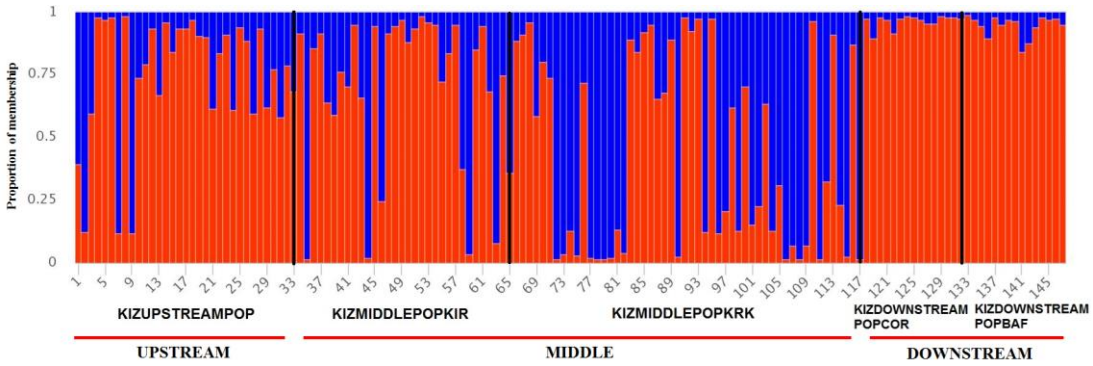


Figure 3.12. Graphical representation of the cluster membership of each 149 *Salix alba* genotypes within each 5 subpopulations sampled from the Kızılırmak River

3.4.3. Principal Coordinate Analysis (PCoA)

Principal Coordinatet Analysis (PCoA) was carried out using nine populations in order to get confirmation of population structure determined with STRUCTURE and Fst analysis. PCoA analysis displayed that 74% of the total variation was explained by the first tree axes with 30, 23 and 21%, respectively. The Göksu and Kızılırmak populations were separated by the first Principal Coordinate (30%). The downstream population- Çorum was an exception in the first and second quadrant, separated by both first and the second axes. The downstream population- Çorum seem to be the most distant from the other populations (Figure 3.13). Similar results were also obtained from structure and Fst analysis.

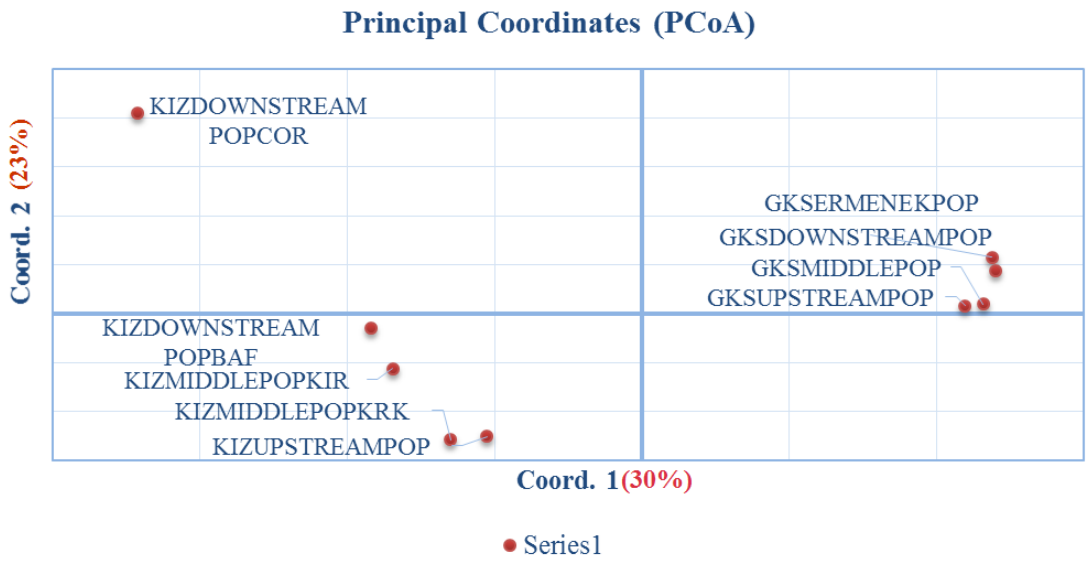


Figure 3.13. Principal Coordinate Analysis (PCoA) of the 249 genotypes from 9 populations based on the Nei's genetic distances.

3.5. Analyses of Molecular Variance (AMOVA)

Analyses of Molecular Variance revealed that there was a significant differentiation ($p<0.001$) among populations depending on Global locus-by-locus analysis of the 9 populations from two river systems (Table 3.16) showed results of the AMOVA, and the fixation indices. The fixation index values (Fst=0.07) were significant and consistent with the pairwise Fst results. The great portion of total variation (93.26%) was found within population in rivers. About small portion (3.47%) of the total variation was among population within river and 3.27% was among the rivers.

Table 3.16. Analyses of molecular variance (AMOVA) as weighted averages over loci in the rivers.

Source of variation	Sum of squares	Variance components	Percentage of total variation	Fixation Indices
Among rivers	74.69	0.22	3.27	Fct: 0.03*
Among populations within river	129.35	0.23	3.47	Fsc: 0.04*
Within populations	3093.18	6.14	93.26	Fst: 0.07*
Total	3297.22	6.58	100	-

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

The Göksu River populations were subdivided into two groups depending on both distance method and Bayesian analysis. The Ermenek population, tributary of the Göksu River, formed one group while the upstream, middle and downstream populations formed another group. When, AMOVA was carried out later, the "among groups" component total variance was found to be 0.86% (no significant differentiation). The highest variance was found among genotypes within populations. It was important that the Fst value found AMOVA results (Fst=0.02) was in full agreement with the pairwise Fst value (Table 3.17) as expected.

Table 3.17. Analyses of molecular variance as weighted averages over loci in the Göksu populations

Source of variation	Sum of squares	Variance components	Percentage of total variation	Fixation Indices
Among groups	15.21	0.05	0.86	Fct: 0.008*
Among populations within groups	24.56	0.10	1.64	Fsc: 0.02*
Within populations	1353.19	6.19	97.50	Fst: 0.02
Total	1392.96	6.34	100	-

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

Based on both distance method and Bayesian analysis, similarly two groups were detected in the Kızılırmak River system. One group was composed of the upstream,

and middle populations. The other group included the downstream populations-Çorum and Bafra. AMOVA results produced reduced variance components due to among trees within population (93.94%), among population (3.44%), and among groups (2.82%). The population differentiation (Fst) was compatible with the pairwise Fst value estimation from ARLEQUIN (Table 3.18).

Table 3.18. Analyses of molecular variance (AMOVA) results as weighted averages over loci based on the Kızılırmak subpopulations

Source of variation	Sum of squares	Variance components	Percentage of total variation	Fixation Indices
Among groups	31.71	0.18	2.82	Fct: 0.03*
Among populations within groups	57.87	0.22	3.44	Fsc: 0.04*
Within populations	1739.99	6.08	93.94	Fst: 0.06*
Total	1829.57	6.49	100	-

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

CHAPTER 4

DISCUSSION

The first population genetic study of the white willow (*Salix alba* L.), one of the most valuable tree species in river ecosystems, were undertaken to examine genetic structure of this species in Göksu and Kızılırmak Rivers which are two different major river systems in Turkey. By using microsatellite data obtained in this study, the ecological and evolutionary factors that affect the genetic structure of *S. alba* were put together. This type of information will be invaluable to establish conservation and breeding strategies of genetic resources of the species.

4.1. Data Quality

Microsatellite markers are coded with prefixes SB-, gS- and W- were developed for willow, and WPMC- and PMGC- for poplar. Primers previously amplified in *S. alba* included SB24, SB80, SB194, SB196, SB243 (Barker *et al.,* 2003; Singh *et al.,* 2013a; Singh *et al.,* 2013b; Singh *et al.,* 2014). It was found that 15 SSR loci out of 20 were new to be used in *S. alba* in this study. Remaining loci were selected from similar studies which were related other *Salix* species by considering polymorphism levels. Besides, they were selected from each chromosome when linkage map was available. Unlike the *Salix* SSR primers with tree repetitive motifs, SSR markers derived from poplar species had a dinucleotide repeat motifs except for three markers (SB196, SB243 and WPMS18).

Previously reported that some error caused by null alleles, stuttering, large allele dropout and false allele size due to misreading of electropherogram may occur while scoring microsatellite data. In this study, no null allele was detected in all studied subpopulations. Only some loci showed evidences of null alleles that are specific to a certain population. The poor primer annealing due to nucleotide sequence divergence such as point mutations or indels in one or both flanking primers, differential amplification of size-variant alleles, and PCR failure due to inconsistent DNA template quality or low template quantity are the potential causes of null alleles (Dakin and Avise, 2004). Furthermore, the observed null alleles (non-amplifying alleles) may be due to geographically restricted mutations in primer binding sites. Therefore, the amplifications can be prevented. Since the identification of null allele depended on deviations from the HWE, it is possible to add incorrect null alleles by nonrandom mating caused events such as inbreeding (Chakraborty *et. al*., 1992). Most of the statistical tools used to detect potential scoring errors work through testing of heterozygous deficiencies, which may be indicative of genotyping errors. Especially, when only a few loci are detected.

Studied SSR markers which were developed from different *Salix* species were used since there was no species-specific marker for *S. alba*. It was observed that 20 primers were polymorphic in studied populations. Allele size ranges of primers identified in the current study are compatible with those reported in the previous studies. Nevertheless, allele number of some studied SSR markers are different from that of the literature (Lian *et al.,* 2003; Stamati *et al.,* 2003; Lin *et al.,* 2009; Puschenreiter *et al.,* 2010; Sochor *et al.,* 2013; Singh *et al.,* 2013a). The reasons for this can be listed as follows; the distribution of species, climatic and habitat environments requirements, and the genetic diversity exposed to environmental and anthropogenic factors. In the future, it can potentially be clarified using species-specific loci.

Linkage disequilibrium (LD) occurs as a result of differentiation among populations isolation by distance, asexual reproduction, linkage between the alleles, selection and genetic drift (Agapow and Burt, 2001). No significant LD was detected between polymorphic loci. SB24, SB80, SB194, SB243 and W784 loci are found on chromosomes XI, VIII, VI, XIII, and D, respectively based on a genetic linkage map

of *S. viminalis* (Hanley *et al.,* 2002). PMGC2709, PMGC2889, PMGC2163 and WPMS18 microsatellite markers developed from *Populus* species are located on the chromosomes II, XVII, X and I, respectively (Gaudet *et al.,* 2008). There is no information in the literature about location of remaining 11 primers on the chromosome of *Salix* species. Linkage among loci is not expected if populations are sexual while alleles transforms freely into new genotypes in the sexual reproduction process. To date, no genetic map of *S. alba* has been created. For this reason, it is a point that should be considered in future studies.

While *Salix* species rarely develop roots sucker, they reproduce easily with vegetative materials that subjected to downstream deposition (Rood *et al.,* 2003; Kuzovkina *et al.,* 2008). Vegetative reproduction is common in *Populus* species belonging to the same family (Castiglione *et al.,* 2010). *S. alba* is colonized mainly by a sexual reproduction. This is supported that multilocus genotyping showed no clonal duplication among sampled genotypes in two river system. This finding is not surprising since similar results were reported in other *Salix* species in *S. arctica* (Steltzer *et al.,* 2008); different willow species (Douhovnikoff *et al.,* 2010), *S. huakoana* (Kikuchi *et al.,* 2011).

4.2. Genetic Diversity of Loci

The genetic diversity consists of two elements; the number of alleles and the abundance (or evenness) of the alleles. Both will cause to increase the expected heterozygosity. The mean number of different alleles per locus (Na) are not equally relevant to the average number of effective alleles (Ne) in present study. Ne is the number of equal-frequency alleles that are required to give the same He, as in the study, because of the fact that allele frequencies are not equally contributing. In this case, low-frequency alleles have little effect on the number of effective alleles. This will cause to Ne to be lower than Na.

Allelic richness (Ar), Polymorphic Information Content (PIC), Probability of identity (PI), Shannon Index (I), and expected heterozygosity (He) were preferred for the selection of markers in *S. alba.* Ar is estimated from SSR dataset, regardless of the

number of samples. In the current study, 16 of the 20 SSR loci exhibited high Ar (> 0.3). Only locus SB196 was found to be low informative (<0.25). Thirteen of 19 SSR primers had informative markers with high PIC values (> 0.5) with low PI values smaller than 0.20. Twelve of the 20 most diverse loci are also characterized by I index. It is thought that Shannon‟s information index (I) may be a better measure of allelic and genetic diversity since, it is not bounded by 1 unlike He (Sherwin *et al.,* 2006). The additional parameters are in agreement with the expected heterozygosity. All the measures were considered together for determining the discriminating efficiency of a primer. It is recommended that the Sare03, Sare04, Sare08, SB24, SB233, SB243, W293, gSIMCT024, gSIMCT052, PMGC2709, PMGC2889 and PMGC2163 loci can be used effectively in the genetic studies of *S. alba*. Some loci had higher allelic parameter values (Ne, PIC, He). This is in contrast to the results of other relevant studies (King *et al.,* 2010; Singh *et al.,* 2013a, b; 2014; Rungis *et al.,* 2017) due to different sampling locations and varying number of samples of the current study and other studies. Besides, *S. alba* populations from different geographic regions may have different evolutionary history.

The expected heterozygosity, also addressed as genetic diversity, is the fundamental criterion for assessing genetic variability (Nei, 1972). The excess of heterozygosity was generally observed for 9 of 20 loci. Five of them has been showing statistically significant deviation from Hardy-Weinberg equilibrium due to excess of some heterozygotes in the studied loci and presence of some rare single homozygotes. Excess of heterozygosity may be result of selection and dispersal of trees with heterozygotes genotypes by humans. Rest of the loci had positive fixation index value. Similarly five of them show statistically significant deviation from Hardy-Weinberg equilibrium due to inbreeding (nonrandom mating) and the population substructuring caused by loss and fragmentation of natural habitats of the species. Urbanization, erosion and flood control works, including the construction of embankments, dams, beaches and borders have caused the destruction of old riparian habitats of *S.alba* species in two river systems. Negative impact of humans on the natural populations of white willow caused genetic erosion in genetic heritage of the species.

F statistics are very useful for getting information about amount of inbreeding and population differentiation. In this study nine of 20 loci showed excess of heterozygosity. This shows that loci in each population is affected independently of changes in environmental factors. The highest Fst values were found for loci SB24, SB80, SB265, SB493, and PMGC2709 with the lowest number of migrants (Nm) values that mostly contributed to the differentiation of populations.

4.3. Genetic Diversity of Populations

Genetic diversity is an important element in dynamics of populations, because it is directly relevant to evolutionary potential of populations (Hughes *et al.,* 2008). Genetic diversity is generally assessed by presenting estimated allelic richness (private alleles, different allele number), polymorphism content, and expected heterozygosity.

Some authors have suggested that private alleles are important for conservation. For example, Funk *et al.,* (2007) described subspecies for conservation and contained the criteria of "unique alleles or haplotypes". Analysis of private alleles showed high genetic diversity in the sampled populations. The private allele occurs at any frequency but unique to only one population that interbreeds freely without restriction on gene flow, and so should have a distinct gene pool relative to the rest of the population. The number of private alleles in the studied population was highly variable. Highest number of private alleles were found in the upstream population- Göksu River (15) and the middle population- Kırıkkale (19) from the Kızılırmak River. This means that these populations may have a higher effective population size than the other populations, since they had a slightly higher mean observed heterozygosity. These populations need to be considered for conservation and new breeding program for future generations due to having the highest number of private alleles. Interestingly, only one private allele was identified in the downstream population- Çorum.

The pattern of genetic diversity is also reflected in the percentage of polymorphic loci within a population. All studied populations were highly polymorphic.

Percentage of polymorphic loci were calculated as 100% for all populations except upstream population (95%), middle population- Kırıkkale (95%) and downstream population- Çorum (90%) in Kızılırmak River.

Natural selection, genetic drift and inbreeding reduce the genetic variation within population while mutation and gene flow increase the genetic diversity within population. Moderate levels of genetic diversity in both Göksu and Kızılırmak Rivers populations were observed (Mean Ho of Göksu populations = 0.65; mean Ho of Kızılırmak populations = 0.64). All studied populations significantly deviated from Hardy-Weinberg equilibrium. Observed heterozygosity was slightly higher than expected heterozygosity in all populations except downstream population- Bafra in Kızılırmak River. This could be due to negative assortative mating, selection (heterozygote advantage), or the proportion of heterozygote individuals in founder population. These populations were characterized by negative values of the Fixation Index (Fis) due to an excess of heterozygotes with respect to HW equilibrium, though at different level. Negative Fis value is expected in the Salicaceae family members which are dioecious and highly heterozygous. Besides, wind pollination, seed dispersal via animals and vegetative material movements promote the variation within populations. A slightly excess number of homozygotes in downstream population- Bafra from Kızılırmak River could indicate inbreeding and (positive) assortative mating. Our result is consistent with the previous studies carried out in *S. eriocephala* (Lin *et al.,* 2009; Lauron *et al.,* 2013); *S. caprea* (Perdereau *et al.,* 2014) and *Salix* spp. (Ukwubile *et al.,* 2014), which all populations in these studies exhibited high values of observed heterozygosity. Besides, the majority of *S. hukaoana* and *S. daphnoides* subpopulations exhibited negative Fis values (Kikuchi *et al.,* 2011; Sochor *et al.,* 2013). There are also studies of different willow species that are in contrast with our results. The reason for positive Fis values is that different SSR primers are used in different species, sample size and studied populations and their locations. Downstream accumulation of genetic diversity has been reported for various river plants (Lundqvist and Andersson 2001; Liu *et al.,* 2006). However, this trend is controversial, since seed dispersal types and effective populations‟ size are very effective on it. There was no significant increase in genetic diversity of downstream populations from both Göksu and Kızılırmak Rivers.

Similar results found in riparian willow *S. hukaoana* study that was conducted by Kikuchi *et al.* (2011).

In all studied populations from two different rivers (Göksu and Kızılırmak), the Garza-Williamson indices were lower than M<0.68 which is the critical value. Obtained values by simulations based on empirical data indicated that there was a reduction in population size caused by genetic bottleneck. Due to the reduction and fragmentation of habitats, recent demographic bottlenecks have reduced the population size of Göksu and Kızılırmak river populations. The trend towards negative Fis values may be attributed to the heterozygosity excess due to populations" experience of past genetic bottleneck (Kikuchi *et al.,* 2011).

4.4. Genetic Differentiation of Populations

Differentiation among populations is a major issue in evolution. Life story characteristics such as pollen and seed dispersal and stage of succession in the ecosystem may affect the variation within and between populations. Genetic drift increases the differentiation whereas gene flow that occurs between populations through pollen, seed or fruit dispersal reduces the differentiation between populations. Mutations have limited effect on differentiation. Furthermore, natural selection is regarded as weak.

All members of the family Salicaceae generally display low Fst value between the populations when compared with other forest trees. Small, light seeds are readily dispersed by wind and water and can be transported efficiently to remote locations. Additionally, vegetative propagation in different locations from shoots can become naturally via broken branches floating in river (Hörandl *et al.,* 2002). Genetic drift is expected to play a minor role for wind- pollinated trees with large and continuous distribution. In these trees, gene flow can be desirable level, which balancing the allele frequencies between populations. Climax species in the ecosystem under consideration are more affected by this phenomenon (Gösta *et al.,* 2013).

Moderate level of differentiation (Fst = 0.07)mexisted between Göksu and Kızılırmak Rivers populations. Fst value among Göksu and Kızılırmak Rivers populations were found to be 0.02 (little genetic differentiation) and 0.05 (moderate genetic differentiation), respectively. This is consistent with the studies carried out in different willow species (Puschenreiter *et al.,* 2010; Trybush *et al.,* 2012; Sochor *et al.,* 2013; Berlin *et al.,* 2014; Perdereau *et al.,* 2014; Ukwubile *et al.,* 2014). To assess the effect of two rivers on genetic variance, analysis of molecular variance (AMOVA) was also performed. Genetic differentiation was tested among river systems and found to be significant (Fst value =0.07). The majority of variation was seen among populations within rivers. This was caused by the higher gene flow between individuals or populations. Although long-distance gene flow is not common in willow species due to short viability of seed, human mediated vegetative materials movements lead to genetic similarities between populations at distant locations. In Anatolia, for thousands of years, people in rural areas commonly moved and scattered willow and poplar vegetative materials to different area of the country for later use of building material, animal feed, and fuel wood.

Various factors may limit or increase gene flow levels in plant species. Pairwise Fst values showed clearly that the populations of two rivers were highly differentiated. The highest pairwise Fst values with the lowest number of migrants were obtained from the Kızılırmak downstream populations- Çorum and Bafra. This was mainly caused by constructed barriers (Boyabat, Altınkaya and Derbent Dams) which fragmented populations and reduced gene flow. The Göksu River populations consisted of one sub-clusters whereas, the remaining sub-cluster had the Kızılırmak upstream population, middle population- Kırşehir and Kırıkkale. The Kızılırmak downstream populations- Çorum and Bafra were genetically most distant to the remaining populations. The reason may be that Black Sea Mountains located between the downstream populations- Çorum, Bafra and the remaining populations. This geographical isolation (mountain barrier) limits the gene flow (the number of migrants) between different populations. In addition, the Kızılırmak downstream population- Çorum was differed from other populations due to two monomorphic marker (SB196, WPMS18) of 20 SSR loci. This situation was also confirmed by Principal coordinate analysis (PCoA). Due to the discovery and visualization of both the similarities and dissimilarities of the data, the Kızılırmak downstream population-

Çorum was separated by both first and the second axes. This region is located in mountains, so the topographic structure of study area is different from other studied regions. Furthermore, little anthropogenic effect is seen on the region. The highest number of migrant values (Nm) were found between upstream and downstream populations of Göksu River and between middle populations- Kırşehir and Kırıkkale of Kızılırmak River.

Genetic structure studies are necessary to verify genetic isolation of subpopulations any species from different locations. Genetic differences are often correlated with geographic distance between populations. Two hundred and fifty-nine *S. alba* genotypes were placed to two gene pools by genetic STRUCTURE analysis. The first cluster was generally composed of trees sampled from Göksu River populations whereas the second cluster was composed of Kızılırmak Rivers populations with high membership values. It is clearly seen that white willow populations in two different river basin have two different founder populations. Moderate level of genetic differentiation was found (Fst=0.07) between these two founder populations. It seems that geographical distance has little effect on genetic differentiation of *S. alba* populations from Kızılırmak and Göksu rivers. Geographical barriers such as mountains (Middle Anatolian and Taurus Mountains) and a plain (Ereğli Plain) caused moderate level of differentiation of populations of two river systems by preventing extensive gene flow.

The Göksu and Kızılırmak rivers were analyzed also separately by considering sampled populations of each river. In spite of geographic barriers, four population of Göksu River were genetically close to each other. The Göksu River populations have two gene pools with a high admixture due to high level of gene flow among populations. It can be quite clearly said that two groups did not correlate with the geographic origin based on the proportion of membership. Altitude differences and Gezende Dam may cause slight differentiation of Ermenek population from other populations in Göksu River. After defining number of cluster based on both distance method and Bayesian analysis, this little genetic differentiation of Ermenek population was checked via AMOVA by considering effect of altitude. Ermenek population which tributary of the Göksu River belonged to first group and the upstream, middle and downstream populations of Göksu River belonged to second

group. AMOVA result showed that differentiation between two groups was found to be low and not significant (Fst = 0.02; P = 0.000). Therefore, there is no obvious barriers for gene flow between populations. Seed dispersal in river systems occurs via hydrochory (dispersion by water) or anemochory (dispersion by wind) (Imbert and Lefèvre, 2003). It was noted that the secondary hydrochory (dispersing seeds with water after initial dispersed by wind) facilitates long-distance seed distribution in willow (Seiwa *et al.,* 2008). Hydrochory may have contributed to the increased long-distance gene flow in *S. alba.* However, *Salix* seeds stay viable only for a few weeks (Maroder *et al.,* 2000). Generally, seeds of *Salix* are vulnerable to long-distance dispersal. Combination of hydrochory and anemochory are the possible mechanisms for long distance gene flow in *S. alba*. However, in the current study, human mediated vegetative material dispersal and asexual propagation appear to be more important for long distance gene flow.

Five different populations of the Kızılırmak River were clustered into two genetic groups. Bafra and Çorum populations were found to be most distant to others among all populations and possess a homogeneous gene pool (Fst= 0.06*). The highest admixture was observed in the middle population- Kırıkkale large areas for germination and seedling are available. Besides, human-mediated dispersal, anthropogenic changes in habitats such as pollution, use of natural resources, and climate change are also common in Kırıkkale.

It is worth mentioning that all the results from the current study are important for future studies towards shedding light on the population structure of S. *alba* in Turkey. Obtained information from this study will provide baseline information in new breeding and improvement programs. However, in further studies, sexual and asexual regeneration strategies in *S. alba* should be investigated taking account the natural patterns of river flow and rates of sediment delivery along rivers.

CHAPTER 5

CONCLUSION

River ecosystems should be protected because of their genetic variability and biodiversity in terms of plant and animal species. *Salix* species is known as pioneer species of riparian ecosystems. Restoration with willow species will provide benefits such as pollution reduction, mitigation, biomass accumulation and biofuel production, CO_2 storage, micro climate management, biodiversity shelter and contributing to the improvement of landscape management in the urban environment. Economic importance of *S. alba* is that white willow wood is used in a number of commercial areas such bioenergy, phytoremediation and construction. Therefore, it is important to use and conserve genetic resource of native white willow species.

Since *Salix excelsa* was used an a synonym of *Salix alba* and was not well differentiated from *S. alba*, this study examines genetic diversity structure of *Salix alba* populations (259 genotypes) from two major river system in Turkey by using 20 microsatellite markers. According to obtained results only some loci showed evidences of null alleles that are specific to certain populations. No significant linkage disequilibrium and clonal duplication were detected in populations. It is recommended that the Sare03, Sare04, Sare08, SB24, SB233, SB243, W293, gSIMCT024, gSIMCT052, PMGC2709, PMGC2889 and PMGC2163 loci can be used effectively in the future genetic studies dealing *S. alba* since these loci were highly informative with respect to Allelic richness (Ar), Polymorphic Information Content (PIC), Probability of identity (PI), Shannon Index (I), and expected heterozygosity (He) parameters.

The upstream population of Göksu and the middle population-Kırıkkale of Kızılırmak need to be considered for conservation and new breeding program for future generations due to having the highest private alleles number. Moderate levels of genetic diversity in both Göksu and Kızılırmak River populations were observed. Wind pollination and seed dispersal via wind, water and animals and vegetative material movements by human appear to promote the high level of gene flow among populations. Furthermore, all studied populations from Göksu and Kızılımak Rivers experienced some level of genetic bottlenecks which cause an increase inbreeding and loss of genetic variation. Recent demographic bottlenecks may have also responsible for reduction of the population sizes of Göksu and Kızılırmak river populations.

Due to high level of gene flow, moderate level of differentiation existed between Göksu and Kızılırmak River populations (Fst=0.07). Göksu River populations showed little genetic differentiation (Fst=0.02) while Kızılırmak River populations contain moderate genetic differentiation (Fst=0.05) which indicates that the majority of variation was found among populations within rivers. Cluster and PCoA analysis showed clearly that the populations from two river systems were highly differentiated. The highest pairwise Fst values with the lowest number of migrants were obtained from the Kızılırmak downstream populations. Therefore, these populations were genetically most distant to the remaining populations. This was mainly caused by constructed barriers (dams) and mountain barrier (geographical isolation) limiting the gene flow (the number of migrants) among populations.

The results of the study demonstrate that white willow populations in two different river systems have two different founder populations with very high membership values originating different genetic resources. Despite geographic barriers, populations within Göksu River as well as populations within Kızılırmak rivers were genetically close to each other. This may be due to human mediated vegetative material dispersal and asexual propagation.

This study provide important insights for efficient conservation, management, utilization and breeding of genetic resources of economically valuable *S. alba*. The

white willow is seen as ecologically and economically important tree in Turkey. It is pioneer tree species of riparian ecosystems and regarded as a potential tree for industrial plantations. By considering importance of the species, the new conservation and breeding programs should be implemented. In order to conserve existing riparian ecosystems of willow species, effective legislation on conservation issue should be established and a proper management policy should be put into action for these natural reserves. This requires a comprehensive cooperation among government institutions, local people and scientists at national levels.

REFERENCES

Acar, P. Molecular Phylogenetic of Turkish *Salix* L. Species (PhD), Thesis, unpublished). Middle East Technical University, 2017.

Agapow, P. and Burt, A. Indices of multilocus linkage disequilibrium. Molecular Ecology Notes, 1(1-2), 101-102, 2001.

Allendorf, F.W. Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biol* 5: 181–190, 1986.

Allendorf, F.W., Luikart, GH. Conservation and the genetics of populations. Chapters 6, 9. Blackwell Publishing, Malden, Oxford, 642 pp, 2007.

Allendorf, FW., Luikart, G.H., Aitken, S.N Conservation and the genetics of populations. West Sussex: Wiley-Blackwell, 2012.

Argus, G.W. An experimental study of hybridization and pollination in *Salix* (willow). *Can. J. Bot.* 52: 1613-1619, 1974.

Argus, G.W. Infrageneric Classification of *Salix* (Salicaceae) in the New World, *Systematic Botany Monographs,* 52, 1-121, 1997.

Argus, G.W. The genus *Salix* (Salicaceae) in the southeastern United States. *Systematic Botany Monographs.* 9: 1–170, 1986.

Arıhan, O. and Güvenç, A. Ankara çevresinde yetiĞen söğüt (*Salix* L.) türleri, *Ot Sistematik Botanik Dergisi*, 15-52, 2009.

Arıhan, O. and Güvenç, A. Studies on the anatomical structure of stems of willow (*Salix* L.) species (Salicaceae) growing in Ankara province, *Turkish Journal of Botany,* 35, 535-551, 2011.

Arnaud-Haond, S., Belkhir, K. Genclone: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes*. 7:15–17, 2007.

Arsenov, D., Zupunski, M., Borisev, M., Nikolic, N., Orlovic, S., Pilipovic, A., Pajevic, S. Exogenously Applied Citric Acid Enhances Antioxidant Defense and Phytoextraction of Cadmium by Willows (*Salix* spp.). *Water Air Soil Pollut*, 228: 221, 2017. doi: 10.1007/s11270-017-3405-6.

Avcı, M. Türkiye'nin Doğal Söğütleri ve Coğrafi Dağılışları, *İstanbul Üniversitesi Edebiyat Fakültesi Coğrafya Bölümü Coğrafya Dergisi*, pp.1-24, 1, 1999.

Ball, J., Carle, J., and Lungo, A. D. Contribution of poplars and willows to sustainable. FAO, Unasylva 56, 3-9, 2005.

Barcaccia, G., Meneghetti, S., Albertini, E., Triest, L. and Luchhin, M. Linkage mapping in tetraploid willows: segregation of molecular markers and estimation of linkage phases support an allotetraploid structure for *Salix alba* × *Salix fragilis* interspecific hybrids. *Heredity* 90:169–180, 2013.

Barcaccia, G., Meneghetti, S., Lucchin, M.and de Jon, H. Genetic Segregation and Genomic Hybridization Patterns Support an Allotetraploid Structure and Disomic Inheritance for *Salix* Species. *Diversity* 6(4) 633-651, 2014. doi:10.3390/d6040633.

Barker J . H. A., Pahlıch A., Trybush, Edwards S. K . J . and Karp A. Microsatellite markers for diverse *Salix* species. *Molecular Ecology Notes* 3, 4–6, 2003. doi: 10.1046/j.1471-8286 .2003.00332.x.

Barret, S.C. H., Yakimowski, S. B., Field, D.L. and Pickup, M. Ecological genetics of sex ratios in plant populations. *Phil. Trans. R. Soc. B* 365, 2549–2557, 2010doi:10.1098/rstb.2010.0002

Barsoum, N. Relative contributions of sexual and asexual regeneration strategies in *Populus nigra* and *Salix alba* during the first years of establishment on a braided gravel bed river. *Evolutionary Ecology* 15(4-6): 255-279, 2001.

Berlin, S.,Trybush, S. O., Fogelqvist, J., Gyllenstrand, N., Hallingbäck, H. R., Åhman, I., Nordh, N-E., Shield, I., Powers, S. J., Weih, M., Lagercrantz, U., Rönnberg-Wästljung, A-C., Karp, A., Hanle, S. J. Genetic diversity, population structure and phenotypic variation in European *Salix viminalis* L. (Salicaceae). *Tree Genetics & Genomes* 10:1595–1610, 2014. Doi: 10.1007/s11295-014-0782-5.

Bosco, C., de Rigo, D., Dewitte, O., Poesen, J. and Panagos. Modelling soil erosion at European scale: towards harmonization and reproducibility. *Natural Hazards and Earth System Sciences,*15, 225-245, 2015. doi: 10.5194/nhess-15-225-2015.

Bosi, G. Kluwer Academic Publishers. Printed in the Netherlands. Abundance, diversity and seasonal succession of dytiscid and noterid beetles (Coleoptera: Adephaga) in two marshes of the Eastern Po Plain (Italy). *Hydrobiologia* 459: 1–7, 2001.

Botstein, D., White, R.L., Skolnick, M., Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*. 32:314–331, 1980.

Bozzi, J., Liepelt, A., Ohneiser, S.S., Gallo, L.A., Marchelli, P., Leyer, I., Ziegenhagen, B., and Mengel, C. Characterization Of 23 Polymorphic SSR Markers in *Salix humboldtiana* (Salicaceae) using Next-Generation Sequencing

and Cross-Amplification From Related Species. *Applications Plant Sciences,* 3 (4): apps. 1400120, 2015. doi:10.3732/apps.1400120.

Brookfield, J.F.Y. A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology,* 5:453–455, 1996.

Brown, A.H.D., Feldman, M.W., Nevo, E. Multilocus structure of natural populations of Hordeum spontaneum. *Genetics*, 96: 523– 536, 1980.

Caballero, A., Garcı´a-Dorado, A. Allelic diversity and its implications for the rate of adaptation. *Genetics,* 195: 1373–1384, 2013.

Camerini, G. Factors affecting Lymantria dispar mortality in a willow wood in northern Italy. *Bulletin of Insectology,* 62 (1): 21-25, 2009.

Castiglione, S., Cicatelli A., Lupı, R. , Patrıgnanı, G. , Fossatı ,T. , Brundu, G. , Sabattı, M., Van Loo, M., Lexer, C.Genetic structure and introgression in riparian populations of *Populus alba* L. *Plant Biosyst,* 144(3): 656–668, 2010.

Cavalli-Sforza, L.L. The DNA revolution in population genetics. *Trends in Genetic*, 14(2): 60-651, 1998. https://doi.org/10.1016/S0168-9525(97)01327-9

Chakraborty, R., De Andrade M., Daiger S.P., Budowle, B. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Ann Hum Genet,* 56: 45–57, 1992.

Chase, M.W., Zmarzty, S., Lledo, M.D., Wurdack, K.J., Swensen, S.M. and Fay, M.F. When in doubt, put it in the Flacourtiaceae: a molecular phylogenetic analysis based on plastid rbcL DNA sequences. *Kew Bulletin* 57, 141–181, 2002.

Ciftci, A., Karatay, H., Kücükosmanoğlu, F., Karahan, A., Kaya. Z. Genetic differentiation between clone collections and natural populations of European black poplar (*Populus nigra* L.) in Turkey. *Tree Genetics & Genomes*, 13:69, 2017. Doi:10.1007/s11295-017-1154-8.

Council of the European Union, 8/9 March 2007, note 7224/1/07.

Cronk, Q.C., Needham I., and Rudall, P. J. Evolution of catkins: inflorescence morphology of selected Salicaceae in an evolutionary and developmental context. *Front. Plant Sci.* 6, 1030, 2015.

Cronquist, A., *An intergrated system of classification of flowering plants.* Columbia University Press, New York, 1981.

Dakin, E. E., Avise, J. C. Microsatellite null alleles in parentage analysis. *Heredity*, 93(5), 504-509, 2004.

Davis, P.H.R., *Flora of Turkey and the East Aegean Islands*, Univ. Press, Edinburgh, pp. 694-716, 1965-1988.

Delph, L.F. *Sexual Dimorphism in Life History*. In: Geber M.A., Dawson T.E., Delph L.F. (eds) Gender and Sexual Dimorphism in Flowering Plants. Springer, Berlin, Heidelberg, 1999. Doi: https://doi.org/10.1007/978-3-662-03908-3_6.

Dobes, C., and Vitek, E. Documented chromosome number checklist of Austrian vascular plants. *Museum of Natural History*, Vienna, Austria. 2000.

Donner, J., *Distribution maps to P.H. Davis 'Flora of Turkey 1-10' Linzer Biologische Betrage* 2: pp.381-515, 1990.

Douhovnikofj, V., Goldsmith, G. R., Tape, K.D., Huang, C., Sur, N. and Bret-Harte, M. S. Clonal Diversity in an Expanding Community of Arctic *Salix* spp. and a Model for Recruitment Modes of Arctic Plants. *Arctic, Antarctic, and Alpine Research,* Vol. 42, No. 4, , pp. 406-411, 2010.

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

Druskovic, B. IOPB chromosome data 9. *Int Org Plant Biosyst Newsl*. 24:11–14. 1995

Dubiec, A., Góüdü, I., and Mazgajski, T.D. Green plant material in avian nests, *Avian Biology Research* 6 (2), 133–146, 2013.

Earl, D.A, Bridgett, V. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conser Genet Resour* 4(2): 359–361, 2012.

Endler, J. A. *Natural Selection in the Wild*, Princeton: Princeton University Press. 1986

Enerji Atlası, http://www.enerjiatlasi.com/akarsular/goksu-nehri.html and http://www.enerjiatlasi.com/akarsular/kizilirmak.html. Accessed August 2017.

Esbrí, J.M., Cacovean, H., Higueras, P. Usage Proposal of a common urban decorative tree (*Salix alba* L.) to monitor the dispersion of gaseous mercury: A case study from Turda (Romania). *Chemosphere*, 2017. doi: 10.1016/j.chemosphere.2017.11.007.

Evanno, G., Regnaut, S., Goudet, J. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol*. 14: 2611–2620, 2005.

Excoffier, L., Lischer, H.E.L. 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, 2010. doi: 10.1111/j.1755-0998.2010.02847.x.

Faccoli, M., Favaro, R., Smith, M. T., Wu, J. Life history of the Asian longhorn beetle *Anoplophora glabripennis* (*Coleoptera cerambycidae*) in southern Europe. *Agricultural and Forest Entomology*, volume 17, Issue 2, Pages 188–196, 2015.

Falush, D., Stephens, M., Pritchard, J.K. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 7(4):574–578, 2007.

Falush, D., Wirth, T., Linz, B., Pritchard, J.K., Stephens, M., Kidd, M., Blaser, M.J., Graham, D.Y., Vacher, S., Perez-Perez, G.I., Yamaoka, Y., Mégraud, F., Otto, K., Reichard, U., Katzowitsch, E., Wang, X., Achtman, M., Suerbaum, S. Traces of human migrations in *Helicobacter pylori* populations. *Science,* 7;299(5612):1582-5, 2003.

Fang, Z. On the distribution and origin of *Salix* in the world, Wiley-Balchwell press, Institue of Botany, Botanical Society of China: *Journal of Systematic and Evolution,* 25(4): 307-313, 1987.

Fischer, G., Prieler, S., van Velthuizen, H. Biomass potentials of miscanthus, willow and poplar: results and policy implications for Eastern Europe, Northern and Central Asia. *Biomass Bioenerg,* 28:119e32, 2005.

Foulley, J., Ollivier, L. Estimating allelic richness and its diversity. *Livestock Sci,* 101: 150–158, 2006.

Francis, R.M. POPHELPER: an R package and web app to analyse and visualise population structure. *Mol Ecol Resour.* 17(1):27-32, 2017. doi: 10.1111/1755-0998.12509.

Funk, W.C., Mullins, T.D., Haig, S.M. Conservation genetics of snowy plovers (*Charadrius alexandrinus*) in the Western Hemisphere: population genetic structure and delineation of subspecies. *Conserv Genet*. 8:1287–1309, 2007. doi: 10.1007/s10592-006-9278-7.

Futuyma, D.J. *Evolution.* Third Edition, 2013.

Garza, J., Williamson, E. Detection of reduction in population size using data from microsatellite loci. *Mol Ecol*. 10:305–318, 2001. doi:10.1046/j.1365-294x.2001.01190.x.

Gaudet, J. FSTAT version 1.2: a computer program to calculate F-statistics. *J Hered.* 86:485–486, 1995.

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

Gemayel, R., Cho, J., Boeynaems, S. and Verstrepen, K.J. Beyond Junk-Variable Tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. *Genes,* 3:461-480, 2012.

Giachetti, G. and Sebastiani, L. Effects of tannery waste on growth dynamics and metal uptake in *Salix alba* L. *Plant Biosystems*, Vol. 141, No. 1, pp. 22 – 30, 2007.

Gösta, E., Inger, E., and David, C. *Genetics Applied to Forestry An Introduction*. Third edition. ISBN 978-91-576-9187-3, 2013.

Greenbaum, G., Templeton, A., Zarmi, Y., Bar-David, S. Allelic Richness following Population Founding Events – A Stochastic Modeling Framework Incorporating Gene Flow and Genetic Drift. *PLOS ONE*, 2014. Doi:10.1371/journal.pone.0115203.

Güner, A. and Zielinski, J., *Salix rizeensis* (Salicaceae): a new willow from N. E., Turkey, *Karaca Arboretum* 1: 1-5, 1993.

Güner, A., *Salix,* L. In: Güner, A., Özhatay, N., Ekim, T. and Başer, K.H.C. (eds.), *Flora of Turkey and the East Aegean Islands*, Vol.11, Edinburgh: Edinburgh University Press, 2000.

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

Gupta, A. Singh, N.B., Choudhary, P. J, S harma and Sankhayan, H. P. Estimation of Genetic Variability, Heritability and Genetic Gain for Wood Density and Fibre Length in 36 Clones of White Willow (*Salix alba* L.). *International Journal of Agriculture, Environment & Biotechnology*. 7(2): 299-304, 2014. Doi: 10.5958/2230-732X.2014.00247.2.

Håkansson, A. Chromosome numbers and meiosis in certain *Salices*. *Hereditas* 41: 454- 482, 1955. doi: 10.1111/j.1601-5223.1955.tb03004.x.

Han, S.H., Kim, D.H., Lee, J.C. Cadmium and zinc interaction and phytoremediation potential of seven *Salix capreaclones*. *J Ecol Field Biol*, 33:245–251, 2010.

Hanley, S., Barker, J. H. A., Van Ooijen, Aldam J. W. C., Harris, S. L. , Åhman. I., Larsson, S., Karp, A. A genetic linkage map of willow (*Salix viminalis*) based on AFLP and microsatellite markers. *Theor Appl Genet*, 105:1087–1096, 2002. Doi:10.1007/s00122-002-0979-0.

Hanley, S.J. and Karp, A. Genetic strategies for dissecting complex traits in biomass willows (*Salix* spp.). *Tree Physiol,* 34(11):1167-80, 2014. doi:10.1093/treephys/tpt089.

Hanley, S.J., Mallot, M.D., Karp, A. Alignment of a *Salix* linkage map to the *Populus* genomic sequence reveals macrosynteny between willow and poplar genomes. *Tree Genetics and Genomes*, 3, 35–48, 2006.

Hardig, T.M., Anttila, C.K., and Brunsfeld, S.J. A Phylogenetic Analysis of *Salix* (Salicaceae) Based on matK and Ribosomal DNA Sequence Data, *Journal of Botany,* 1-12, 2010.

Hartl, D. L., Clark, A. G., *Principles of Population Genetics*, Fourth Edition, Sinauer Associates, Inc., 23 Plumtree Road, Sunderland, MA 01375, USA, 2007.

Hartl, D.L., Clark, A.G. *Principles of population* genetics. Sinauer Associates, Inc. Publishers. Mass (USA): Sunderland, 1997.

Hearne, C. M., Ghosh, S. & Todd, J. A. Microsatellites for linkage analysis of genetic-traits. *Trends Genet.* 8, 288–294, 1992.

Heijden, E. W. van der and Kuyper, Th. W. Ecological Strategies of Ectomycorrhizal Fungi of *Salix repens*: Root Manipulation versus Root Replacement. *Oikos,* Vol. 103, No. 3, pp. 668-680, 2003.

Hinchee, M.A.W., Mullinax, L.N., Rottmann, W. Wood biomassand purpose-grown trees as teedstocks for renewable energy. In:Mascia PN, Scheffran J, Thomas SR, Widholm JM (eds) Biotechnology in agriculture and forestry, vol 66., Plantbiotechnology for sustainable production of energy and co-products. Springer, Berlin, pp 155–208, 2010.

Holm, S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6:65-70, 1979.

Hörandl, E., Florineth, F., Hadacek, F. Weiden in Österreich und angrenzenden Gebieten. Universität für Bodenkultur, Wien, 2002.

Hoshikawa, T., Kikuchi, S., Nagamıtsu, T., and Tomaru, N. Eighteen microsatellite loci in *Salix arbutifolia* (Salicaceae) and cross-species amplification in *Salix* and *Populus* species. *Molecular Ecology Resources*, Volume 9, Issue 4, Pages 1202–1205, 2009. doi: 10.1111/j.1755-0998.2009.02607.x.

Houston, D.T., Rigo, D., Caudullo, G. *Salix alba* in Europe: distribution, habitat, usage and threats. In: "European Atlas of Forest Tree Species" (San-Miguel-Ayanz J, Rigo D, Caudullo G, Houston Durrant T, Mauri A eds), 2016.

Hubisz, J.M., Falush, D., Stephens, M., Pritchard, J.K. Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour,* 9(5):1322–1332, 2009.

Hughes, R.A., Inouye, B.D., Johnson, M.T.J., Underwood, N., Vellend, M. Ecological consequences of genetic diversity. *Ecol Lett,* 11: 609–623, 2008.

Imbert, E. and Lefèvre, F. Dispersal and gene flow of *Populus nigra* (Saliacaceae) along a dynamic river system Journal of Ecology 91 447–56. 2003.

Isebrands, J. G., Richardson, J. Poplars and willows: trees for society and the environment. CABI; FAO, 2014.

IUCN Red List of Threatened Species, http://www.iucnredlist.org/. Accessed August 2017.

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

Johnson, O. and More, D. *Collins tree guide*. Collins, 2006.

Jonah, P., Bello, L., Lucky, O., Midau, A. and Moruppa, S. Review: The importance of molecular markers in plant breeding programmes. *Glob J Sci Front Res,* Volume 11 Issue 5 Version 1.0, 2011.

Kacalkova, L., Tlustos, P., Szakova, J. Phytoextraction of risk elements by willow and poplar trees. *Int. J. Phytoremediat*. 17, 414–421, 2015.

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

Kalinowski, S.T., Taper, M.L., Marshall, T.C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol.* 16:1099–1106, 2007. doi:10.1111/j.1365-294X.2007.03089.x.

Kamvar, Z.N., Tabima, J.F., Grünwald, N.J. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281, 2014. https://doi.org/10.7717/peerj.281.

Kantarcı, D., *Türkiye Bitkileri Listesi*, Nezahat Gökyiğit Botanik Bahçesi, 2012.

Karp, A., Richter, G.M., Shield, I.F., Hanley, S.H. Genetics, genomics and crop modelling: integrative approaches to the improvement of biomass willows. In: Carpita N, Buckeridge MS, McCann MC (eds) *Plants and bioenergy*. Chapter 7. Springer, New York, 2013.

Kavak, S. 2017. *Salix excelsa*. (amended version published in 2014) The IUCN Red List of Threatened Species 2017: e.T19620474A112568440.

Kaya, Z. and Raynal, D.J., Biodiversity and conservation of Turkish forests. *Biol. Conserv.* 97(2): 131–141, 2001.

Khalilia, Z., Mirzaie-Nodoushanb, H., Ghahremaninejada, F. and Maassoumib, A. A. Chromosome number and meiosis behavior in several *Salix* species of Iran. Caryologia: International Journal of Cytology, Cytosystematics and Cytogenetics Vol. 65, No. 4, 258–262, 2012.

Khela, S. *Salix alba*. The IUCN Red List of Threatened Species 2013: e.T203465A2765960.

Kikuchi, S., Suzuki, W. , Sashimura, N. Gene flow in an endangered willow *Salix hukaoana* (Salicaceae) in natural and fragmented riparian landscapes. *Conserv Genet,* 12:79–89, 2011. doi: 10.1007/s10592-009-9992-z.

Kikuchi, S., Suzuki, W., Ban, N., Kanazashi, A., Yoshimaru, H. Characterization of eight polymorphic microsatellites in endangered willow *Salix hukaoana. Molecular Ecology Notes*, 5, 869–870, 2005.

King, R. A., Harris, S.L., Karp, A., Barker, J. H. A Characterisation and inheritance of nuclear microsatellite loci for use in population studies of the allotetraploid

Salix alba–Salix fragilis complex. *Tree Genetics & Genomes*, 6:247–258, 2010. doi: 10.1007/s11295-009-0245-6.

Klug, W. S., Cummings, M. R., Spencer, C. A., *Concepts of Genetics*, Eight Edition, Pearson Education Inc., Pearson Prentice Hall, Upper Saddle River, NJ 07458, 2006.

Koutsos, T.M., Dimopoulos, G.C., Mamolos, A. P. Spatial evaluation model for assessing and mapping impacts on threatened species in regions adjacent to Natura 2000 sites due to dam construction. *Ecological Engineering,* 36: 1017–1027, 2010.

Kuzovkina, Y. A. and Quigley, M.F. Willows Beyond Wetlands: Uses Of *Salix* L. Species For Environmental Projects. *Water, Air, and Soil Pollution,* 162: 183–204, 2005.

Kuzovkina, Y. A. and Vietto, L. An update on the cultivar registration of *Populus* and *Salix* (Salicaceae). *Skvortsovia*: 1(2): 133 – 148, 2014.

Kuzovkina, Y.A., Weih, M., Romero, M.A., Charles, J., Hurst, S., McIvor, I. *Salix*: botany and global horticulture. *Horticultural Reviews*, 34, 447–489, 2008.

Lauron-Moreau, A., Pitre, F. E., Brouillet, L. and Labrecque, M. Microsatellite Markers of Willow Species and Characterization of 11 Polymorphic Microsatellites for *Salix eriocephala* (Salicaceae), a Potential Native Species for Biomass Production in Canada. *Plants,* 2, 203-210, 2013. doi:10.3390/plants2020203.

Lei, Y., Jiang Y.i, Chen, K, Duan, B., Zhang, S., Korpelainen, H., Niinemets, Ü. and Li, C., Epron, D. Reproductive investments driven by sex and altitude in sympatric *Populus* and *Salix* trees. *Tree Physiol.* 28:1-12, 2017. doi:10.1093/treephys/tpx075.

Lewis, P. O., and Zaykin, D. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html, 2001.

Lian, C., Nara, K., Nakaya, H., Zhou, Z., Wu, B., Miyashita, N., Hogetsu, T. Development of microsatellite markers in polyploid *Salix reinii*. *Mol Ecol Notes,* 1: 160-161, 2001.

Lian, C.L., Oishi, R., Miyashita, N., Nara, K., Nakaya, H., Wu, B.Y., Zhou, Z.H., Hogetsu, T. Genetic structure and reproduction dynamics of *Salix reinii* during primary succession on Mount Fuji, as revealed by nuclear and chloroplast microsatellite analysis. *Mol Ecol*, 12:609–618, 2003.

Lin, J., Gibbs, J.P., Smart, L.B. Population genetic structure of native versus naturalized sympatric shrub willows (*Salix*; Salicaceae). *Am J Bot*, 96:771–785, 2009.

Liu, S. *Salix*. In *Flora of China, Cycadaceae through Fagaceae*; Fang, Z.F., Zhao, S.D., Skvortsov, A.K., Eds.; Science Press Beijing: St. Louis, MO, USA, pp. 139–279, 1999.

Liu, Y., Wang, Y., Huang, H. High interpopulation genetic differentiation and unidirectional linear migration patterns in *Myricaria laxiflora* (Tamaricaceae), an endemic riparian plant in the Three Gorges Valley of the Yangtze River. *Am J Bot,* 93:206–215, 2006.

Logan, T. J. „Reclamation of chemically degraded soils", *Adv. Soil Sci.* 17, 13–35. 1992.

Luna, R.K., Plantation trees. Publishing Corporation. IBD, Dehradun, pp 640-645, 1995.

Lundqvist, E., Andersson, E. Genetic diversity in populations of plants with different breeding and dispersal strategies in a freeflowing boreal river system. *Hereditas,* 135:75–83, 2001.

Madan, R.K., Levitt, J. A review of toxicity from topical salicylic acid preparations. *J Am Acad Dermatol,* 70: 788-792, 2014.

Maddison, W.P., Knowles, LL. "Inferring Phylogeny Despite Incomplete Lineage Sorting", *Syst. Biol.* 55(1), 21–30, 2006.

Major, J.E. Mosseler, A., Malcolm, J.W., Heartz, S. Salinity tolerance of three *Salix* species: Survival, biomass yield and allocation, and biochemical efficiencies. *Biomass and Bioenergy,* Volume 105, Pages 10-22, 2017.

Manel, S., Gaggiotti, O.E., Waples, R.S. Assignment methods: matching biological questions with appropriate techniques. *TRENDS in Ecology and Evolution* Vol. 20 No.3, 2005.

Maroder, H. L., Prego, I. A., Facc Iuto, G. R. And Maldonado, S. B.). Storage Behaviour of *Salix alba* and *Salix matsudana* Seeds. Annals of Botany 86: 1017-1021, 2000. doi:10.1006/anbo.2000.1265.

Marshall, T.C., Slate, J., Kruuk, L., Pemberton, J.M. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol.* 7:639–655, 1998. doi: 10.1046/j.1365-294x.1998.00374.x.

Mason, A.S. SSR Genotyping. In: Batley J (ed) Plant Genotyping. *Springer*, New York, NY, pp 77-89, 2015.

Migeon, A., Richaud, P., Guinet, F., Chalot, M., Blaudez, D. Metal Accumulation by Woody Species on Contaminated Sites in the North of France. *Water Air Soil Pollut* 204:89–10, 2009. doi: 10.1007/s11270-009-0029-5.

Miner, J., Hoffhines, A. The discovery of aspirin's antithrombotic effects. *Tex Heart Inst J*, 34: 179-186, 2007.

Mitchell, A. F., Dahlstrom, P., Sunesen, Darter, E. A field guide to the trees of Britain and northern Europe. Collins, 1974.

Mleczek, M., Rutkowski, P., Rissmann, I., Kaczmarek, Z., Golinski, P., Szentner, K., Strazynska, K., Stachowiak A. Biomass productivity and phytoremediation potential of *Salix alba* and *Salix viminalis*. Elsevier. *Biomass and bioenergy* 34: (9), 1410-1418, 2010. doi:10.1016/j.biombioe.2010.04.012.

Monedero, E., Hernández, J. J.and Collado, R. Combustion-Related Properties of Poplar, Willow and Black Locust to be used as Fuels in Power Plants. *Energies*, Vol.10 No.7 pp.997 ref.44, 2017. doi:10.3390/en10070997.

Nei, M. 1972. Genetic distance between populations. *American Naturalist*, 106, 283-392.

Nei, M. 1977. F-statistics and analysis of gene diversity in subdivided populations. *Annals of Human Genetics,* 41, 225-233

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

Nei, M., *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY, USA, 1987.

Newman, D., Pilson, D. Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. *Evolution* (N Y) 51: 354–362, 1997.

Oliveira, E.J., Pádua, J.G., Zucchi, M.I., Vencovsky, R. and Vieira, M.L.C. Origin, evolution and genome distribution of microsatellites. *Genet Mol Biol,* 29:294-307, 2006.

Oosterhout, C., Hutchinson, W. F., Wills, D.P.M., Shipley, P. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538, 2004.

Orman Atlası, T.C. Orman Ve Su Ġleri Bakanlığı, Orman Genel Müdürlüğü, 2013.

Pajević, S., Bororišev, M., Nikolić, N., Krstić, B., Pilipović, A., and Ororlović, S. Prophytoremediation Capacity Of Poplar (*Populus* Spp.) and Willow (*Salix* Spp.) Clones In Relation To Photosynthesis. Arch. Biol. Sci., Belgrade, 61 (2), 239-247, 2009. doi:10.2298/ABS0901239P.

Palmé, A.E., Semerikov, V., Lascoux, M. (2003) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity*, 91:465–474.

Paray, P.A., Bhat, S. J. A., Bhat, A. and Parrey, A.A., Candidate Plus Tree Selection of *Salix alba* – A Multipurpose Tree Species, *Int. J. Pure App. Biosci.* 5(2): 387-395, 2017. doi: http://dx.doi.org/10.18782/2320-7051.2711.

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

Perdereau, A. C., Kelleher, C. T., Douglas, G. C., Hodkinson, T. R. High levels of gene flow and genetic diversity in Irish populations of *Salix caprea* L. inferred from chloroplast and nuclear SSR markers. *BMC Plant Biol.* 1:202, 2014. 10.1186/s12870-014-0202-x.

Petit, R. J., A. El Mousadik, and O. Pons. Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, 12:844-855, 1998.

Phumichai, C., Phumichai, T. and Wongkaew, A. Novel chloroplast microsatellite (cpSSR) markers for genetic diversity assessment of cultivated and wild *Hevea rubber*. *Plant Mol Biol Report* 33:1486-1498, 2015.

Pohjonen, V. Energy forests with *Salix* as a Carbon Dioxide sink, 2008.

Pourrut, B., Lopareva-Pohu, A., Pruvot, C., Garçon, G., Verdin, A., Waterlot, C., Bidar, G., Shirali, P., Douay, F. Assessment of fly ash-aided phytostabilisation of highly contaminated soils after an 8-year field trial. Part 2. Influence on plants. *Science of the Total Environment,* 409, 4504–4510, 2011.

Praciak, A, Pasiecznik, N, Sheil, D, van Heist, M, Sassen, M, Correia, CS, Dixon, C, Fyson, G, Rushford, K., Teeling, C (eds). The CABI encyclopedia of forest trees, CABI, Oxfordshire, 2013, UK. ISBN: 9781780642369.

Preston, C.D., Pearman, D.A., Dines, T.D. New atlas of the British and Irish flora. Oxford University Press, Oxford, 2002.

Pritchard, J.K., Stephens, M., Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* 155(2):945-59, 2000.

Puschenreiter, M., TürktaǴ,M., Sommer, P., Wieshammer, G., Laaha, G., Wenzel, W.W., Hauser, M.T. Differentiation of metallicolous and non-metallicolous *Salix caprea* populations based on phenotypic characteristics and nuclear microsatellite (SSR) markers. Published in final edited form as: Plant Cell Environ.; 33(10): 1641–1655, 2010. doi:10.1111/j.1365-3040.2010.02170.x.

Raskin, I. Salicylate, A new plant hormone. *Plant Physiol* 99: 799-803, 1992.

Raven, J. A. „The physiology of *Salix*", in R. Watling and J. A. Raven (eds), *Willow Symposium. Proceedings of The Royal Society of Edinburgh*, vol. 98, The Royal Society of Edinburgh, Edinburgh, pp. 49–62, 1992.

Raymond, M., Rousset, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Heredity*, 86:248-249, 1995.

Reisch, C., Schurm, S. And Poschlod, P. Spatial Genetic Structure and Clonal Diversity in an Alpine Population of *Salix herbacea* (Salicaceae). Annals of Botany 99: 647–651, 2007. doi:10.1093/aob/mcl290.

Rigo, D. de.Scientific Topics Focus 2, mri10a15+, 2016.

Rood, S.R., Kalischuk, A.R., Polzin, M.L., Braatne, J.H. Branch propagation, not cladoptosis, permits dispersive, clonal reproduction of riparian cottonwoods. *For Ecol Manag*, 186:227–242, 2003.

Rostunov, A., Konchina, T., Zhestkova, E., Gusev, D., Kharitono, S. The Dependence of Morphological and Physiological Indicators of the Leaves of Woody Plants on the Degree of Technogenic Pollution. *Environment. Technology*. Resources, Rezekne, Latvia Proceedings of the 11th International Scientific and Practical Conference. Volume I, 235-239, 2017.

Rousset, F. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol. Ecol. Resources 8: 103-106, 2008.

Rungıs, D., Bérubé, Y., Zhang, J., Ralph, S., Rıtland, C. E., Ellıs, B. E., Douglas, C., Bohlmann, J., Rıtland, K. Robust simple sequence repeat markers for spruce (*Picea* spp.) from expressed sequence tags. *Theoretical and Applied Genetics,* 109, 1283-1294, 2004.

Rungis, D., Laivins, M., Gailite, A., Korica, A., Lazdina, D., Skipars, V., Veinberga, I. Genetic analysis of Latvian *Salix alba* L. and hybrid populations using nuclear and chloroplast DNA markers. *iForest* 10: 422-429, 2017. doi: 10.3832/ifor2004-009.

Saini, B.C., Sharma P., *Salix* - a multipurpose tree for future agro-forestry. ENVIS *For. Bull*. 1(1): 18-20, 2001.

Sastry, T.C. and Kanathekar, K.V. CSIR., Plants for reclamation of wastelands. CSIR, New Delhi, India, 1990.

Savill, P. S.). The silviculture of trees used in British forestry. CABI, 2013.

Schlotterer, C. The evolution of molecular markers- just a matter of fashion. *Nat. Rev. Genet.* 5, 63–69, 2004.

Seiwa, K., Tozawa, M., Ueno, N., Kimura, M., Yamasaki, M., Maruyama, K. Roles of cottony hairs in directed seed dispersal in riparian willows. *Plant Ecol* 198:27–35, 2008.

Sherwin, W.B., Jabot, F., Rush, R., Rossetto, M. Measurement of biological information with applications from genes to landscapes. *Molecular Ecology,* 15, 2857–2869, 2006.

Singh, M.K., Singh, N.B., Sankhyan, H.P., and Naik, P.K. Congruence of random amplification of polymorphic deoxyribonucleic acid (RAPD) and simple sequence repeats (SSR) markers in genetic characterization of willow (*Salix* spp). Vol. 13(32), pp. 3217-3229, 2014. doi: 10.5897/AJB013.1544.

Singh, N. B., Singh, M. K., Naik, P. K., Thakur, S. and Sharma, J. P. Analysis of genetic diversity in female, male and half sibs willow genotypes through

RAPD and SSR markers. *African Journal of Biotechnology*. Vol. 12(29), pp. 4578-4587, 2013b. doi: 10.5897/AJB2013.12120.

Singh, N.B., Joshi, S., Choudhary, P., Sharma, J.P. SSR Dna Marker Aided Genetic Diversity Assessment Of Selected Willow Clones. *Genetika,* Vol. 45, No.2,527-536, 2013a.

Skvortsov, A.K. and Edmondson, J. D., L. *In: Davis P.H. (ed) Flora of Turkey and the Eaat Aegean Islands*, Salix*,* 7, 694-716, 1970.

Skvortsov, A.K., *Willows of Russia and adjacent countries: Taxonomical and Geographical Revision* (transl from: Skvortsov AK (1968) Willows of the USSR: Taxonomic and Geographic Revision. Nauka, Moscow), ed: Zinovjev, A.G., Argus, G.W., Tahvanainen, J. and Roininen, H., Joensuu Unviersity, Joensuu, 307, 1999.

Slatkin, M. A measure of population Subdivision based on microsatellite allele frequencies. *Genetics,* 139:457–462, 1995.

Sochor, M., Vasut, R.J., Bartova, E., Majesky, L., Mracek, J. Can gene flow among populations counteract the habitat loss of extremely fragile biotopes? An example from the population genetic structure in Salix daphnoides. *Tree Genet Genomes,* 9:1193–1205, 2013.

Sokal, R.R., Rohlf, F.J. Biometry: The Principles and Practice of Statistics in Biological Research, 3rd edn. New York: W. H. Freeman & Co, 1995.

Stamati, K., Blackie, S., Brown, J.W.S., Russell, J. A set of polymorphic SSR loci for subarctic willow (*Salix lanata, S. lapponum* and *S. herbacea*). *Mol Ecol Notes,* 3: 280-282, 2003.

Stamati, K., Hollingsworth, P.M., Russell, J. Patterns of clonal diversity in three species of sub-arctic willow (*Salix lanata, Salix lapponum* and *Salix herbacea*). *Plant Syst Evol*, 269:75–88, 2007.

Steltzer, H., Hufbauer, R.A., Welker J. M., Casalis, M., Sullivan, P. F., and Chimner, R.. Frequent sexual reproduction and high intraspecific variation in *Salix arctica*: Implications for a terrestrial feedback to climate change in the High Arctic. *Journal Of Geophysical Research*, Vol. 113, G03S10, 2008. doi:10.1029/2007JG000503.

Stettler, R. F. Cottonwood and the river of time: on trees, evolution, and society.– University of Washington Press, Washington, 2011.

Stott, K.G. Willows in tbe service of man. In: Watling R, Ravens JA, eds. Proceedings: IVillotu Symposium, Edinburgh: Royal Botanic Gardens, *Edinburgh Botanical Society*, 169-182, 1992.

Suda, Y. and Argus, G.W. Chromosome numbers of some North American Salix. Brittonia 20, 191–197, 1968.

Szulkin, M., Bierne, N., David, P. Heterozygosity-fitness correlations: a time for reappraisal. *Evolution*, 64, 1202–1217, 2010.

Taberlet, P., Luikart, G. Non-invasive genetic sampling and individual identification. *Biol J Linn Soc,* 68: 41–55, 1999.

Taylor, D. R., Ingvarsson, P. K. Common features of segregation distortion in plants and animals. *Genetica*, 117, 27–35, 2003. doi:10.1023/A:1022308414864.

Terzioğlu, S., Serdar, B., Karaköse, M., CoGkunçelebi, K. and Gültepe, M., New data on *Salix anatolica* (Salicaceae) endemic to Turkey, *PHYTOTAXA*, vol.167, pp.111-118, 2014.

The European Atlas of Forest Tree Species. http://forest.jrc.ec.europa.eu/european-atlas-of-fores. Accessed August 2017.

Tillesse, V., Nef, L., Charles, J., Hopkin, A., Augustin, S. Damaging poplar Insects - Internationally important species. International Poplar Commission, FAO, Rome, 2007.

Toplu, F., Tunçtaner, K., Tulukçu, M., Kahraman, T., Küçükosmanoğlu. F. KirGehir Yöresine Uygun Söğüt (*Salix* L.) Klonlarının Belirlenmesine GiGkin AraGtırmalar. Teknik Bülten No: 203. Çevre Ve Orman Bakanlığı Kavak Ve Hızlı GeliGen Orman Ağaçları AraGtırma Enstitüsü, 2004.

Toro, M., Ferna´ndez, J., Caballero, A. Molecular characterization of breeds and its use in conservation. *Livest Sci,* 120: 174–195, 2009.

Trybush, S. O ., Jahodová, Š., Čížková, L., Karp, A., Hanley, S.J. High Levels of Genetic Diversity in *Salix viminalis* of the Czech Republic as Revealed by Microsatellite Markers. *Bioenerg. Res*. 5:969–977, 2012. doi: 10.1007/s12155-012-9212-4.

Tunçtaner, K. ÇeGitli Söğüt Klonlarının Genetik Varyasyonları ve Türkiye"nin DeğiGik Yörelerine Adaptasyonları Üzerine AraGtırmalar. Tarım, Orman ve Köy GöGleri Bakanlığı, Orman Genel Müdürlüğü, Kavak ve Hızlı GeliGen Orman Ağaçları AraGtırma Enstitüsü,1990-2, Teknik Bülten No: 150, Gzmit.

Tunçtaner, K. Türkiye" de Kavak ve Söğütlerde Gen Kaynaklarının Korunması ve Değerlendirilmesi. Orman Bakanlığı, Kavak ve Hızlı GeliGen Tür Orman Ağaçları AraGtırma Müdürlüğü. ÇeGitli yayınlar serisi no: 3.Gzmit, 1993.

Tunçtaner, K., Akbulut, T., Tulukçu, M. Bazı Kavak Ve Söğüt Klonlarının Göller Bölgesine Adaptasyonları Ve Yonga Levha Endüstrisinde Değerlendirilme Olanakları. Teknik Bülten No: 194, GZMGT. Orman Bakanlığı, Kavak ve Hızlı GeliGen Orman Ağaçları AraGtırma Enstitüsü, 2001.

Turkish Plants Data Service (TÜBGVES),Version 2.0 BETA http://www.tubives.com/index.php?sayfa=karsilastir. Accessed August 2017.

Turner, J. G., Davis, J. M. L., Guven, K., Proceedings of the Royal Society of Edinburgh. *Section B. Biological Sciences,* 98, 105, 1992.

Tuskan, G. A., Gunter, L. E., Yang, Z. K., Yin T., Sewell, M. M., DiFazio, S. P.. Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. Canadian Journal of Forest Research 34: 85–93, 2004.

Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U. The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray). *Science* 313, 1596–1604, 2006.

Ukwubile, C. A., Ahuchaogu, C. E., Tajudeen, B. L.. Congruence of random amplification of polymorphic deoxyribonucleic acid (RAPD) and simple sequence repeats (SSR) markers in genetic characterization of willow (*Salix* spp.). *European Journal of Biotechnology and Bioscience*, 2 (1): 11-23, 2014.

Uludağ, S., Zoralioğlu, T., Kılıçaslan, H., Karabulut, S. 2004. Farklı Çelik Boy Ve Kalınlıklarının Kaliteli Söğüt Fidanı YetiĞtirilmesine Olan Etkilerinin AraĞtırılması. Teknik Bülten No: 198. T. C. Çevre Ve Orman Bakanlığı Kavak Ve Hızlı GeliĞen Orman Ağaçları AraĞtırma Müdürlüğü Ġzmit.

Uotila, P. Salicaceae. – In: Euro+Med Plantbase - the information resource for Euro-Mediterranean plant diversity, 2011. http://www.emplantbase.org/home.html

Vamerali, T., Bandiera, M., Coletto, L., Zanetti, F., Dickinson, N. M., Mosca, G. Phytoremediation trials on metal- and arsenic-contaminated pyrite wastes (Torviscosa, Italy). Environmental Pollution 157: 887–894, 2009.

Vande, W.I., Van Camp, N., Van de Casteele, L., Verheyen, K., Lemeur, R. Short-rotation forestry of birch, maple, poplar and willow in Flanders (Belgium) II. Energy production and CO2 emission reduction potential. *Biomass Bioenerg*; 31: 276e83, 2007.

Vashegyi, Á., Mezôsi, G., Barta, K., Farsang, A., Dormány, G., Bartha, B., Pataki, S., Erdei, L. Phytoremediation of heavy metal pollution: A case study. *Acta Biologica Szegediensis*, Volume 49(1-2):77-79, 2005.

Vassilev, A., Perez-Sanz, A., Cuypers, A.and Vangronsveld, J. Tolerance of Two Hydroponically Grown *Salix* Genotypes to Excess Zinc. Journal of Plant Nutrition, 30: 1471–1482, 2007. Doi: 10.1080/01904160701555671.

Velioğlu, E. And Akgül, S. Poplars and Willows in Turkey: Country Progress Report of the National Poplar Commision Time period: 2012-2015, 2016.

Verstrepen, K.J., Jansen, A., Lewitter, F., Fink, G.R. Intragenic tandem repeats generate functional variability. *Nat. Genet.* 37, 986–990, 2005.

Verwijst, T. Willows: An underestimated resource for environment and society. THE *FORESTRY CHRONICLE*, VOL. 77, NO. 2, 2001.

Vieira, M. L. C., Santini L., Diniz, A.L. and Munhoz, C.de F. Microsatellite markers: what they mean and why they are so useful. *Genetics and Molecular Biology*, 39, 3, 312-328, 2016. doi: http://dx.doi.org/10.1590/1678-4685-GMB-2016-0027.

Vries, De, M. G. Conservation of natural ecosystems of poplar and willow. *Forestry Chron.*77(2), 255–257, 2001.

Vries, S.M.G. Conservation of natural ecosystems of poplar and willow. *The Forestry Chronicle*, Vol. 77, No. 2, 2001.

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

Wang, Z. S., Du, S. H., Dayanandan, S., Wang, D. S., Zeng, Y. F., Zhang, J. G. Phylogeny reconstruction and hybrid analysis of *Populus* (Salicaceae) based on nucleotide sequences of multiple single-copy nuclear genes and plastid fragments. *PLoS ONE,* 9:e103645, 2014.

Web site of the International Populus Genome Consortium (2014) Available at http://www.ornl.gov/sci/ipgc/ssr_resources.htm. Accessed March 2014.

White, T. L., Adams, W. T., Neale, D. B. Forest genetics. Year of Publication 2007. ISBN 9781845932855. doi: 10.1079/9781845932855.

Williamson-Natesan, E.G. Comparison of methods for detecting bottlenecks from microsatellite loci. *Conservat Genet.* 6:551–562, 2005.

Wright, S. 1951. The genetical structure of populations. Ann. Eugenics 15, 323-354.

Wu, J., Nyman, T., Wang, D., Argus, G.W., Yang, Y. and Chen, J. Phylogeny of *Salix* subgenus *Salix* s.l. (Salicaceae): delimitation biogeography and reticulate evolution. *BMC Evolutionary Biology,* 15:31, 2015.

Xu, Y. Molecular plant breeding, Cabi, 2010.

Yang, H.B., Kang, W.H., Nahm, S.H., Kang, B.C. Methods for Developing Molecular Markers. In: Koh HJ., Kwon SY., Thomson M. (eds) Current Technologies in Plant Molecular Breeding. Springer, *Dordrecht*, 2015. doi: https://doi.org/10.1007/978-94-017-9996-6_2.

Yasaman, S. H., Reza, V., Mohammed, R. Z., Bahman, S. H. Patterns of species richness and diversity of willow *Salix alba* pests in the West Azerbaijan Province of Iran. *Applied Ecology And Environmental Research* 15(3): 1227-1242, 2017. Doi: http://dx.doi.org/10.15666/aeer/1503_12271242.

Zabihi, N.A., Mahmoudabady, M., Soukhtanloo, M., Hayatdavoudi, P., Beheshti, F., Niazmand, S. *Salix alba* attenuated oxidative stress in the heart and kidney of hypercholesterolemic rabbits. *Avicenna J Phytomed,* 2017. Epub ahead of print.

Zielinski, J. and Tomaszevski, D. *Salix anatolica* (Salicaceae), a new species from Turkey, *Annual Botany Fennici,* 45: 386-388, 2007.

Van Der Schoot, J., Pospiskova, M., Vosman, B., Smulders, M.J.M. Development and characterization of microsatellitemarkers in black poplar (*Populus nigra* L.) *Thoer Appl Genet* 101(1–2):317–322, 2000.

Smulders, M.J.M., Van Der Schoot, J., Arens, P., Vosman, B. Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.) *Mol Ecol Notes* 1:188–190, 2001.

APPENDICES

APPENDIX A:

GENOMIC DNA ISOLATION

Total DNA was isolated from young leaves of each sample with a modified CTAB (cetyl trimethylammonium bromide) protocol (Doyle and Doyle, 1987). The procedure for DNA isolation *Salix* leaf samples is given below:

1. 0.1 gram fresh leaf tissue from every individual was put in autoclave-sterilized mortar and grounded 1000 µl extraction buffer -2X CTAB (CTAB (pH : 8.0), Tris HCl (pH:8.0), EDTA, and NaCl**).**
2. Liquid mixture was poured into 1.5 ml eppondorf tubes and 700 µl CTAB, 200 µl β-mercapto-ethanol and 5 µl Proteinase K were added to extraction buffer.
3. Tubes were incubated at 65 ºC for at least 30 minutes, mixing twice at every 10 minutes.
4. After incubation, they were centrifuged at 15000 rpm, +4ºC for 15 minutes.
5. The aqueous phase (top part) of the mixture was transferred into the new eppendorf tubes and 0,8Volume Phenol: Chloroform: Isoamyl alcohol (25:24:1) was added. If phase is not clear, this step is repeated.
6. After adding phenol step, they were centrifuged at 15000 rpm, +4ºC for 15 minutes.
7. After centrifugation, the supernatant was transferred to new tubes and 0, 8 Volume Chloroform IAA (24:1) was added to mixture and were shaken gently.
8. The mixture was centrifuged at 15000 rpm +4ºC for 15 minutes.
9. Supernatant were taken and 0, 7-1V isopropanol were added to extraction mixture. And then mixed very gently. The diffused DNA becomes visible in the alcohol.

10. The samples were incubated at -20ºC for 2 hours. After cold incubation, they were centrifuged at 13000 rpm in +4ºC for 8 minutes.
11. The pellet is washed by 500 ml cold 70% ethanol.
12. Supernatant were discarded and the tubes with pellet were inverted on a clean tissue paper and allowed to dry for 60 minutes.
13. After being sure that DNA pellet are dry, they were hydrated with 50-75 µl with TE (Tris HCL (Ph 7.0) and EDTA) and resuspended overnight at +4ºC.

APPENDIX B

BUFFERS CHEMICALS AND EQUIPMENTS

Buffers and solutions for DNA isolation

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

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

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

28 ml NaCl is completed with 100 mL distilled water

Phenol, (AMRESCO)**:** Pure phenol

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

Ethanol: 70% in distilled water

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

TE buffer: 10mm Tris HCL (pH:7) 10mm ethylene diamine tetra acetic acid disodium salt (EDTA)

Isopropanol, (FLUKA) : Pure Isopropanol, ice cold

Buffers and solutions for PCR

Sterile water

Taq DNA Polymerase (SIGMA Red *Taq*): 1U/µl

10X PCR buffer including MgCl2 (SIGMA)

dNTPs (SIGMA): 10mM

DNA: 10ng/ µl

Primer Pairs: 10µM

Agarose Gel Electrophoresis Buffers and Gel System

10X TBE Buffer: 108 gr Trizma Base, (SIGMA), 55 gr Boric Acid, (SIGMA)

Running Buffers: X TBE prepared in distilled water

Ethyidium Bromide,(SIGMA):4 mg/ ml

Agarose, (SIGMA): 3 % Agorose Gel

40 ml EDTA, (FLUKA) (0.5 M, pH:8) completed with 1000 ml with distilled water

Low molecular weight DNA Ladder (SIGMA)

Equipments

Autoclave: Yamato

Centrifuge: Nüve- NF048

Electrophoresis System: Thermo Scientific

Thermocyclers: Eppendorf- Mastercycler

Deepfreezer: UĞUR- Freezer

Magnetic Stirrer: Labor Brand – Hotplate L-81

Refrigerator: Siemens

UV Transilluminator: Vilbor Lourmant

Vortex: Nüve- NM110

Water Bath: Memmert

Oven: Dedeoğlu

Micropipettes: Gılson

pHmeter: Hanna Inst.

APPENDIX C

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

1. Registration of pcr product (by Customer)

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

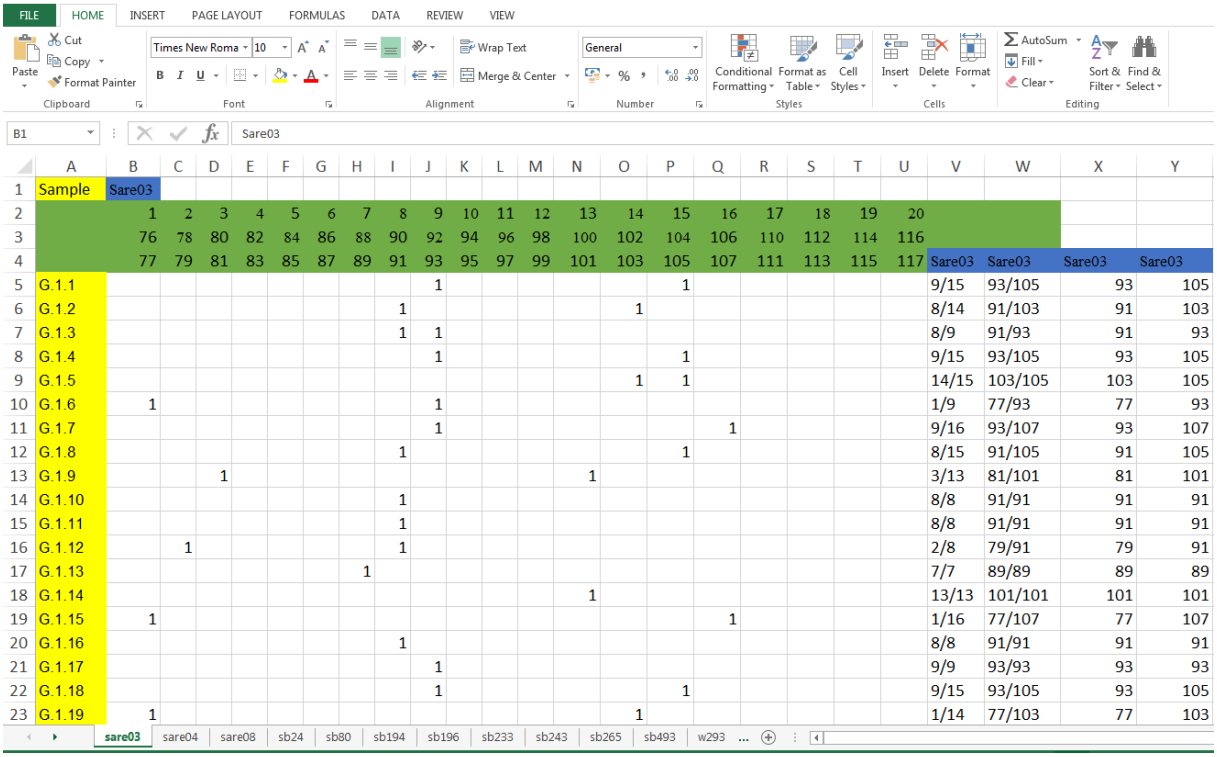
3. denaturation

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

5. Genemapper v.5 analysis

APPENDIX D

A PART OF EXCEL MATRIX SHOWING GENOTYPES ALLELE SIZES



Sample	Sare03																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20				
	76	78	80	82	84	86	88	90	92	94	96	98	100	102	104	106	110	112	114	116				
	77	79	81	83	85	87	89	91	93	95	97	99	101	103	105	107	111	113	115	117	Sare03	Sare03	Sare03	Sare03
G.1.1									1						1						9/15	93/105	93	105
G.1.2								1						1							8/14	91/103	91	103
G.1.3								1	1												8/9	91/93	91	93
G.1.4									1						1						9/15	93/105	93	105
G.1.5														1	1						14/15	103/105	103	105
G.1.6	1								1												1/9	77/93	77	93
G.1.7									1							1					9/16	93/107	93	107
G.1.8								1							1						8/15	91/105	91	105
G.1.9			1										1								3/13	81/101	81	101
G.1.10								1													8/8	91/91	91	91
G.1.11								1													8/8	91/91	91	91
G.1.12		1						1													2/8	79/91	79	91
G.1.13							1														7/7	89/89	89	89
G.1.14													1								13/13	101/101	101	101
G.1.15	1															1					1/16	77/107	77	107
G.1.16								1													8/8	91/91	91	91
G.1.17									1												9/9	93/93	93	93
G.1.18									1						1						9/15	93/105	93	105
G.1.19	1													1							1/14	77/103	77	103

APPENDIX E

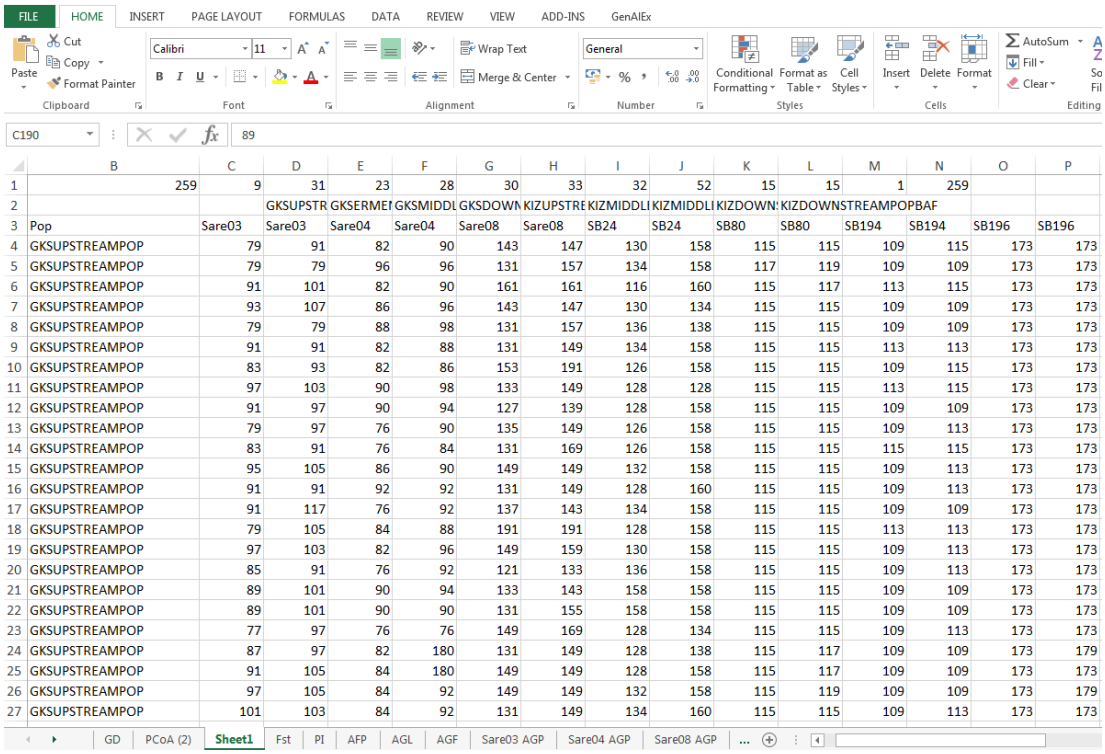
EXAMPLES OF DATA FILE FORMATS

Genepop data format

File Edit Format View Help

```
Title line:  "PopulationData.txt"
Sare03
Sare04
Sare08
SB24
SB80
SB194
SB196
SB233
SB243
SB265
SB493
W293
W784
gSlMCT01
gSlMCT02
gSlMCT05
WPMS18
PMGC2709
PMGC2889
PMGC2163
Pop
1_      ,       079091 082090 143147 130158 115115 109115 173173 180180 101104 135135 212216 112114 152158 400400 300300 201211 220226 194194 192202 198216
1_      ,       079079 096096 131157 134158 117119 109109 173173 180186 104110 135135 212212 112114 152158 400400 302302 197199 220220 196208 192206 198212
1_      ,       091101 082090 161161 116160 115117 113115 173173 176178 104107 135135 212212 112114 152158 400400 302302 185199 220220 202208 188196 198212
1_      ,       093107 086096 143147 130134 115115 109109 173173 216216 104104 135135 210212 112114 152158 400400 296296 199201 214220 214224 196202 198204
1_      ,       079079 088098 131157 136138 115115 109109 173173 204216 101104 135135 210212 112114 152158 400400 302302 197197 220220 192214 206206 198204
1_      ,       091091 082088 131149 134158 115115 113113 173173 198214 104110 135135 212212 112120 152158 400400 296302 203207 220220 208214 202202 198216
1_      ,       083093 082086 153191 126158 115115 109115 173173 180180 101107 135135 210212 112114 152158 400400 294296 197213 220220 160208 204204 198198
1_      ,       097103 090098 133149 128128 115115 113115 173173 180216 101104 135135 212214 112114 152158 400400 296302 185211 220220 202208 202202 198212
1_      ,       091097 090094 127139 128158 115115 109109 173173 180180 104107 135135 212212 114116 152158 400400 296300 197199 220220 190190 194202 198204
1_      ,       079097 076090 135149 126158 115115 109113 173173 180180 101104 135135 210216 112114 152158 400400 296298 199199 220220 200208 192202 198212
1_      ,       083091 076084 131169 126158 115115 115115 173173 180180 104104 135135 212212 112116 152158 400400 296296 207213 217217 198208 196196 198214
1_      ,       095105 086090 149149 132158 115115 109113 173173 174180 104107 135135 212216 000000 148152 400400 296302 191211 220220 198208 194194 198216
1_      ,       091091 092092 131149 128160 115115 109113 173173 180202 104107 135135 210216 112114 148152 400400 298298 191197 220220 202208 202202 198216
1_      ,       091117 076092 137143 134158 115115 109109 173173 180202 104110 135135 212212 112114 152158 400400 296300 199211 220220 208214 192194 198216
1_      ,       079105 084088 191191 128158 115115 113113 173173 180216 104104 135135 210212 112112 148152 400400 296300 197211 217217 198198 192212 198214
1_      ,       097103 082096 149159 130158 115115 109113 173173 180202 101107 135135 210210 112114 152158 400400 296300 191197 217220 208214 192202 198212
1_      ,       085091 076092 121133 136158 115115 109113 173173 180216 101104 135135 210212 112114 152158 400400 300302 185199 220220 160208 202202 198212
1_      ,       089101 090094 133143 158158 115115 109109 173173 180216 101104 135135 210210 112114 152158 400400 296302 185199 220220 208208 196202 198212
1_      ,       089101 090090 131155 158158 115115 109109 173173 180206 000000 135135 210210 112114 152158 400400 296300 000000 220220 202208 192200 198212
1_      ,       077097 076076 149169 128134 115115 109113 173173 180202 101104 135135 210210 130130 152158 400400 296296 199213 220220 190190 192200 198198
1_      ,       087097 082180 131149 128138 115117 109109 173179 180202 101104 135135 210212 114114 152158 286286 296300 191191 220220 190194 192192 198216
1_      ,       091105 084180 149149 128158 115117 109109 173173 180202 101104 135135 212212 112114 152158 400400 296300 201213 220220 198208 194202 198216
1_      ,       097105 084092 149149 132158 115119 109109 173179 176184 101107 109135 210212 112114 152158 286400 300302 185213 214220 194208 192204 198212
1_      ,       101103 084092 131149 134160 115115 109113 173173 216216 101104 135135 210214 112114 150152 400400 298302 185197 220226 198208 190200 198216
1_      ,       097105 092180 131149 128160 115115 109109 173173 202210 101104 137139 210212 112114 152158 400400 296296 197199 220226 188208 196202 198216
1_      ,       097105 086090 139147 134158 115115 109113 173173 180202 101113 135135 210212 112114 152158 400400 300300 197221 220220 198208 184196 198216
1_      ,       091105 090120 131147 158158 115115 109109 173173 216216 101104 135135 210210 112114 152158 400400 296300 189199 220220 208214 180202 198198
1_      ,       093105 084090 133147 128158 115115 109109 173173 180180 107110 135135 212212 112112 152158 400400 296302 191211 220226 188188 180202 198216
1_      ,       097103 086094 139147 126158 115115 109109 173173 180212 101110 135135 210212 112116 152158 400400 298298 199201 223226 198208 202202 198216
1_      ,       105111 094102 139147 128158 115119 109113 173173 180180 101104 135135 210210 108108 152158 400400 300302 189199 223226 198214 182192 198216
1_      ,       091093 092098 149149 128158 115115 109115 173173 180202 101113 135135 210210 108108 152158 400400 298298 213217 220220 198214 204204 198216
Pop
2_      ,       089103 088096 139139 128158 115119 109113 173173 180180 101107 133133 212212 112130 152158 400400 300300 195217 220220 188208 182202 198216
2_      ,       103103 096096 143143 134158 119119 109113 173173 212202 104104 135135 208212 112114 152158 400400 300302 199199 220220 194208 192192 198216
2_      ,       089095 088098 139139 134158 115115 109109 173173 180212 101104 135135 212216 112114 152158 400400 296298 185199 220220 202208 192204 198212
2_      ,       103103 082088 153155 128158 115115 113113 173173 180202 101104 135135 210212 112114 152158 286400 300300 199205 220220 188188 192212 198216
2_      ,       103111 082088 131147 128158 115119 109113 173173 180180 101107 135135 210212 112114 150152 400400 294296 213213 220220 208208 192202 198216
2_      ,       097103 088094 131161 126158 115115 109113 173173 212202 101107 135135 212212 112118 152158 400400 300302 197213 220220 188208 182192 198216
2_      ,       091097 082096 149191 134158 115115 109113 173173 216216 101104 135135 210210 112114 152158 400400 296296 197203 214223 190190 182194 198216
```

GenAIEx data format

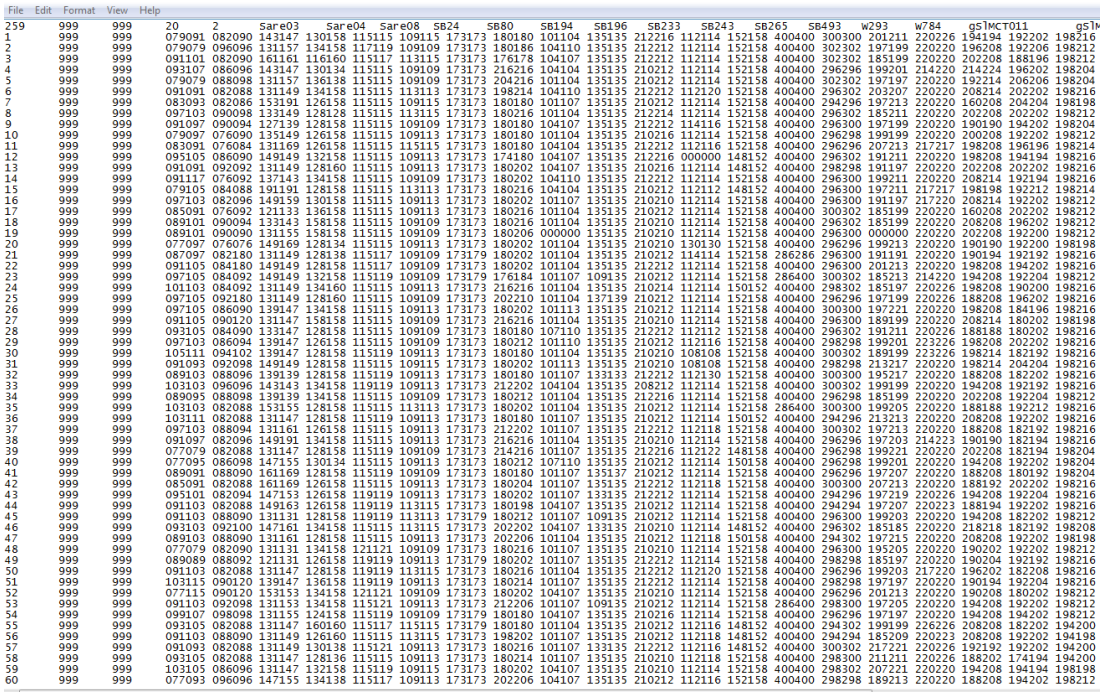


	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		259	9	31	23	28	30	33	32	52	15	15	1	259		
2				GKSUPSTR	GKSERMEI	GKSMIDDL	GKSDOWN	KIZUPSTRE	KIZMIDDLI	KIZMIDDLI	KIZDOWN	KIZDOWNSTREAM	POPBAF		
3	Pop	Sare03	Sare03	Sare04	Sare04	Sare08	Sare08	SB24	SB24	SB80	SB80	SB194	SB194	SB196	SB196
4	GKSUPSTREAMPOP	79	91	82	90	143	147	130	158	115	115	109	115	173	173
5	GKSUPSTREAMPOP	79	79	96	96	131	157	134	158	117	119	109	109	173	173
6	GKSUPSTREAMPOP	91	101	82	90	161	161	116	160	115	117	113	115	173	173
7	GKSUPSTREAMPOP	93	107	86	96	143	147	130	134	115	115	109	109	173	173
8	GKSUPSTREAMPOP	79	79	88	98	131	157	136	138	115	115	109	109	173	173
9	GKSUPSTREAMPOP	91	91	82	88	131	149	134	158	115	115	113	113	173	173
10	GKSUPSTREAMPOP	83	93	82	86	153	191	126	158	115	115	109	115	173	173
11	GKSUPSTREAMPOP	97	103	90	98	133	149	128	128	115	115	113	115	173	173
12	GKSUPSTREAMPOP	91	97	90	94	127	139	128	158	115	115	109	109	173	173
13	GKSUPSTREAMPOP	79	97	76	90	135	149	126	158	115	115	109	113	173	173
14	GKSUPSTREAMPOP	83	91	76	84	131	169	126	158	115	115	115	115	173	173
15	GKSUPSTREAMPOP	95	105	86	90	149	149	132	158	115	115	109	113	173	173
16	GKSUPSTREAMPOP	91	91	92	92	131	149	128	160	115	115	109	113	173	173
17	GKSUPSTREAMPOP	91	117	76	92	137	143	134	158	115	115	109	109	173	173
18	GKSUPSTREAMPOP	79	105	84	88	191	191	128	158	115	115	113	113	173	173
19	GKSUPSTREAMPOP	97	103	82	96	149	159	130	158	115	115	109	113	173	173
20	GKSUPSTREAMPOP	85	91	76	92	121	133	136	158	115	115	109	113	173	173
21	GKSUPSTREAMPOP	89	101	90	94	133	143	158	158	115	115	109	109	173	173
22	GKSUPSTREAMPOP	89	101	90	90	131	155	158	158	115	115	109	109	173	173
23	GKSUPSTREAMPOP	77	97	76	76	149	169	128	134	115	115	109	113	173	173
24	GKSUPSTREAMPOP	87	97	82	180	131	149	128	138	115	117	109	109	173	179
25	GKSUPSTREAMPOP	91	105	84	180	149	149	128	158	115	117	109	109	173	173
26	GKSUPSTREAMPOP	97	105	84	92	149	149	132	158	115	119	109	109	173	179
27	GKSUPSTREAMPOP	101	103	84	92	131	149	134	160	115	115	109	113	173	173

GDA data format

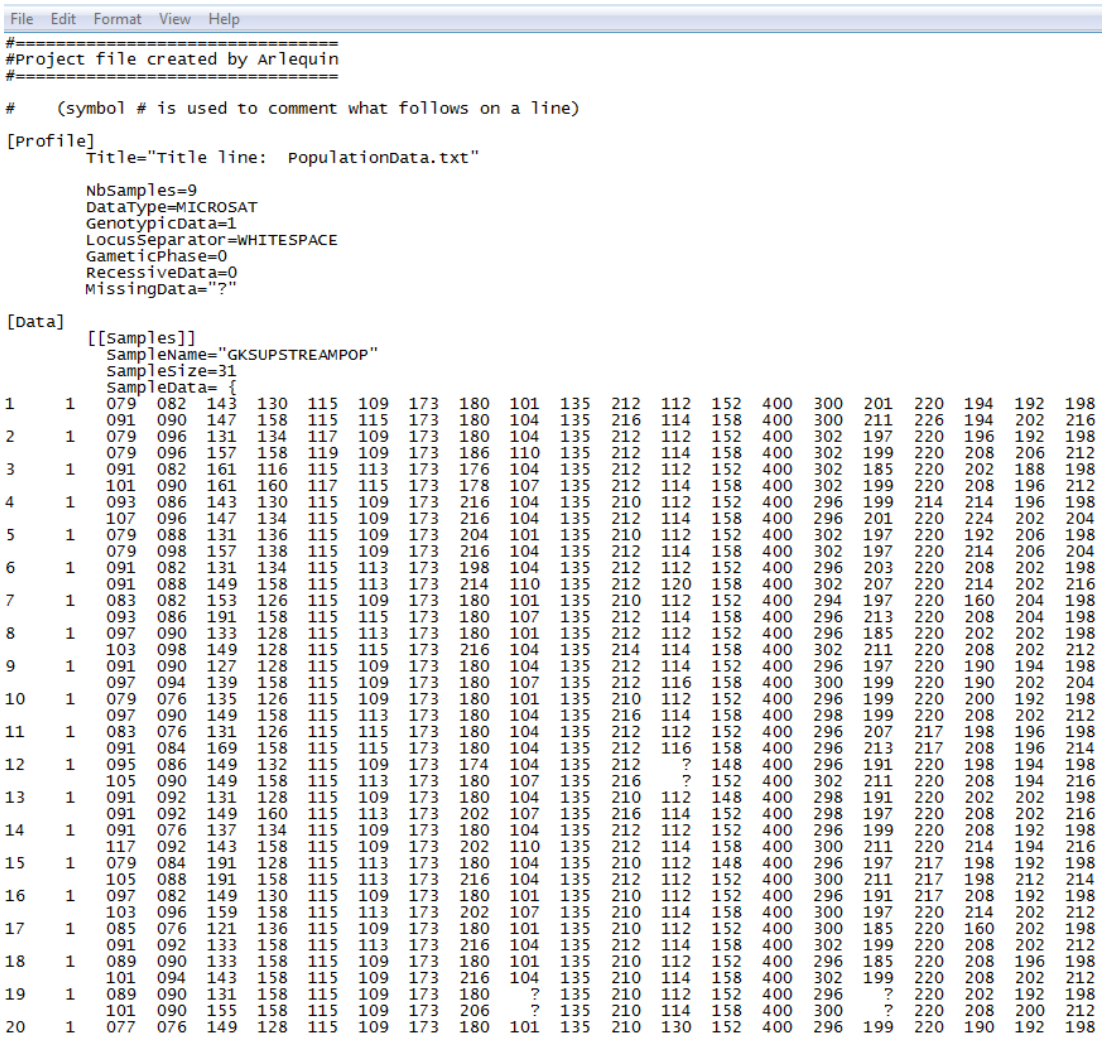
```
#NEXUS

[!
Data from salix
]

begin gdadata;
     dimensions nloci=20 npops=9;
     format tokens missing=? datapoint=standard;
     locusallelelabels

               1 'Sare03'   [/ 77 79 81 83 85 87 89 91 93 95 97 99 101 103 105 107 111 113 115 117],
               2 'Sare04'   [/ 76 82 84 86 88 90 92 94 96 98 100 102 104 106 108 110 112 114 120 122 124 126 130 180],
               3 'Sare08'   [/ 121 127 131 133 135 137 139 141 143 145 147 149 153 155 157 159 161 163 167 169 191],
               4 'SB24'   [/ 114 116 118 124 126 128 130 132 134 136 138 140 142 144 158 160],
               5 'SB80'   [/ 115 117 119 121],
               6 'SB194'   [/ 109 113 115],
               7 'SB196'   [/ 167 170 173 179],
               8 'SB233'   [/ 170 172 176 178 180 184 186 190 198 200 202 204 206 208 210 212 214 216 218 220],
               9 'SB243'   [/ 101 104 107 110 113],
               10 'SB265'   [/ 109 133 135 137 139],
               11 'SB493'   [/ 208 210 212 214 216 226 238 242 244],
               12 'W293'   [/ 108 112 114 116 118 120 122 124 126 130 136],
               13 'W784'   [/ 148 150 152 158 162 168],
               14 'gSlMCT011'   [/ 274 276 280 284 286 290 400],
               15 'gSlMCT024'   [/ 288 290 294 296 298 300 302 304],
               16 'gSlMCT052'   [/ 185 189 191 195 197 199 201 203 205 207 209 211 213 215 217 219 221 225 227 231 243 263 271 281],
               17 'WPMS18'   [/ 211 214 217 220 223 226 232],
               18 'PMGC2709'   [/ 160 166 172 180 184 186 188 190 192 194 196 198 200 202 204 208 210 212 214 216 218 224],
               19 'PMGC2889'   [/ 174 178 180 182 184 186 188 190 192 194 196 198 200 202 204 206 208 210 212 222],
               20 'PMGC2163'   [/ 186 188 194 198 200 204 206 208 210 212 214 216 218 224];

matrix
GKSUPSTREAMPOP:
G.4.1   79/91   82/90   143/147 130/158 115/115 109/115 173/173 180/180 101/104 135/135 212/216 112/114 152/158 400/400 300/300 201/211 220/226 194/194 192/202 198/216
G.4.2   79/79   96/96   131/157 134/158 117/119 109/109 173/173 180/186 104/110 135/135 212/212 112/114 152/158 400/400 302/302 197/199 220/220 196/208 192/206 198/212
G.4.3   91/101  82/90   161/161 116/160 115/117 113/115 173/173 176/178 104/107 135/135 212/212 112/114 152/158 400/400 302/302 185/199 220/220 202/208 188/196 198/212
G.4.4   93/107  86/96   143/147 130/134 115/115 109/109 173/173 216/216 104/104 135/135 210/212 112/114 152/158 400/400 296/296 199/201 214/220 214/224 196/202 198/204
G.4.5   79/79   88/98   131/157 136/138 115/115 109/109 173/173 204/216 101/104 135/135 210/212 112/114 152/158 400/400 302/302 197/197 220/220 192/214 206/206 198/204
G.4.6   91/91   82/88   131/149 134/158 115/115 113/113 173/173 198/214 104/110 135/135 212/212 112/120 152/158 400/400 296/302 203/207 220/220 208/214 202/202 198/216
G.4.7   83/93   82/86   153/191 126/158 115/115 109/115 173/173 180/180 101/107 135/135 210/212 112/114 152/158 400/400 294/296 197/213 220/220 160/208 204/204 198/198
G.4.8   97/103  90/98   133/149 128/128 115/115 113/115 173/173 180/216 101/104 135/135 212/214 112/114 152/158 400/400 296/302 185/211 220/220 202/208 202/202 198/212
G.4.10  91/97   90/94   127/139 128/158 115/115 109/109 173/173 180/180 104/107 135/135 212/212 114/116 152/158 400/400 296/300 197/199 220/220 190/190 194/202 198/204
G.4.11  79/97   76/90   135/149 126/158 115/115 109/113 173/173 180/180 101/104 135/135 210/216 112/114 152/158 400/400 296/298 199/199 220/220 200/208 192/202 198/212
G.4.12  83/91   76/84   131/169 126/158 115/115 115/115 173/173 180/180 104/104 135/135 212/212 112/116 152/158 400/400 296/296 207/213 217/217 198/208 196/196 198/214
G.4.13  95/105  86/90   149/149 132/158 115/115 109/113 173/173 174/180 104/107 135/135 212/216 ?/?     148/152 400/400 296/302 191/211 220/220 198/208 194/194 198/216
G.4.14  91/91   92/92   131/149 128/160 115/115 109/113 173/173 180/202 104/107 135/135 210/216 112/114 148/152 400/400 298/298 191/197 220/220 202/208 202/202 198/216
G.4.15  91/117  76/92   137/143 134/158 115/115 109/109 173/173 180/202 104/110 135/135 212/212 112/114 152/158 400/400 296/300 199/211 220/220 208/214 192/194 198/216
G.4.16  79/105  84/88   191/191 128/158 115/115 113/113 173/173 180/216 104/104 135/135 210/212 112/112 148/152 400/400 296/300 197/211 217/217 198/198 192/212 198/214
G.4.17  97/103  82/96   149/159 130/158 115/115 109/113 173/173 180/202 101/107 135/135 210/210 112/114 152/158 400/400 296/300 191/197 217/220 208/214 192/202 198/212
G.4.18  85/91   76/92   121/133 136/158 115/115 109/113 173/173 180/216 101/104 135/135 210/212 112/114 152/158 400/400 300/302 185/199 220/220 160/208 202/202 198/212
G.4.19  89/101  90/94   133/143 158/158 115/115 109/109 173/173 180/216 101/104 135/135 210/210 112/114 152/158 400/400 296/302 185/199 220/220 208/208 196/202 198/212
G.4.20  89/101  90/90   131/155 158/158 115/115 109/109 173/173 180/206 ?/?     135/135 210/210 112/114 152/158 400/400 296/300 ?/?     220/220 202/208 192/200 198/212
G.4.21  77/97   76/76   149/169 128/134 115/115 109/113 173/173 180/202 82/101  135/135 210/210 130/130 152/158 400/400 296/296 199/213 220/220 190/190 192/200 198/198
G.4.22  87/97   82/180  131/149 128/138 115/117 109/109 173/179 180/202 101/104 135/135 210/212 114/114 152/158 286/286 296/300 191/191 220/220 190/194 192/192 198/216
G.4.23  91/105  84/180  149/149 128/158 115/117 109/109 173/173 180/202 101/104 135/135 212/212 112/114 152/158 400/400 296/300 201/213 220/220 198/208 194/202 198/216
G.4.24  97/105  84/92   149/149 132/158 115/119 109/109 173/179 176/184 101/107 109/135 210/212 112/114 152/158 286/400 300/302 185/213 214/220 194/208 192/204 198/212
G.4.25  101/103 84/92   131/149 134/160 115/115 109/113 173/173 216/216 101/104 135/135 210/214 112/114 150/152 400/400 298/302 185/197 220/226 198/208 190/200 198/216
G.4.26  97/105  92/180  131/149 128/160 115/115 109/109 173/173 202/210 101/104 137/139 210/212 112/114 152/158 400/400 296/296 197/199 220/226 188/208 196/202 198/216
G.4.27  97/105  86/90   139/147 134/158 115/115 109/113 173/173 180/202 101/113 135/135 210/212 112/114 152/158 400/400 300/300 197/221 220/220 198/208 184/196 198/216
G.4.28  91/105  90/120  131/147 158/158 115/115 109/109 173/173 216/216 101/104 135/135 210/210 112/114 152/158 400/400 296/300 189/199 220/220 208/214 180/202 198/198
```

GeneClone data format



```
File  Edit  Format  View  Help
259     999     999     20      2       Sare03    Sare04  Sare08  SB24    SB80    SB194   SB196   SB233   SB243   SB265   SB493   W293    W784    gS1MCT011       gS1M
1       999     999     079091 082090 143147 130158 115115 109115 173173 180180 101104 135135 212216 112114 152158 400400 300300 201211 220226 194194 192202 198216
2       999     999     079079 086096 131157 134158 117119 109109 173173 180186 104110 135135 212212 112114 152158 400400 302302 197199 220226 196208 192206 198212
3       999     999     091101 082090 161161 116160 115117 113115 173173 176178 104107 135135 212212 112114 152158 400400 302302 185199 220220 202208 188196 198212
4       999     999     093107 086096 143147 130134 115115 109109 173173 216216 104104 135135 210212 112114 152158 400400 296296 199201 214220 214224 196202 198204
5       999     999     079079 088098 131157 136138 115115 109109 173173 204216 101104 135135 210212 112114 152158 400400 302302 197197 220220 192214 206206 198204
6       999     999     091091 082088 131149 134158 115115 113113 173173 198214 104110 135135 212212 112120 152158 400400 296302 203207 220220 208214 202202 198216
7       999     999     083093 082086 153191 126158 115115 109115 173173 180180 101107 135135 210212 112114 152158 400400 294296 197213 220220 160208 204204 198198
8       999     999     097103 090098 133149 128128 115115 113115 173173 180216 101104 135135 212214 112114 152158 400400 296302 185211 220220 202208 202202 198212
9       999     999     091097 090094 127139 128158 115115 109109 173173 180180 104107 135135 212212 114116 152158 400400 296300 197199 220220 190190 194202 198204
10      999     999     079097 076090 135149 126158 115115 109113 173173 180180 101104 135135 210216 112114 152158 400400 296298 199199 220220 200208 192202 198212
11      999     999     083091 076084 131169 126158 115115 115115 173173 180180 104104 135135 212212 112116 152158 400400 296296 207213 217217 198208 196196 198214
12      999     999     095105 086090 149149 132158 115115 109113 173173 174180 104107 135135 212216 000000 148152 400400 296302 191211 220220 198208 194194 198216
13      999     999     091091 092092 131149 128160 115115 109113 173173 180202 104107 135135 210216 112114 148152 400400 298298 191197 220220 202208 202202 198216
14      999     999     091117 076092 137143 134158 115115 109109 173173 180202 104110 135135 212212 112114 152158 400400 296300 199211 220220 208214 192194 198216
15      999     999     079105 084088 191191 128158 115115 113113 173173 180216 104104 135135 210212 112112 148152 400400 296300 197211 217217 198198 192212 198214
16      999     999     097103 082096 149159 130158 115115 109113 173173 180202 101107 135135 210210 112114 152158 400400 296300 191197 217220 208214 192202 198212
17      999     999     085091 076092 121133 136158 115115 109113 173173 180216 101104 135135 210212 112114 152158 400400 300302 185199 220220 160208 202202 198212
18      999     999     089101 090094 133143 158158 115115 109109 173173 180216 101104 135135 210210 112114 152158 400400 296302 185199 220220 208208 196202 198212
19      999     999     089101 090090 131155 158158 115115 109109 173173 180206 000000 135135 210210 112114 152158 400400 296300 000000 220220 202208 192200 198212
20      999     999     077097 076076 149169 126134 115115 109113 173173 180202 101104 135135 210210 130130 152158 400400 296296 199213 220220 190190 192200 198198
21      999     999     087097 082180 131149 128138 115117 109109 173179 180202 101104 135135 210212 114114 152158 286286 296300 191191 220220 190194 192192 198216
22      999     999     091105 084180 149149 128158 115117 109109 173173 180202 101104 135135 212212 112114 152158 400400 296300 201213 220220 198208 194202 198216
23      999     999     097105 084092 149149 132158 115119 109109 173179 176184 101107 109135 210212 112114 152158 286400 300302 185213 214220 194208 192204 198212
24      999     999     101103 084092 131149 134160 115115 109113 173173 216216 101104 135135 210214 112114 150152 400400 298302 185197 220226 198208 190200 198216
25      999     999     097105 092180 131149 128160 115115 109109 173173 202210 101104 137139 210212 112114 152158 400400 296296 197199 220226 188208 196202 198216
26      999     999     097105 086090 139147 134158 115115 109113 173173 180202 101113 135135 210212 112114 152158 400400 300300 197221 220220 198208 184196 198216
27      999     999     091105 090120 131147 158158 115115 109109 173173 216216 101104 135135 210210 112114 152158 400400 296300 189199 220220 208214 180202 198198
28      999     999     093105 084090 133147 128158 115115 109109 173173 180180 107110 135135 212212 112112 152158 400400 296302 191211 220226 188188 180202 198216
29      999     999     097103 086094 139147 126158 115115 109109 173173 180212 101110 135135 210212 112116 152158 400400 298298 199201 223226 198208 202202 198216
30      999     999     105111 094102 139147 128158 115119 109113 173173 180180 101104 135135 210210 108108 152158 400400 300302 189199 223226 198214 182192 198216
31      999     999     091093 092098 149149 128158 115115 109115 173173 180202 101113 135135 210210 108108 152158 400400 298298 213217 220220 198214 204204 198216
32      999     999     089103 088096 139139 128158 115119 109113 173173 180180 101107 133133 212212 112130 152158 400400 300300 195217 220220 188208 182202 198216
33      999     999     103103 096096 143143 134158 119119 109113 173173 212202 104104 135135 208212 112114 152158 400400 300302 199199 220220 194208 192192 198216
34      999     999     089095 088098 139139 134158 115115 109109 173173 180212 101104 135135 212216 112114 152158 400400 296298 185199 220220 202208 192204 198212
35      999     999     103103 082088 153155 128158 115115 113113 173173 180202 101104 135135 210212 112114 152158 286400 300300 199205 220220 188188 192212 198216
36      999     999     103111 082088 131147 128158 115119 109113 173173 180180 101107 135135 210212 112114 150152 400400 294296 213213 220220 208208 192202 198216
37      999     999     097103 088094 131161 126158 115115 109113 173173 212202 101107 135135 212212 112118 152158 400400 300302 197213 220220 188208 182192 198216
38      999     999     091097 082096 149191 134158 115115 109113 173173 216216 101104 135135 210210 112114 152158 400400 296296 197203 214223 190190 182194 198216
39      999     999     077079 082088 131147 128158 115119 109109 173173 214216 101107 135135 212216 112122 148158 400400 296298 199221 220220 202208 182194 198204
40      999     999     077095 086098 147155 130134 115115 109113 173173 180212 107110 135135 210212 112114 150158 400400 296298 199201 220220 194208 192202 198204
41      999     999     089091 088090 161169 128158 115119 109109 173173 180180 101107 135137 210212 112114 152158 400400 296296 197207 220220 188208 180192 198204
42      999     999     085091 082088 161169 126158 115115 109113 173173 180204 101107 135135 212212 112118 152158 400400 300300 207213 220220 188192 202202 198216
43      999     999     095101 082094 147153 126158 119119 109113 173173 180202 101107 133135 212212 112114 152158 400400 294296 197219 220226 194208 192204 198216
44      999     999     091103 082088 149163 126158 119119 113115 173173 180198 104107 135135 210212 112114 152158 400400 294294 197207 220223 188194 192202 198216
45      999     999     091103 088090 131131 128158 119119 113113 173179 180212 101107 109135 210212 112114 152158 400400 296300 199203 220220 194208 182202 198212
46      999     999     093103 092100 147161 134158 115115 113115 173173 202202 104107 133135 210210 112114 148152 400400 296302 185185 220220 218218 182192 198208
47      999     999     089103 088090 131161 128158 115115 109113 173173 202206 101104 135135 210212 112118 150158 400400 294302 197215 220220 208208 192202 198198
48      999     999     077079 082090 131131 134158 121121 109109 173173 180216 101107 135135 210210 112114 152158 400400 296300 195205 220220 190202 192202 198212
49      999     999     089089 088092 121131 126158 119119 109113 173179 180202 101107 135135 212212 112114 152158 400400 298298 185197 220220 190204 192192 198216
50      999     999     091103 082088 131147 128158 119119 113115 173173 180216 101104 135135 212212 112120 152158 400400 296296 199203 217220 196202 182208 198216
51      999     999     103115 090120 139147 136158 119119 109113 173173 180214 101107 135135 212212 112114 152158 400400 298298 197197 220220 190194 192204 198216
52      999     999     077115 090120 153153 134158 121121 109109 173173 180202 104107 135135 210210 112114 152158 400400 296296 201213 220220 190208 180202 198212
53      999     999     091103 092098 131153 134158 115121 109113 173173 212206 101107 109135 210212 112114 152158 286400 298300 197205 220220 194208 192202 198212
54      999     999     099107 098098 131155 124158 115119 109109 173179 180180 104107 135135 210216 112114 152158 400400 296296 197197 220220 194208 194202 198212
55      999     999     093105 082088 131147 160160 115117 115115 173179 180180 101104 135135 210212 112116 148152 400400 294302 199199 226226 208208 182202 194200
56      999     999     091103 088090 131149 126160 115115 113115 173173 198202 101107 135135 210212 112118 148152 400400 294294 185209 220223 208208 192202 194198
57      999     999     091093 082088 131149 130138 115121 109113 173173 180216 101107 133135 212212 112116 148152 400400 300302 217221 220226 192192 192202 194200
58      999     999     093105 082088 131147 128136 115115 109113 173173 180214 101107 135135 210210 112118 152158 400400 298300 211211 220226 188202 174194 194200
59      999     999     103105 086096 131147 132158 115119 109115 173173 180202 104107 135135 210210 112114 152158 400400 298302 207221 220220 194208 194194 198198
60      999     999     077093 096096 147155 134138 115117 109113 173173 202206 104107 135135 210212 112116 152158 400400 298298 189213 220220 188208 194202 198212
```

Arlequin data format



```
File  Edit  Format  View  Help
#=================================
#Project file created by Arlequin
#=================================

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

[Profile]
        Title="Title line:  PopulationData.txt"

        NbSamples=9
        DataType=MICROSAT
        GenotypicData=1
        LocusSeparator=WHITESPACE
        GameticPhase=0
        RecessiveData=0
        MissingData="?"

[Data]
        [[Samples]]
          SampleName="GKSUPSTREAMPOP"
          SampleSize=31
          SampleData= {
1     1   079  082  143  130  115  109  173  180  101  135  212  112  152  400  300  201  220  194  192  198
            091  090  147  158  115  115  173  180  104  135  216  114  158  400  300  211  226  194  202  216
2     1   079  096  131  134  117  109  173  180  104  135  212  112  152  400  302  197  220  196  192  198
            079  096  157  158  119  109  173  186  110  135  212  114  158  400  302  199  220  208  206  212
3     1   091  082  161  116  115  113  173  176  104  135  212  112  152  400  302  185  220  202  188  198
            101  090  161  160  117  115  173  178  107  135  212  114  158  400  302  199  220  208  196  212
4     1   093  086  143  130  115  109  173  216  104  135  210  112  152  400  296  199  214  214  196  198
            107  096  147  134  115  109  173  216  104  135  212  114  158  400  296  201  220  224  202  204
5     1   079  088  131  136  115  109  173  204  101  135  210  112  152  400  302  197  220  192  206  198
            079  098  157  138  115  109  173  216  104  135  212  114  158  400  302  197  220  214  206  204
6     1   091  082  131  134  115  113  173  198  104  135  212  112  152  400  296  203  220  208  202  198
            091  088  149  158  115  113  173  214  110  135  212  120  158  400  302  207  220  214  202  216
7     1   083  082  153  126  115  109  173  180  101  135  210  112  152  400  294  197  220  160  204  198
            093  086  191  158  115  115  173  180  107  135  212  114  158  400  296  213  220  208  204  198
8     1   097  090  133  128  115  113  173  180  101  135  212  112  152  400  296  185  220  202  202  198
            103  098  149  128  115  115  173  216  104  135  214  114  158  400  302  211  220  208  202  212
9     1   091  090  127  128  115  109  173  180  104  135  212  114  152  400  296  197  220  190  194  198
            097  094  139  158  115  109  173  180  107  135  212  116  158  400  300  199  220  190  202  204
10    1   079  076  135  126  115  109  173  180  101  135  210  112  152  400  296  199  220  200  192  198
            097  090  149  158  115  113  173  180  104  135  216  114  158  400  298  199  220  208  202  212
11    1   083  076  131  126  115  115  173  180  104  135  212  112  152  400  296  207  217  198  196  198
            091  084  169  158  115  115  173  180  104  135  212  116  158  400  296  213  217  208  196  214
12    1   095  086  149  132  115  109  173  174  104  135  212    ?  148  400  296  191  220  198  194  198
            105  090  149  158  115  113  173  180  107  135  216    ?  152  400  302  211  220  208  194  216
13    1   091  092  131  128  115  109  173  180  104  135  210  112  148  400  298  191  220  202  202  198
            091  092  149  160  115  113  173  202  107  135  216  114  152  400  298  197  220  208  202  216
14    1   091  076  137  134  115  109  173  180  104  135  212  112  152  400  296  199  220  208  192  198
            117  092  143  158  115  109  173  202  110  135  212  114  158  400  300  211  220  214  194  216
15    1   079  084  191  128  115  113  173  180  104  135  210  112  148  400  296  197  217  198  192  198
            105  088  191  158  115  113  173  216  104  135  212  112  152  400  300  211  217  198  212  214
16    1   097  082  149  130  115  109  173  180  101  135  210  112  152  400  296  191  217  208  192  198
            103  096  159  158  115  113  173  202  107  135  210  114  158  400  300  197  220  214  202  212
17    1   085  076  121  136  115  109  173  180  101  135  210  112  152  400  300  185  220  160  202  198
            091  092  133  158  115  113  173  216  104  135  212  114  158  400  302  199  220  208  202  212
18    1   089  090  133  158  115  109  173  180  101  135  210  112  152  400  296  185  220  208  196  198
            101  094  143  158  115  109  173  216  104  135  210  114  158  400  302  199  220  208  202  212
19    1   089  090  131  158  115  109  173  180    ?  135  210  112  152  400  296    ?  220  202  192  198
            101  090  155  158  115  109  173  206    ?  135  210  114  158  400  300    ?  220  208  200  212
20    1   077  076  149  128  115  109  173  180  101  135  210  130  152  400  296  199  220  190  192  198
```

Fstat data formats

File Edit Format View Help

```
9 20 400 3
Sare03
Sare04
Sare08
SB24
SB80
SB194
SB196
SB233
SB243
SB265
SB493
W293
W784
gSlMCT011
gSlMCT024
gSlMCT052
WPMS18
PMGC2709
PMGC2889
PMGC21631
1       079091 082090 143147 130158 115115 109115 173173 180180 101104 135135 212216 112114 152158 400400 300300 201211 220226 194194 192202 198216
1       079079 096096 131157 134158 117119 109109 173173 180186 104110 135135 212212 112114 152158 400400 302302 197199 220220 196208 192206 198212
1       091101 082090 161161 116160 115117 113115 173173 176178 104107 135135 212212 112114 152158 400400 302302 185199 220220 202208 188196 198212
1       093107 086096 143147 130134 115115 109109 173173 216216 104104 135135 210212 112114 152158 400400 296296 199201 214220 214224 196202 198204
1       079079 088098 131157 136138 115115 109109 173173 204216 101104 135135 210212 112114 152158 400400 302302 197197 220220 192214 206206 198204
1       091091 082088 131149 134158 115115 113113 173173 198214 104110 135135 212212 112120 152158 400400 296302 203207 220220 208214 202202 198216
1       083093 082086 153191 126158 115115 109115 173173 180180 101107 135135 210212 112114 152158 400400 294296 197213 220220 160208 204204 198198
1       097103 090098 133149 128128 115115 113115 173173 180216 101104 135135 212214 112114 152158 400400 296302 185211 220220 202208 202202 198212
1       091097 090094 127139 128158 115115 109109 173173 180180 104107 135135 212212 114116 152158 400400 296300 197199 220220 190190 194202 198204
1       079097 076090 135149 126158 115115 109113 173173 180180 101104 135135 210216 112114 152158 400400 296298 199199 220220 200208 192202 198212
1       083091 076084 131169 126158 115115 115115 173173 180180 104104 135135 212212 112116 152158 400400 296296 207213 217217 198208 196196 198214
1       095105 086090 149149 132158 115115 109113 173173 174180 104107 135135 212216 000000 148152 400400 296302 191211 220220 198208 194194 198216
1       091091 092092 131149 128160 115115 109113 173173 180202 104107 135135 210216 112114 148152 400400 298298 191197 220220 202208 202202 198216
1       091117 076092 137143 134158 115115 109109 173173 180202 104110 135135 212212 112114 152158 400400 296300 199211 220220 208214 192194 198216
1       079105 084088 191191 128158 115115 113113 173173 180216 104104 135135 210212 112112 148152 400400 296300 197211 217217 198198 192212 198214
1       097103 082096 149159 130158 115115 109113 173173 180202 101107 135135 210210 112114 152158 400400 296300 191197 217220 208214 192202 198212
1       085091 076092 121133 136158 115115 109113 173173 180216 101104 135135 210212 112114 152158 400400 300302 185199 220220 160208 202202 198212
1       089101 090094 133143 158158 115115 109109 173173 180216 101104 135135 210210 112114 152158 400400 296302 185199 220220 208208 196202 198212
1       089101 090090 131155 158158 115115 109109 173173 180206 000000 135135 210210 112114 152158 400400 296300 000000 220220 202208 192200 198212
1       077097 076076 149169 128134 115115 109113 173173 180202 101104 135135 210210 130130 152158 400400 296296 199213 220220 190190 192200 198198
1       087097 082180 131149 128138 115117 109109 173179 180202 101104 135135 210212 114114 152158 286286 296300 191191 220220 190194 192192 198216
1       091105 084180 149149 128158 115117 109109 173173 180202 101104 135135 212212 112114 152158 400400 296300 201213 220220 198208 194202 198216
1       097105 084092 149149 132158 115119 109109 173179 176184 101107 109135 210212 112114 152158 286400 300302 185213 214220 194208 192204 198212
1       101103 084092 131149 134160 115115 109113 173173 216216 101104 135135 210214 112114 150152 400400 298302 185197 220226 198208 190200 198216
1       097105 092180 131149 128160 115115 109109 173173 202210 101104 137139 210212 112114 152158 400400 296296 197199 220226 188208 196202 198216
1       097105 086090 139147 134158 115115 109113 173173 180202 101113 135135 210212 112114 152158 400400 300300 197221 220220 198208 184196 198216
1       091105 090120 131147 158158 115115 109109 173173 216216 101104 135135 210210 112114 152158 400400 296300 189199 220220 208214 180202 198198
1       093105 084090 133147 128158 115115 109109 173173 180180 107110 135135 212212 112112 152158 400400 296302 191211 220226 188188 180202 198216
1       097103 086094 139147 126158 115115 109109 173173 180212 101110 135135 210212 112116 152158 400400 298298 199201 223226 198208 202202 198216
1       105111 094102 139147 128158 115119 109113 173173 180180 101104 135135 210210 108108 152158 400400 300302 189199 223226 198214 182192 198216
1       091093 092098 149149 128158 115115 109115 173173 180202 101113 135135 210210 108108 152158 400400 298298 213217 220220 198214 204204 198216
2       089103 088096 139139 128158 115119 109113 173173 180180 101107 133133 212212 112130 152158 400400 300300 195217 220220 188208 182202 198216
2       103103 096096 143143 134158 119119 109113 173173 212202 104104 135135 208212 112114 152158 400400 300302 199199 220220 194208 192192 198216
2       089095 088098 139139 134158 115115 109109 173173 180212 101104 135135 212216 112114 152158 400400 296298 185199 220220 202208 192204 198212
2       103103 082088 153155 128158 115115 113113 173173 180202 101104 135135 210212 112114 152158 286400 300300 199205 220220 188188 192212 198216
2       103111 082088 131147 128158 115119 109113 173173 180180 101107 135135 210212 112114 150152 400400 294296 213213 220220 208208 192202 198216
2       097103 088094 131161 126158 115115 109113 173173 212202 101107 135135 212212 112118 152158 400400 300302 197213 220220 188208 182192 198216
2       091097 082096 149191 134158 115115 109113 173173 216216 101104 135135 210210 112114 152158 400400 296296 197203 214223 190190 182194 198216
2       077079 082088 131147 128158 115119 109109 173173 214216 101107 135135 212216 112122 148158 400400 296298 199221 220220 202208 182194 198204
2       077095 086098 147155 130134 115115 109113 173173 180212 107110 135135 210212 112114 150158 400400 296298 199201 220220 194208 192202 198204
```

Structure data format

File Edit Format View Help

		Sare03	Sare04	Sare08	SB24	SB80	SB194	SB196	SB233	SB243	SB265	SB493	W293	W784	gSlMCT011	gSlMCT024	gS
G41	1	079	082	143	130	115	109	173	180	101	135	212	112	152	400	300	201	220	194	192	198
G41	1	091	090	147	158	115	115	173	180	104	135	216	114	158	400	300	211	226	194	202	216
G42	1	079	096	131	134	117	109	173	180	104	135	212	112	152	400	302	197	220	196	192	198
G42	1	079	096	157	158	119	109	173	186	110	135	212	114	158	400	302	199	220	208	206	212
G43	1	091	082	161	116	115	113	173	176	104	135	212	112	152	400	302	185	220	202	188	198
G43	1	101	090	161	160	117	115	173	178	107	135	212	114	158	400	302	199	220	208	196	212
G44	1	093	086	143	130	115	109	173	216	104	135	210	112	152	400	296	199	214	214	196	198
G44	1	107	096	147	134	115	109	173	216	104	135	212	114	158	400	296	201	220	224	202	204
G45	1	079	088	131	136	115	109	173	204	101	135	210	112	152	400	302	197	220	192	206	198
G45	1	079	098	157	138	115	109	173	216	104	135	212	114	158	400	302	197	220	214	206	204
G46	1	091	082	131	134	115	113	173	198	104	135	212	112	152	400	296	203	220	208	202	198
G46	1	091	088	149	158	115	113	173	214	110	135	212	120	158	400	302	207	220	214	202	216
G47	1	083	082	153	126	115	109	173	180	101	135	210	112	152	400	294	197	220	160	204	198
G47	1	093	086	191	158	115	115	173	180	107	135	212	114	158	400	296	213	220	208	204	198
G48	1	097	090	133	128	115	113	173	180	101	135	212	112	152	400	296	185	220	202	202	198
G48	1	103	098	149	128	115	115	173	216	104	135	214	114	158	400	302	211	220	208	202	212
G410	1	091	090	127	128	115	109	173	180	104	135	212	114	152	400	296	197	220	190	194	198
G410	1	097	094	139	158	115	109	173	180	107	135	212	116	158	400	300	199	220	190	202	204
G411	1	079	076	135	126	115	109	173	180	101	135	210	112	152	400	296	199	220	200	192	198
G411	1	097	090	149	158	115	113	173	180	104	135	216	114	158	400	298	199	220	208	202	212
G412	1	083	076	131	126	115	115	173	180	104	135	212	112	152	400	296	207	217	198	196	198
G412	1	091	084	169	158	115	115	173	180	104	135	212	116	158	400	296	213	217	208	196	214
G413	1	095	086	149	132	115	109	173	174	104	135	212	-9	148	400	296	191	220	198	194	198
G413	1	105	090	149	158	115	113	173	180	107	135	216	-9	152	400	302	211	220	208	194	216
G414	1	091	092	131	128	115	109	173	180	104	135	210	112	148	400	298	191	220	202	202	198
G414	1	091	092	149	160	115	113	173	202	107	135	216	114	152	400	298	197	220	208	202	216
G415	1	091	076	137	134	115	109	173	180	104	135	212	112	152	400	296	199	220	208	192	198
G415	1	117	092	143	158	115	109	173	202	110	135	212	114	158	400	300	211	220	214	194	216
G416	1	079	084	191	128	115	113	173	180	104	135	210	112	148	400	296	197	217	198	192	198
G416	1	105	088	191	158	115	113	173	216	104	135	212	112	152	400	300	211	217	198	212	214
G417	1	097	082	149	130	115	109	173	180	101	135	210	112	152	400	296	191	217	208	192	198
G417	1	103	096	159	158	115	113	173	202	107	135	210	114	158	400	300	197	220	214	202	212
G418	1	085	076	121	136	115	109	173	180	101	135	210	112	152	400	300	185	220	160	202	198
G418	1	091	092	133	158	115	113	173	216	104	135	212	114	158	400	302	199	220	208	202	212
G419	1	089	090	133	158	115	109	173	180	101	135	210	112	152	400	296	185	220	208	196	198
G419	1	101	094	143	158	115	109	173	216	104	135	210	114	158	400	302	199	220	208	202	212
G420	1	089	090	131	158	115	109	173	180	-9	135	210	112	152	400	296	-9	220	202	192	198
G420	1	101	090	155	158	115	109	173	206	-9	135	210	114	158	400	300	-9	220	208	200	212
G421	1	077	076	149	128	115	109	173	180	101	135	210	130	152	400	296	199	220	190	192	198
G421	1	097	076	169	134	115	113	173	202	104	135	210	130	158	400	296	213	220	190	200	198
G422	1	087	082	131	128	115	109	173	180	101	135	210	114	152	286	296	191	220	190	192	198
G422	1	097	180	149	138	117	109	179	202	104	135	212	114	158	286	300	191	220	194	192	216
G423	1	091	084	149	128	115	109	173	180	101	135	212	112	152	400	296	201	220	198	194	198
G423	1	105	180	149	158	117	109	173	202	104	135	212	114	158	400	300	213	220	208	202	216
G424	1	097	084	149	132	115	109	173	176	101	109	210	112	152	286	300	185	214	194	192	198
G424	1	105	092	149	158	119	109	179	184	107	135	212	114	158	400	302	213	220	208	204	212
G425	1	101	084	131	134	115	109	173	216	101	135	210	112	150	400	298	185	220	198	190	198
G425	1	103	092	149	160	115	113	173	216	104	135	214	114	152	400	302	197	226	208	200	216
G426	1	097	092	131	128	115	109	173	202	101	137	210	112	152	400	296	197	220	188	196	198
G426	1	105	180	149	160	115	109	173	210	104	139	212	114	158	400	296	199	226	208	202	216
G427	1	097	086	139	134	115	109	173	180	101	135	210	112	152	400	300	197	220	198	184	198
G427	1	105	090	147	158	115	113	173	202	113	135	212	114	158	400	300	221	220	208	196	216
G428	1	091	090	131	158	115	109	173	216	101	135	210	112	152	400	296	189	220	208	180	198
G428	1	105	120	147	158	115	109	173	216	104	135	210	114	158	400	300	199	220	214	202	198
G429	1	093	084	133	128	115	109	173	180	107	135	212	112	152	400	296	191	220	188	180	198
G429	1	105	090	147	158	115	109	173	180	110	135	212	112	158	400	302	211	226	188	202	216
G430	1	097	086	139	126	115	109	173	180	101	135	210	112	152	400	298	199	223	198	202	198
G430	1	103	094	147	158	115	109	173	212	110	135	212	116	158	400	298	201	226	208	202	216
G431	1	105	094	139	128	115	109	173	180	101	135	210	108	152	400	300	189	223	198	182	198
G431	1	111	102	147	158	119	113	173	180	104	135	210	108	158	400	302	199	226	214	192	216

APPENDIX F

A SUMMARY OF THE STATSTICS USED IN THESIS

The statistics used in the MICRO-CHECKER

Brookfield (1996) Null Allele Estimator 1

Assuming there are no null allele homozygotes or ignoring all non-amplified samples as degraded DNA, human error, etc., the Brookfield (1996) estimate of the null allele frequency is given by:

r = (He - Ho) / (1 + He)

The Confidence Interval

Bonferroni (Dunn-Sidak) adjusted 95% confidence interval is used.The values for a randomised locus are placed in an ordered list. If, for example, the 95% Confidence Interval is selected, the top 2.5% of the values and the bottom 2.5% of the values are removed from the list. The highest and lowest values in the remaining list provide the range of the Confidence Interval.

The Bonferroni (Dunn-Sidák) adjusted Confidence Interval is based on the formula:

$$a' = 1 - (1 - a)^{1/k}$$

where is 0.05 (the complement of 0.95 or 95%) and k is the number of comparisons (Sokal & Rohlf, 1995).

The statistics used in GenAlEx 6.5

Number of different alleles (Na)

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

Effective number of alleles (Ne)

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

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

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

No. of private alleles

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

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.

Expected heterozygosity (He)

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

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

Expected Heterozygosity averaged across populations (Mean He)

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

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

Observed Heterozygosity (Ho)

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

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

Observed heterozygosity, averaged across populations (Mean Ho)

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

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

Fixation Index (F)

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

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

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;

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

F statistics

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

Individual Heterozygosity (H_I)

H-indiv = the proportion of loci that are heterozygous across an individual, where nH is the number of heterozygous loci, and nL is the number of loci. When compared across individuals H-indiv can offer important clues about the amount and distribution of inbreeding in populations.

$$H_I = \frac{nH}{nL}$$

Average within Population heterozygosity (Hs)

Identical to the mean He, being the average of the within population expected heterozygosity across populations.

$$H_e = 1 - \sum p_i^2 \; ; \quad H_S = \overline{H_E} = \frac{\sum H_{Es}}{k}$$

Total expected heterozygosity (Ht)

Ht is the expected heterozygosity if all populations were pooled (no subdivision). Calculated as 1 minus the sum of the average allele frequencies over populations.

$$H_T = 1 - \sum_{i=1}^{h} \overline{p}_i^2$$

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

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

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

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

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

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

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

The genetic structure of populations can be estimated by the formula

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

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

Number of Migrants (Nm)

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

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

Nei's Genetic Distance

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

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

Nei's Genetic Identity

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

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

Standard Error

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

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

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

Garza-Williamson index (G-W)

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 G

DAMS AND HYDROELECTRIC POWER PLANT ON THE GÖKSU RIVER

There are 5 hydroelectric power plants with 86 MW installed power on Göksu River. These hydroelectric power plants with a total installed power of 86 MW will meet 0,422 percent of electricity generated from HPP in Turkey and 0,113 percent of total electricity consumption. Dam and / or HPP located on the Göksu River are as follows;

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

DAMS AND HYDROELECTRIC POWER PLANT ON THE KIZILIRMAK RIVER

There are 15 dams and / or hydroelectric power plants on Kızılırmak. With these hydroelectric power plants having a total installed capacity of 2.085 MW, the electricity generated from hydroelectric power plants in Turkey is 5,277 percent and the electricity consumption is 1,419 percent. Dam and / or HPP located on Kızılırmak are as follows;

No:	Power plant name	Province / District	Maximum Operating Level	Minimum Tail Water Elevation	Installed Power	Construction Year
1	**Çermikler Dam and HPP**	Sivas, ğarkıÇla	1.194 m		25 MW	2013
2	**Yamula Dam** and HPP	Kayseri, Kocasinan	1.100 m		100 MW	2002
3	**Bayramhacılı Dam and HPP**	NevÇehir, Avanos	980 m		47 MW	2008
4	**Sarıhıdır HPP**	NevÇehir, Ürgüp	938 m		6 MW	2007
5	**Cemel HPP**	NevÇehir, Avanos	922 m	902 m	20 MW	2011
6	**Hirfanlı Dam** and HPP	KırÇehir	851 m		128 MW	1953
7	**Kesikköprü Dam**	Ankara, Bala	786 m		76 MW	1959
8	**Kapulukaya Dam** and HPP	Kırıkkale	724 m		54 MW	1979
9	**Ülkün HPP** (under construction)	Çorum, Uğurludağ	534 m	515 m	(24) MW	-
10	**Obruk Dam** and HPP	Çorum	510 m		211 MW	1996
11	**Pirinçli HPP**	Çorum, Dodurga	444 m		19 MW	2010
12	**Kargı Kızılırmak Dam and HPP**	Çorum	405 m	342 m	102 MW	2011
13	**Boyabat Dam** and HPP	Sinop, Durağan	330 m		513 MW	2011
14	**Altınkaya Dam** and HPP	Samsun, Bafra	190 m		703 MW	1980
15	**Derbent Dam** and HPP	Samsun, Bafra	58 m		56 MW	1984
16	**Saraçbendi HPP**	Sivas, Gemerek			25 MW	2009

APPENDIX H

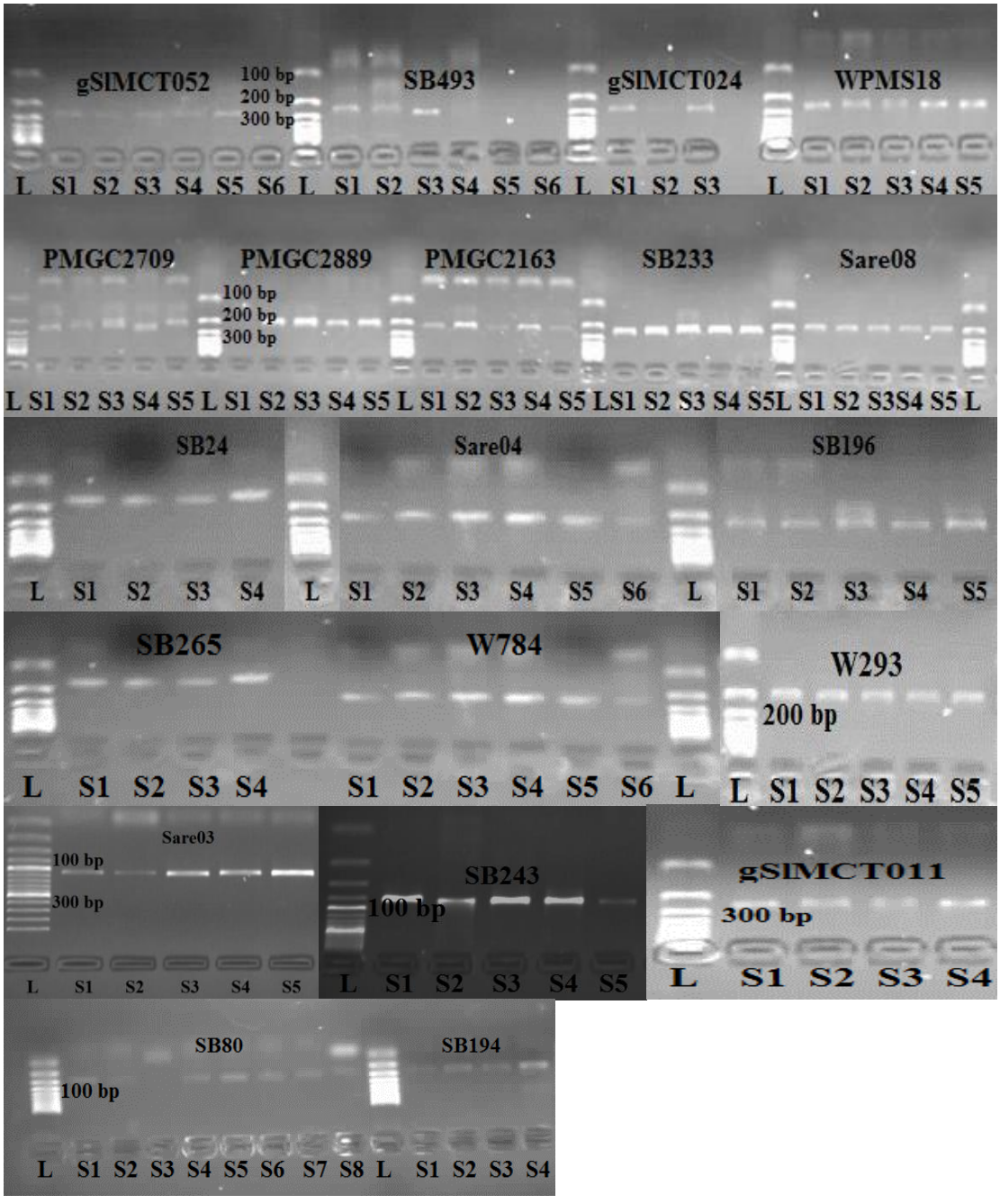
SUMMARY OF PRIVATE ALLELES BY POPULATION

Pop	Locus	Allele	Freq
GKSUPSTREAMPOP	Sare03	117	0,016
GKSUPSTREAMPOP	Sare04	84	0,097
GKSUPSTREAMPOP	Sare04	180	0,048
GKSUPSTREAMPOP	Sare08	127	0,016
GKSUPSTREAMPOP	Sare08	137	0,016
GKSUPSTREAMPOP	SB24	116	0,016
GKSUPSTREAMPOP	SB233	176	0,032
GKSUPSTREAMPOP	SB233	184	0,016
GKSUPSTREAMPOP	SB233	210	0,016
GKSUPSTREAMPOP	SB265	139	0,016
GKSUPSTREAMPOP	gSlMCT052	191	0,100
GKSUPSTREAMPOP	PMGC2709	224	0,016
GKSUPSTREAMPOP	PMGC2889	190	0,016
GKSUPSTREAMPOP	PMGC2889	200	0,048
GKSERMENEKPOP	Sare03	115	0,043
GKSERMENEKPOP	Sare08	163	0,022
GKSERMENEKPOP	SB493	208	0,022
GKSERMENEKPOP	W293	122	0,022
GKSERMENEKPOP	gSlMCT052	205	0,065
GKSERMENEKPOP	PMGC2709	218	0,043
GKSERMENEKPOP	PMGC2889	208	0,022
GKSERMENEKPOP	PMGC2163	208	0,022
GKSMIDDLEPOP	SB233	208	0,018
GKSMIDDLEPOP	gSlMCT052	231	0,018
GKSMIDDLEPOP	gSlMCT052	271	0,018
GKSMIDDLEPOP	gSlMCT052	281	0,018
GKSMIDDLEPOP	PMGC2709	172	0,054
GKSMIDDLEPOP	PMGC2889	174	0,018
GKSMIDDLEPOP	PMGC2163	206	0,018
GKSDOWNSTREAM	Sare03	113	0,017
GKSDOWNSTREAM	Sare04	104	0,017
GKSDOWNSTREAM	SB24	114	0,017
GKSDOWNSTREAM	SB233	170	0,034
GKSDOWNSTREAM	SB233	200	0,017
GKSDOWNSTREAM	gSlMCT052	227	0,017
GKSDOWNSTREAM	PMGC2709	166	0,017
GKSDOWNSTREAM	PMGC2709	210	0,033
GKSDOWNSTREAM	PMGC2163	218	0,033

Pop	Locus	Allele	Freq
KIZUPSTREAMPOP	Sare03	83	0,015
KIZUPSTREAMPOP	Sare04	114	0,061
KIZUPSTREAMPOP	SB24	118	0,015
KIZUPSTREAMPOP	SB233	176	0,015
KIZUPSTREAMPOP	SB233	198	0,030
KIZUPSTREAMPOP	SB493	226	0,015
KIZUPSTREAMPOP	gSlMCT052	243	0,016
KIZUPSTREAMPOP	PMGC2709	186	0,015
KIZUPSTREAMPOP	PMGC2709	214	0,015
KIZUPSTREAMPOP	PMGC2889	200	0,015
KIZUPSTREAMPOP	PMGC2163	188	0,030
KIZMIDDLEPOPKIR	Sare03	117	0,031
KIZMIDDLEPOPKIR	SB196	167	0,016
KIZMIDDLEPOPKIR	SB233	206	0,063
KIZMIDDLEPOPKIR	gSlMCT024	294	0,033
KIZMIDDLEPOPKIR	gSlMCT052	217	0,016
KIZMIDDLEPOPKIR	WPMS18	211	0,031
KIZMIDDLEPOPKIR	WPMS18	214	0,016
KIZMIDDLEPOPKIR	PMGC2709	198	0,016
KIZMIDDLEPOPKIR	PMGC2889	178	0,016
KIZMIDDLEPOPKRK	Sare04	82	0,010
KIZMIDDLEPOPKRK	Sare04	112	0,010
KIZMIDDLEPOPKRK	Sare08	169	0,010
KIZMIDDLEPOPKRK	SB24	116	0,010
KIZMIDDLEPOPKRK	SB233	184	0,010
KIZMIDDLEPOPKRK	SB265	109	0,019
KIZMIDDLEPOPKRK	SB493	242	0,010
KIZMIDDLEPOPKRK	W293	130	0,038
KIZMIDDLEPOPKRK	W784	168	0,010
KIZMIDDLEPOPKRK	gSlMCT052	191	0,031
KIZMIDDLEPOPKRK	PMGC2709	166	0,010
KIZMIDDLEPOPKRK	PMGC2709	180	0,010
KIZMIDDLEPOPKRK	PMGC2709	184	0,010
KIZMIDDLEPOPKRK	PMGC2889	208	0,010
KIZMIDDLEPOPKRK	PMGC2889	210	0,010
KIZMIDDLEPOPKRK	PMGC2163	186	0,048
KIZMIDDLEPOPKRK	PMGC2163	200	0,029
KIZMIDDLEPOPKRK	PMGC2163	208	0,010
KIZDOWNSTREAMPOPCOR	Sare08	133	0,067
KIZDOWNSTREAMPOPBAF	Sare03	115	0,100
KIZDOWNSTREAMPOPBAF	Sare04	126	0,067
KIZDOWNSTREAMPOPBAF	Sare04	130	0,067
KIZDOWNSTREAMPOPBAF	SB196	170	0,033
KIZDOWNSTREAMPOPBAF	gSlMCT052	213	0,067
KIZDOWNSTREAMPOPBAF	WPMS18	202	0,067
KIZDOWNSTREAMPOPBAF	PMGC2889	222	0,067

APPENDIX I

PCR PRODUCTS OF 20 SSR MARKERS RUN IN 3 % AGAROSE GEL



CURRICULUM VITAE

General Information

Date of Issue	:
ID Number	: 24007214628
Name Surname	: Funda Özdemir Değrimenci
Mail Address	: METU, Department of Biology, Lab 251, Ankara
Birth Date and Place	: 25.03.1983 Antalya
Phone	: 0 312 210 51 60 GSM: 0 505 376 16 13
e-mail	: ozfunda@metu.edu.tr

Education

Period of Study	Degree	University	Department
2011/2017	Doctorate	METU	Biology
2008/2011	Master	METU	Biology
2002/2007	Bachelor	Gazi University	Biology Education

Academic and Professional Experience

Period	Position	Location	Department
2008-2017	Research and Teaching Assistant	METU	Biology