MOLECULAR PHYLOGENETICS OF TURKISH *SALIX* L. SPECIES

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

BY

PELİN ACAR

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

MAY 2017

Approval of the Thesis:

MOLECULAR PHYLOGENETICS OF TURKISH *SALIX* L. SPECIES

Submitted by **PELİN ACAR** 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, **Biology Dept., METU**

Examining Committee Members:

Prof.Dr. Musa Doğan _____
Biology Dept., METU

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

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

Prof. Dr. Emine Sümer Aras _____
Biology Dept., Ankara University

Prof. Dr. İrfan Kandemir _____
Biology Dept., Ankara University

Date:05/05/2017

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 : Pelin Acar

Signature :

ABSTRACT

MOLECULAR PHYLOGENETICS OF TURKISH *SALIX* L. SPECIES

ACAR, Pelin
Ph.D., Department of Biological Sciences
Supervisor: Prof. Dr. Zeki KAYA
May 2017, 123 Pages

Chloroplast (*trn*T-F, *mat*K and *rbc*L) and nuclear genome (*ITS*) regions were used to explore the evolutionary relationships of *Salix* species which are native to Turkey. Morphological analysis with the utilization of Principal Component Analysis (PCA) was also carried out.

Such a combined sequence data from cpDNA and nrDNA genes enabled reliable subgenus-level classification of Turkish willows (subgenus *Salix* and *Vetrix*). Morphological and genetic results agree with traditional taxonomic concepts in clustering of two subgenera. Two coding cpDNA gene regions (*mat*K and *rbc*L) are found as conserved whereas non-coding cpDNA (*trn* T-F) and nrDNA (*ITS)* evolved rapidly for Turkish *Salix.*

Turkish willow species appear to be a monophyletic (cpDNA) or potentially monophyletic (nrDNA) with having two well supported clades. Biogeographically, subg. *Vetrix* species are located in high altitude and cool climate whereas subg. *Salix* species are in warmer climate. *S. acmophylla* is always located at distant positions of subg. *Salix* clade. The appearance of the subg. *Vetrix* members (*S. rizeensis* and *S. amplexicaulis*) in subg. *Salix* clade can be explained by occurrence of natural hybrids with species from subg. *Salix* in shared habitat.

New World (America) and Old World (Asia) *Salix* species are scattered throughout Turkish subgenera in all phylogenetic trees. The close molecular relations between subg. *Salix*-OWS and subg. *Vetrix*-NWS were determined by taxonomical relations, rather than geographical distribution. Turkish subgenera were diverged from others in late Pliocene. Chloroplast introgression and incomplete lineage sorting events in *Salix* species of Turkey may explained the high haplotype diversity in the genus.

Key words: *Salix*, Phylogeny, *trn*T-F, *mat*K, *rbc*L, *ITS*

ÖZ

TÜRKİYE *SALIX* L. TÜRLERİNİN MOLEKÜLER FİLOGENİSİ

ACAR, Pelin
Doktora, Biyolojik Bilimler
Tez Yöneticisi: Prof. Dr. Zeki KAYA
Mayıs 2017, 123 sayfa

Türkiye'de doğal olarak yayılış gösteren *Salix* cinsine ait türlerde, evrimsel ilişkileri ortaya çıkarmak için kloroplast (*trn*T-F, *mat*K and *rbc*L) ve nükleer genom bölgeleri (*ITS*) kullanılmıştır. Morfolojik analizler ise Temel Bileşenler Analizi (PCA) ile gerçekleştirilmiştir.

Türkiye'deki *Salix* L. türlerine ait kloroplast ve nükleer gen bölgeleri dizi verileri, alt cins seviyesinde (altcins *Salix* ve *Vetrix*) sınıflandırma yapabilmektedir. Morfolojik ve genetik verilere göre Türkiye söğütlerinin iki alt cinste gruplandığı ve geleneksel taksonomik kavramı ile uyumlu olduğu görülmüştür. Kloroplast bölgesinin iki kodlanan geni (*mat*K ve *rbc*L) korunmuş iken, kloroplastta kodlanmayan (*trn*T-F) ve çekirdek (*ITS)* genlerinin evrimleşme hızının Türkiye *Salix* türleri için yüksek olduğu gözlenmiştir.

Türkiye söğüt türleri çok iyi desteklenen iki grup (altcins *Salix* ve *Vetrix*) kloroplast DNA açısından monofiletik veya çekirdek DNA'ya göre potensiyel olarak monofiletiktir. Biyocoğrafik olarak altcins *Vetrix* türleri yüksek rakım ve serin iklimlerde bulunurken, altcins *Salix* türleri daha ılıman iklimlere uyum göstermektedir. *S.acmophylla* genelde altcins *Salix* grubu üyelerinden uzak konumlanmaktadır. Altcins *Vetrix* grubuna dahil olan *S. rizeensis* ve

S. amplexicaulis türlerinin altcins *Salix*'te yer alması, altcins *Salix*'in diğer üyeleri ile ortak yaşam alanlarında doğal melez yapması ile açıklanabilir.

Yeni Dünya (Amerika kıtası) ve Eski Dünya (Asya kıtası) *Salix* türleri, oluşturulan bütün filogenetik ağaçlarda Türkiye türleri arasında dağılım göstermişlerdir. Altcins *Salix*-Eski Dünya ve altcins *Vetrix*-Yeni Dünya arasındaki yakın moleküler ilişki, coğrafik dağılımdan çok taksonomik ilişki ile belirlenmektedir. Türkiye *Salix* altcinslerinin diğerlerinden yakın zamanda (geç Pliosen) farklılaştığı düşünülmektedir. Türkiye'ye özgü *Salix* türlerinde gözlenen yüksek haplotip çeşitliliği, kloroplast introgresyonu ve türleşmenin henüz tamamlanmamış olması ile açıklanabilir.

Anahtar kelimeler: *Salix,* Filogeni, *trn*T-F, *mat*K*, rbcL, ITS*

To my big family…

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude and sincerest appreciation to my supervisor Prof. Dr. Zeki Kaya for his guidance, continuous advice, invaluable help and understanding throughout this study. I also owe my appreciations to Prof. Dr. Sertaç Önde for his encouragements.

I would like to thank to Prof. Dr. Emine Sümer Aras for her help and criticism throughout the study. I also wish to thank to Prof. Dr. Hayri DUMAN from Gazi University and Prof. Dr. Meral Avcı from Istanbul University for their kindness and supports. All supports provided by Scientific and Technical Research Council of Turkey (TUBITAK) 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 supported by Middle East Technical University (METU) BAP-01-08-2012-013 Project "Türkiye Söğüt Türlerinin Moleküler Filogenetiği".

I want to thank my exclusive coffee friends; Bircan Taşkıran and M. Alev Ateş, my project and field trip partners; Funda Özdemir Değirmenci, Çiğdem Kansu and Asiye Çiftçi, my youngest labmates; Sıla Aygün, Yasin Tokdemir and Burak Yelmen, my sweet roommate; Fatma Küçük Baloğlu, and all other my colleagues from Department of Biological Sciences, Plant Genetics and Tissue Culture Laboratory.

I would like to give my special thanks to my mother Nuray Keske, my father M. Münür Keske, my brother Pamir Keske and my lovely uncle Ali Sözkesen for their love, support and patience over the years.

Finally, I am grateful to my endless love; Çağrı Acar for his great support.

This thesis is dedicated to my cute son.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

AMOVA	Analysis of Molecular Variance
ANK	Ankara University Herbarium
cpDNA	Chloroplast DNA
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraaceticacid disodium salt
ETS	Expressed Sequence Tags
***F*st**	Fixation Index
GAZI	Gazi University Herbarium
HCl	Hydrogen chloride
HUB	Hacettepe University Herbarium
IAA	Isoamyl alcohol
IGS	Intergenic Spacer
ISTO	Istanbul University Forest Department Herbarium
ITS	Internal Transcribed Spacer Region
LCS	Large inverted repeat
***mat*K**	The maturase Kinase
MEGA	Molecular Evolutionary Genetic Analysis
NaCl	Sodium chloride
NCBI	National Center for Biotechnology Information
NJ	Neighbour-joining
nrDNA	Nuclear Ribosomal DNA
nrRNA	Nuclear Ribosomal RNA
NTSYS	Numerical Taxonomy System

NWS	New World *Salix*
ORF	Open Reading Frame
OWS	Old World *Salix*
PCR	Polymerase Chain Reaction
***rbc*L**	Large subunit of Rubisco
rDNA	Ribosomal DNA
RNA	Ribonucleic Acid
Rpm	Revolutions per minute
SSC	Small single-copy region
Subg.	Subgenus
TBE	Tris-Borate-EDTA
TE	Tris EDTA
TOVAG	Tarım Ormancılık ve Veterinerlik Araştırma Grubu
tRNA	Transfer RNA
TÜBİTAK	The Scientific and Technological Research Council of Turkey

CHAPTER 1

INTRODUCTION

The biodiversity and characteristics of natural communities in a country can be measured by number of native species, distribution of these and also their vegetation types. Among the European countries, Turkey can easily be considered as a key country for global biodiversity regarding these criteria.

Anatolia, in other words Asia Minor, lies in the nexus of three phytogeographical regions, Euro-Siberian, Mediterranean and Irano-Turanian regions (Zohary, 1973). Anatolia, the bridge between two continents, has various climatic and geographical conditions. Geographic barriers within Anatolia especially the Anatolian Diagonal, serving as a barrier cause differentiation of taxa at the species and subspecies level (Bilgin, 2011). Due to its location, Anatolia's biogeography ranges from wetlands to marine ecosystem, from plains to mountains, from valleys to steppes resulting in high amount of species diversity. The level of endemism is considered to be quite high with the ratio; 34% (National report of Convention of Biological Diversity, 2009).

The first floristic studies exercised by French researcher Tournefort in North and Northeast Anatolia were pursued by E.Boissier (*Flora Orientalis*) and P.H. Davis (*Flora of Turkey and East Aegean Islands*) in Turkey. The 7th volume of *Flora of Turkey* covers the Salicaceae family with taxonomic key written by Skvortsov and Edmondson (1970). Although *Flora of Turkey* is still the best source of floristic composition of Turkey, some of the data presented need to be revised with the result from further studies.

In Turkey, most forested areas are located in the Black Sea region followed by the Mediterranean and Aegean geographic regions. 26.8% of the total surface area of Turkey is covered by forests (Kaya and Raynal, 2001). One of the most significant and crowded member of the Anatolian dendroflora is Salicaceae species (Skvortsov, 1999).

The taxonomy of *Salix* genus is still a debate among the authorities not only in Turkey, also throughout the world. Linnaeus (1753) claims that; "Species of this genus are extremely difficult to clarify". High morphological variability in willows has been resulted in taxonomic confusion (Argus 1986; Cronquist 1964; Dorn 1975). The traditional methods for identifying the members of Salicaceae species by using morphological traits are not sufficient to classify them due to hybridization, reproductive isolation, and polyploidy in the family.

1.1. The Taxonomic History of Salicaceae: Genus *Salix*

The Salicaceae family is comprised of two genus *Populus* and *Salix* (Cronquist, 1981) though Skvortsov (1999) segregated Chosenia as another genus in Salicaceae. The complex genus *Salix* is in the family Salicaceae in the major group Angiosperms (Table 1.1). The number of species *Salix,* is estimated as 350 (Skvortsov, 1968; 1999), 526 (Fang, 1987), 450 (Argus, 1997) or 500 (Hardig, 2014). Today, it is accepted that the genus *Salix* L. is represented with more or less 500 species (Wu *et al*., 2015). They are extensively found in northern hemisphere and abundantly in China and former Soviet Union (Argus, 1997). 27 of them are naturally found in Turkey (Terzioğlu *et al*., 2014).

Table 1.1. Scientific classification of *Salix* genus

Kingdom	***Plantae***
Division	Magnoliophyta
Class	Magnoliopsida
Subclass	Dilleniidae
Order	Salicales
Family	Salicaceae
Genus	*Salix*

Salicologists have been trying to classify *Salix* into natural groups more than a hundred years. *Salix* classification has always been complicated in literature. Firstly, Nakai (1920) described *Chosenia* as a new genus based on *Salix splendida* because of its floral and bud differences. Kimura (1928) segregated a new genus *Toisusu* including *Salix* species with different ament and bud style. *Salix* was divided into 35 genera by Argus (1974) at first. Consequently, many molecular phylogenetic studies suggested that *Salix* is new monophyletic. Natural group is not necessary to segregate into more subclasses (Leskinen and Alström-Rapaport, 1999). Argus (2000) treated *Chosenia* and *Toisusu* as genera. However Skvortsov (1968; 1999) categorized them as a subgenus since they have only a few different features. In Skvortsov system (1968; 1999) which is widely accepted, *Salix* includes three subgenera; *Salix*, *Chamaetia*, and *Vetrix*. Thus, taxonomist has disagreed about the generic and subgeneric level of *Salix* species.

The first 23 *Salix* species were introduced to World Flora by C. Linnaeus (1753) and this number increased to 31 in his lifetime. In the late 18th century, Schleicher (1807) published a detailed botanical monograph on the willows of Switzerland with 120 species. W. Koch (1828) categorized the European willows with underlying the variability of willows. Wimmer (1866) made clear the identification of European species by the publication, '*Salices Europaeae*'. The monograph on describing new varieties, forms, and hybrids in the European *Salix* was introduced after the publication of '*Synopsis*' by Ascherson and Graebner in 19th century according to Skvortsov (1999).

Worldwide classification of *Salix* genus is accepted after the results of Andersson (1868) at first. Schneider (1916) proceeded to develop worldwide traditional classification schemes on *Salix* species and varieties of section Longifolia. New World *Salix* species were categorized by Argus (1997) with the help of numerical taxonomic methods. One of the other salicologist in *Salix* history is Skvortsov (1968; 1999) having publications on European, Asian, Middle Eastern specimens of *Salix*. Skvortsov and Edmonson revised the Turkish *Salix* species in Flora of Turkey and East Aegean Islands (Davis, 1965). Molecular systematic work involving cpDNA RFLPs of *Salix* species phylogeny in *Salix* sect. Longifoliae Pax was started by Brunsfeld (Brunsfeld *et al*., 1991). DNA based studies for *Salix* genus continued with

Leskinen and Alström-Rapaport (1999). Many molecular works are restricted to sections, Azuma *et al*., (2000) defined the entire *Salix* genus by using macromolecular characters. Other attributions on molecular based phylogenetic analyses of Salicaceae family were done by Hardig *et al*. (2010) which showed the harmony between molecular data (*ITS* and *mat*K) and Argus (1997) study on morphological classification of willows. Recently, two new surveys about evolutionary implication by molecular markers are performed by Wu *et al*. (2015) with study of the new and Old World lineages of subg. *Salix* and Lauron-Moreau *et al.* (2015) with phylogenetic studies on American willows.

The sections in *Salix* genus in Turkey (Skvortsov, 1968; 1999) are listed as follows; Subgenus *Salix*; *Humboldtianae, Amygdalinae, Urbanianae, Pentandrae, Salix, Subalbae, Subgenus Chamaetia; Chamaetia, Retusae, Myrtilloides, Glaucae, Myrtosalix,* Subgenus *Vetrix; Hastatae, Glabrella, Nigricantes, Vetrix, Arbuscella, Vimen, Subviminales, Canae, Villosae, Lanatae, Daphnella, Incubaceae, Flavidae, Helix, Cheilophilae*. The distribution of subgenera of *Salix* L. including *Salix* and *Vetrix* in the world is illustrated in Figure 1.1.

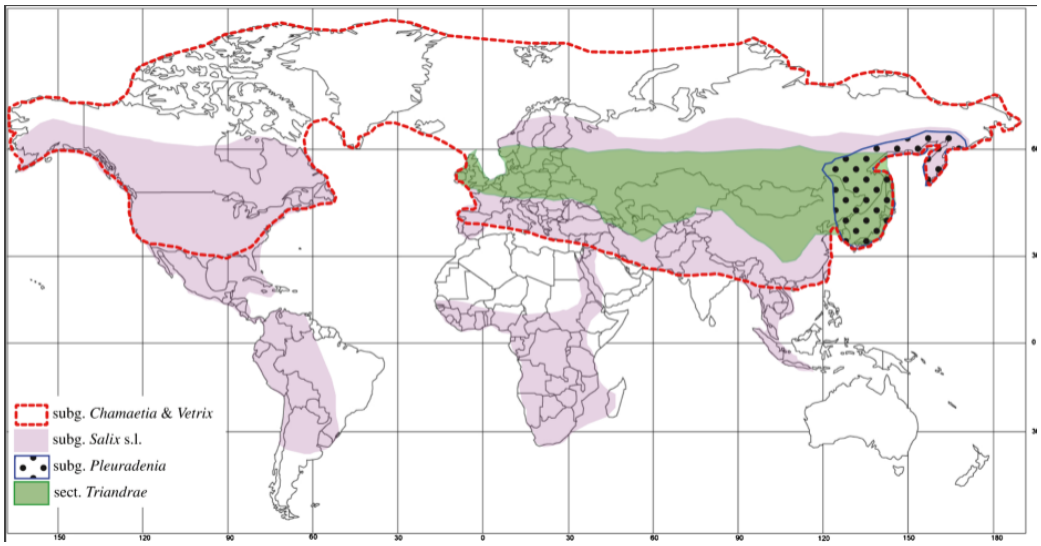

Figure 1.1. The distribution of subgenera belonging to *Salix* L. genus (Wu *et al*., 2015)

1.2. The Ecology and the Morphology of *Salix* genus

In Salicaceae family, two genus, *Salix* and *Populus* have many common characteristics as far as ecology is concerned. However, *Salix* are more adaptable genus when it is compared with the *Populus*. Their habitats vary in much wider ranges. The members of the *Salix* L. genus prefer habitats with more water and higher light conditions (Figure 1.2). Species of the genus can survive from arid areas to wetlands, from beaches to high mountains. This adaptable characeristic of the genus comes from minute features of the seeds. Since they are very small and numerous, they are dispersed by the wind and germinate easily (Skvortsov, 1999).

Salix are deciduous woody plants and their size changes from long trees to miniature shrubs. Leaves are usually alternate and varied from linear to obovate, from entire to minutely glandular-dentate (Figure 1.3). Axillary bud scale in willow species are with single outer scale (Davis, 1965).

Generally many data for systematics of angiosperms are derived from the characteristics of the flowers'. However this type of floral characteristics is absent in *Salix* genus as a result of reduced flowers (Azuma *et al*., 2000). Unisexual willow flowers, called catkins, appear before or with leaves. The number of the stamens of willows are from two to twelve, ovary has usually stalk (Figure 1.4). The pollen has sticky substance and also pollination is performed by insects (Skvortsov, 1999). Palynological investigation of *Salix* species has been another interesting topic for many years (Sohma, 1993; Qureshi *et al*., 2007; Babayi *et al*., 2012).

Figure 1.2. *Salix* sp. Habitat with Higher Light and More Water conditions (Photographed by Pelin Acar)

Figure 1.3. Alternately Arranged Leaves of *Salix cinerea* (Photographed by Pelin Acar)

Figure 1.4. *Salix alba* Male (2 stamens) and Female Catkins (Photographed by Pelin Acar)

According to Argus (2009), there are some criteria what make morphological identification a problem in *Salix* genus;

(1) Because thet are dioecious plants, taxonomically key for both vegetative and reproductive characters required.
(2) *Salix* flowers are very simple, so few floral features to identify.
(3) Floral and vegetative buds do not open at the same time.
(4) Developmental process vary with age.
(5) Climatic conditions affect the phenotypes.
(6) Difficulty of morphological identification due to high hybridization within genus.

Identification problem was tried to be solved by methods in field studies listed below (Skvortsov, 1999);

a. In taxonomical transect methods, one region rich of willows is chosen and then all species are identified.
b. Sampling is done two or three times during one season from the related specimens.

Both options are preferred by salicologists but the first is hardly possible to apply when the complex willow distributed region is concerned. To overcome this problem, numerical taxonomic methods namely Principal Component Analysis (PCA) by Sneath and Sokal (1973) has been applied and NTSYS-pc developed by Rohlf (2000) has been used. The morphometric approach, based on an appropriate character set, allowed scientists to understand the genus having taxonomical complications.

1.3. Hybridization in *Salix* species

Hybridization has been known since the period of Linnaeus and Scopoli who were the first to underline the existence of hybrids in willows (Skvortsov, 1999). Hybridization which is a very common phenomenon for *Salix* species is widely observed within European willows (Grant, 1981).

The willow chromosomes are uniform and small (Skvortsov, 1999). The chromosome number for *Salix* genus is 19, but 22 chromosomes is also seen as a result of fragmentation. The maximum number reported for *S. myrsinites* and *S. glauca* ssp. *callicarpaea* is 2n=190 (Wilkinson, 1944; Löve, 1961). *Salix* is famous to be one of the few woody genera with a large number of polyploid taxa. The polyploidy complexes of *Salix* was first given as a specific example for activity of hybridization in plant in 1950 (Stebbins, 1950). Interspecific crossability in willows gives rise to high polyploidy level. This high number of polyploidy level that resulted in *Salix* hybrids could have been an important mechanism in evolution. Thus, hybrids may have promoted the speciation mechanism in *Salix* genus (Karrenberg, 2002).

New genetic evidence suggests that hybrid speciation can also be common without polyploidy in plants than it is thought. Well establish homoploid plants are recognized and give rise to speciation (Mallet, 2007). Besides, introgressive hybridization is observed between *Salix* species and lead to speciation. Different from the hybridization, the term 'introgressive hybiridization' was first used by Anderson to describe the consequences of repeated backcrossing hybrids to parental types (Anderson and Stebbins, 1954). To illusturate, the gene flow effects on species variability for *Salix* genus is difficult to assess, but low reproductive barriers were pastly investigated in such crosses like *S. bebbiana x petiolaris*, *S. pellita x discolor*, and *S. petiolaris x eriocephala* (Mosseler, 1989). In *Salix* section *Longifoliae,* it was found that, without the occurrence of polyploidy, introgressive hybridization, itself has played a role in the evolution (Argus, 1997). In those years, the introgression event was identified in some European *Salix* taxa (*Salix alba*, *S. rubens* and *S. fragilis*) to demonstrate the first generation and introgressed hybrids by the help of AFLP (Beismann *et al*., 1997). Another specific work with *Salix* hybridization via introgression refer that drought stress of *Salix* could increase the rate of both backcrossing and introgression (Orians *et al*., 1998). Salicologist have still analysed molecular and morphological data from hybrid willows (*Salix sericea* and *S. eriocephala*) and report that chloroplast diversity may suggest both historic introgression, and contemporary hybridization (Hardig *et al.,* 2010). So introgression is more common in willow species.

Every species of *Salix* does not hybridize with anothers randomly. Some willows, sharing the same habitats, make hybrids which can be easily observed from intermediate morphology. It is really hard to examine this type of hybrid zone (Fritz, Orians and Brunsfeld *et al.,* 1991). Three biogeographical region of Turkey Mediterranean, the Irano-Anatolian and the Caucasus give rise to the presence of multiple hybrid zones and local refugial areas within Anatolia (Bilgin, 2011).

According to Skvortsov, intermediate morphological characteristics of *Salix* in nature do not refer to the hybrid species all the time. Insufficient knowledge on species of *Salix* L., caused confusion between hybrids and species. The factors responsible for nature and origin of hybridization still remain unclear in *Salix* genus. So it is obvious that existence of evolutionary forces in *Salix* genus such as hybridism, polyploidy and introgression lead to taxonomical problems.

1.4. The Uses of *Salix* species

The 'Cortex Salicis' derived from bark of willow has been used for modern phytotherapy for many years. The 'Cortex Salicis' is the source of salicin which is the pioneer substance of aspirin when transformed to salicylic acid molecule (Baytop, 1997). Nowadays, it is used for cancer research (Mahdi *et al*., 2006).

Anthropogenic factors that caused dispersal of heavy metals are tried to be remediated by *Salix* species (Evlard *et al*., 2014). Accumulation of Zn and Cu metals in bark and wood system of *Salix* sp. are investigated by many researchers who determined the rate of metal accumulation in *Salix* species in contaminated soil (Pulford and Watson, 2003). The level of accumulation of metal which sink in biological tissues can be changed according to related part of *Salix* tissues. Sander and Ericsson (1998) found concentrations of Zn, Cu, Ni and Cd in stems of *Salix viminalis* increased as a consequence of increasing bark proportions.

Bioenergy production is another popular research topic of the modern world. *Salix* species are outstanding candidates for bioenergy production (Vermerris, 2008). Data from determination of abiotic and biotic resistance genes is increased the use of *Salix* clones in biomass and bioenergy studies (Aravanopoulos *et al*., 1999). Some of *Salix*

clones are more preferred in forest biotechnology for quick growth, wide distribution, resistance to disease and stress features (Herrera, 2006). Przyborowski and Sulima (2010) showed the genetic characterization on *Salix viminalis* with RAPD markers which reveal how suitable it is for bioenergy production (Przyborowski, 2010). Higher degree of heterogeneity and higher level of variation between the *Salix viminalis* genotypes were found to make crossing and producing valuable form for breeding in bioenergy field. *Salix* species have also considerable value in preventing erosion curtain, basket, paper and fence making, ornamental uses and as animal feed (Tunçtaner, 1990).

1.5. The Genus *Salix* in Turkey

Number of *Salix* species vary from 23 (Davis, 1965) to 33 (Kantarcı, 2013). Naturally, 27 *Salix* sp. are found and 4 of them are endemic to Turkey; namely *S. trabzonica*, *S. purpurea* subsp. *leucodermis*, *S. rizeensis* and *S. anatolica* (Skvortsov and Edmonson, 1970; Donner, 1990; Güner and Zielinski, 1993; Güner, 2000, Zielinski and Tomaszewski, 2007) (Table 1.2)

Table 1.2. *Salix* species that are native to Turkey

Number	***Salix species***	***Author***
1	*S. acmophylla*	Boiss.
2	*S. triandra* subsp. *triandra*	L.
	S. triandra subsp. *bornmuelleri*	(Hausskn.) A. Skv.
3	*S. pentandra*	L.
4	*S. pentandroides*	A. Skv.
5	*S. alba*	L.
6	*S. excelsa*	J.F. Gmelin
7	*S. fragilis*	L.
8	*S. babylonica*	L.
9	*S. apoda*	Trautv.
10	*S. trabzonica* (endemic)	A. Skv.
11	*S. myrsinifolia*	Salisb.
12	*S. caucasica*	Andersson
13	S. *pedicellata* subsp. *pedicellata*	Desf.
14	*S.caprea*	L.
15	*S.aegyptiaca*	L.
16	*S.cinerea*	L.
17	*S.pseudomedemii*	E. Wolf
18	*S.pseudodepressa*	A. Skv.
19	*S.viminalis*	L.
20	*S.armenorossica*	A. Skv.
21	*S.elaeagnos*	Scop.
22	*S.elbursensis*	Boiss.
23	*S.amplexicaulis*	Bory and Chaub
24	*S.rizeensis* (endemic)	A. Güner et J. Zielinski
25	*S.purpurea* subsp. *leucodermis* (endemic)	L.
26	*S.wilhelmsiana*	Bieb.
27	*S.anatolica* (endemic)	J.Zielinski and D. Tomaszewski

The phytogeographical distribution of some of the Turkish species is in harmony with geographical region, for instance, (*Salix aegyptiaca*; Iran Turan element in Southeast Anatolia region) whilst some of them are not (*Salix fragilis*; Euro Sibernian element in Southeast Anatolia region) (Avcı, 1999).

In contrast to the Northern part of Europe, Turkey was not totally covered by ice sheets in the last glacial period starting from 2.6 Mya (Late Pliocene) (Erinç, 1978). In this interglacial period, Turkey acted like glacial refuge and reservoirs of diversity (Ledig, 1987). So in vegetational history of Turkey, when Anatolian plateau has been covered by steppes as a result of anthropogenic effects such as grazing and agriculture, most of the trees migrated to higher elevations like Northern Anatolia (Ledig, 1998).

The richest region of Turkey for *Salix* L. species is Northern Anatolia and Black Sea Region, with 23 species. This is followed by Eastern Anatolia region with 15 species.

The region having the least number of *Salix* L. is Southeast Anatolia with 6 species (Arıhan and Güvenç, 2009). *Salix rizeensis* from Rize has been recently introduced to literature by Güler and Zielinski (2000) and Terzioğlu *et al*. (2007) whom made morphological identification of *Salix rizeensis*. Furthermore, *S. myrsinifolia* and *S. purpurea* subsp. *leucodermis*, as new species for Turkey are added to the 11th volume of *Flora of Turkey* (Zielinski and Güner, 2000). In addition to these studies, some other scientists (Yaltırık, 1989; Avcı and Zielinski, 1997; Avcı, 1999; Arıhan and Güvenç, 2011) contributed to increase data source for *Salix* L. genus in Turkey. In forestry, Tuçtaner (1990) investigated the adaptive characters directed by genetic variation in willow clones. Up to today, there is no record about a molecular study on Turkish *Salix* species.

1.6. Molecular Phylogeny

The basic chromosome number of *Salix* is 19 and the chromosomes are uniform and small. About 40% of the willow species are detected as polyploids ranging from tetraploids to octoploids (Suda and Argus, 1968). The polyploidy levels in *Salix* genus mentioned by Argus (1997) vary from 3x to 12x. Today, it is noticed that cytogenetic and morphological characters are not enough to disctinguish plant species and to comprehend the relationships exist in evolutionary perspective. Thus, molecular barcodes are started to be used to clarify the problems in taxonomical and evolutionary research.

In combination or alone coding or non-coding regions of plastid genome loci are used as plant DNA barcodes (Hollinsworth *et al*., 2011). The Consortium for Barcode of Life (CBOL), Plant working group (2009) proposed that *rbc*L and *mat*K are the standard plant barcodes to make discrimination. The nuclear ribosomal internal transcribed spacer regions (*ITS*1 and *ITS*2) were also used as basic plant barcode regions (Kress *et al*., 2015). In addition to *ITS* data, the external transcribed spacer (*ETS*) of *18S–26S* rDNA as the genomic region has potential in phylogenetic studies (Baldwin and Markos, 1998).

There are various studies that inferred the importance of molecular approaches towards in *Salix* L. species. For instance, phylogenetic analyses on section Longifolia

of *Salix* genus (Brunsfeld *et al*.,1991), section separation by genomic ribosomal region *ITS* in Salicaceae family (Leskinen and Alström-Rapaport, 1999), phylogenetic analyses with *rbcL* sequence data on *Salix* genus (Azuma *et al.,* 2000), hybridization, introgression and phylogenetic analyses based on molecular phylogeny survey by chloroplast database in *Salix* L. (Chen *et al*., 2010), regional phylogenetic study on Iranian *Salix* species (Abdollahzadeh *et al.,* 2011), the role of trans-specific selective sweep for DNA barcodes in willows (Percy *et al.,* 2014), and nuclear markers (*ETS* and *ITS*) on subg. *Salix* (Wu *et al*., 2015). Recently, two new surveys about evolutionary implications by molecular markers are published by researchers (Lauren-Moreau *et al*., 2015; Liu *et al.,* 2016)

1.6.1. Chloroplast DNA (cpDNA) Regions

Chloroplasts are photosynthetic organelles of cells. They have their own genomes. In the literature, the first complete chloroplast genome was sequenced in *Nicotiana tabacum* (Shinozaki *et al*., 1986). The chloroplast genome of plants is made up of multiple copies of circular double stranded DNA (Masood *et al*., 2004). The genome comprises of large inverted repeats (IR), including ribosomal RNA (rRNA), six tRNA seven protein genes (Palmer, 1991). The inverted repeats are separated from each other by one large (LSC) and one small single-copy region (SSC). Many chloroplast regions are located in LSC. Numerous researchers prefer the single copy regions as a molecular tool to figure out relationships in evolution (Shaw *et al.*, 2007).

Phylogenetic relationships have been investigated by the help of chloroplast DNA (cpDNA) at different taxonomic levels. Sequencing by the regions of cpDNA has become one of the most popular areas of plant systematic and evolution (Clegg and Zurawski, 1993). Since chloroplast genome is abundant, haploid and conserved in plants, it provides useful data for phylogenetic studies. Both non-coding and coding regions of cpDNA; t-RNA (*trn*T-F), *mat*K (*maturase Kinase*) and *rbc*L (*ribulose-1, 5-bisphosphate carboxylase/oxygenase*) are utilized widely to clarify the relationships between Salicaceae family members.

1.6.2. *trn*T-F

The non-coding regions of chloroplast genome have a potential of high frequency mutations (Taberlet *et al*., 1991). Amplification of non-coding regions increases the resolution of cpDNA for evolutionary studies. The advantages of being conserved region made *trn*T-F suitable for molecular systematic studies of plants (Taberlet *et al*., 1991).

The cpDNA *trn*T-F region is composed of the transfer RNA genes *trn*T (UGU), *trn*L (UAA), and *trn*F (GAA) which are located in tandem and separated by non-coding spacer regions. In other words, it involves *trn*L $^{(UAA)}$ gene and two flanking intergenic spacers (IGS); *trn*T-L and *trn*L-F (Pirie *et al.,* 2007). The *trn*L gene contains group I intron located between the U and the A of the UAA anticodon loop (Kuhsel *et al*., 1990). In contrast to conserved regions [*trn*L $^{(UAA)}$ *5'* exon and *trn*L $^{(UAA)}$ *3'* exon] of *trn*L intron region, the core part (*trn*L intron) is variable (Bakker *et al.,* 2000). It is widely accepted that this region evolves faster than the chloroplast genome (Neuhaus *et al*., 1987).

Recently, using cpDNA *trn*T-F region in combination with *rbc*L and *mat*K region has answered questions in phylogenetic reconstruction, classification, biogeography and molecular dating (Pirie *et al*., 2007). Especially, using *trn*T-F region is effective in phylogenetic studies of groups such as *Populus* in which hybridization has significant role in evolutionary processes (Hamzeh and Dayanandan, 2004). Thus, the *trn*T-F region, including *trn*T-F (IGS), *trn*L intron and *trn*L-F (IGS) are useful to determine relationships *Salix* species.

1.6.3. *mat*K (maturase Kinase)

The chloroplast gene maturase Kinase (*mat*K) which is formerly known as open reading frame (ORF), located in intron of chloroplast gene *trn*K (lysine tRNA). *trn*K intron containing group II intron is different from typical group II introns by encoding *mat*K protein (Hausner *et al.,* 2006). The *mat*K gene, about 1500 base pairs (bp), is positioned on the large single-copy section (Hilu and Liang, 1997). It

encodes a maturase protein used in RNA splicing (Neuhaus and Link, 1987; Wolfe *et al.*, 1992; Mort *et al.*, 2001).

The data from *mat*K sequence can alone generate reliable phylogenetic trees in contrast to phylogenetic trees based on set of other cpDNA gene combination (*rbc*L and *atp*B) (Hilu *et al*., 2003). The *mat*K gene is commonly used in addressing plant molecular systematics and evolutionary questions (Barthet and Hilu, 2007). As it evolves fast in plants, the *mat*K is a marker to construct plant phylogenies in evolutionary studies (Hilu and Liang, 1997; Kelchner, 2002).

The *mat*K lacks indels, indicating the important function in evolution and conservation (Hilu and Liang, 1997). The variation is slightly higher at the sequencs near 5' region than sequence adjacent to the 3' end. Since the conserved 3'end has specific informative sites, it is useful in resolving phylogeny (Hilu and Liang, 1997). Hardig (2010) selected *mat*K gene region to elaborate phylogenetic analyses in *Salix* genus.

1.6.4. *rbc*L

Gene *rbc*L, which codes for the large subunit of *ribulose-1, 5-bis-phosphate carboxylase/oxygenase* (RuBisCO or RuBPCase), is a useful locus for phylogenetic analysis (Clegg *et al*., 1986). In studies, extensive variation in substitution for *rbc*L gene was reported for gymnosperms, dicots, and monocots as a result of life history factors (Bousquet *et al*., 1992). Chase (1993) also argued how valuable and informative *rbc*L gene region is for showing lineages of angiosperms. But at lower taxonomic level, *rbc*L gene data cannot solve the reticulate relationships as a consequence of its conserved nature (Gielly and Taberlet, 1994). Thus, *rbc*L gene has been used for understanding large interfamilial relationships of plants by scale cladistics analysis (Soltis and Soltis 1998; Savolainen *et al.,* 2000). Furthermore, according to some scientists, amplification of *rbc*L gene alone is too conserved to define in closely related taxa (Doebley *et al*., 1990; Plunkett, Soltis, and Soltis, 1997; Xiang, Soltis, and Soltis, 1998). To overcome this phenomenon, the Consortium for Barcode of Life (CBOL) Plant working group supported the efficiency of using combination of *rbc*L and *mat*K regions.

Nucleotide sequences of the chloroplast *rbc*L gene were first used to explore relationships of *Salix* with others genera (*Chosenia* and *Toisusu*) by Azuma *et al*., (2000). Besides, Wu *et al.* (2015) combined the data of *rbc*L region sequences with other chloroplast regions to discover the reticulate evolution mechanisms in willows.

1.6.5. Genomic DNA Region

Nuclear ribosomal DNA has been accepted as influential tool for evaluation studies as it is universal in all organisms and higher number of copies are repeated (Zimmer and Hamby, 1989). Every gene region has potential to determine lineages of different levels of divergence in plants (Youngbae *et al*., 1993). With the increasing complex problem of uniparentally inheritance in phylogeny, researches on nuclear sequence markers in plant systematics were conducted (Alvarez and Wendel, 2003). There are several regions having different rates of sequence variation in the repeated units. Now, genomic sequencing has become easy to use for phylogenetic surveys at all taxonomic levels.

1.6.6. *ITS* (Internal Transcribed Spacer)

The nuclear rDNA units, divided by intergenic spacers, involve the *18S*, *5.8S*, and *26S* coding regions in plants. The *5.8S* coding region, flanked by internal transcribed spacers 1 and 2 (*ITS* 1 and 2), are found between the *18S* and *26S* coding regions (Youngbae *et al*., 1993). The *18S* and *26S* coding regions have been suggested to show phylogenetic relations at the family or higher taxonomic levels in plants (Zimmer *et al*., 1989).

Baldwin (1995) performed a survey about the popularity of *ITS* sequence and found out that one third of the phylogenetic studies are based completely on *ITS* region. Internal transcribed spacer region (*ITS*) is one of the most popular sequences for phylogenetic analysis at generic and infrageneric levels in plants (Alvarez and Wendel, 2003). There are some reasons why *ITS* region is preferred in phylogenetic implications. One of them is ease of access while the other is uniparentally inherited

feature of the *ITS* region which makes it superior over cpDNA loci (Alvarez and Wendel, 2003).

Introgressive hybrid speciation and chloroplast capturing events among *Salix* species is known (Brunsfeld *et al*., 1991). Incomplete lineage sorting (Hardig *et al*., 2010, Percy *et al.,* 2015) has been reported therefore phylogenetic inference based on a chloroplast-encoded gene cannot examine the actual lineages of willows. Therefore performance of nuclear DNA is necessary in *Salix* genus to assimilate the relationships accurately. Various scientists in the world used *ITS* sequence as a tool to elucidate phylogeny of *Salix* (Leskinen and Alström-Rapaport, 1998; Abdollahzadeh *et al*., 2011, Barkalov and Kozyrenk, 2014).

1.7. Justification of the Study

The genus *Salix* L. is represented by more than 500 species in the world and 27 species among them are naturally found in Turkey (Terzioğlu *et al*., 2014). This genus has been important raw material as a biomass production for energy, phytoremediation and pharmacological products (Aravanopoulos, 1999; Pulford and Watson, 2002; Mahdi *et al*., 2006). They also have traditional ornamental and ethnobotanical such as basketry production and handcrafting (Arıhan and Güvenç, 2011). However, in contrast to this situation, there is no genetic information or research present on Turkish *Salix* L. species. This genus is one of the most poorly understood one in Turkey.

Since only the morhological traits of *Salix* do not reflect the evolutionary relationships, molecular analysis were performed for this genus in the world (Brunsfeld *et al.,* 1991; Azuma *et al*., 2000; Hardig *et al*., 2010, Lauren *et al*., 2015). But, there is no extensive molecular studies about Turkish *Salix* species to complete the phylogenetic tree of *Salix* genus in the world. Therefore, in the current study, combined molecular data (large sequence) were used to construct phylogenetic trees of *Salix* genus. Besides, to comprehend the relationship well, the additional PCA were performed in regard to eleven morphological characters for Turkish willows.

The current study will provide both morphological and molecular data for revision of *Salix* genus in the Flora of Turkey. Additionally, the molecular data on *Salix* genus will shed light on the general overview of *Salix* L. phylogeny and the placement of Turkish willow in it. Since they are economically important plants, it will gain worldwide interest. Other disipline such as pharmacy and forest engineering will take advantage of taxonomic clarification. Furthermore, putting DNA sequences from Turkish *Salix* species into NCBI system makes possible to provide data for other researchers.

1.8. Scope of the Study

This study aims to shed a light on general overview of Turkish *Salix* phylogeny. The combination of molecular dataset including 3 cpDNA and 1 nrDNA regions are utilized to determine the phylogenetic inherence of *Salix* L. genus.

The main purposes of the study are as follows:

- To reveal interspecific morphologic, phylogenetic and evolutionary relationships of Turkish *Salix* L. species.
- To construct a reliable molecular phylogenetic tree using DNA sequence data from cpDNA regions of *trn*T-F, *mat*K, *rbc*L and nrDNA *ITS* region for Turkish *Salix* L. species.
- To contribute information and bring a new aspect about Turkish *Salix* systematics by making combination of molecular and morphological datasets.

CHAPTER 2

MATERIALS AND METHODS

2.1. Plant Materials

The plant materials used in this study were collected during field studies extending from 2011 to 2016. The past location records were examined for naturally distributed *Salix* genus in Turkey and the field works were planned according to them (Figure 2.1, Figure 2.2). Species are collected and analyzed within the scope of the two projects, TOVAG 213O154 "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" supported by the Scientific and Technical Research Council of Turkey (TUBITAK) and BAP-01-08-2012-013 "Türkiye Söğüt Türlerinin Moleküler Filogenetiği" supported by Middle East Technical University (METU). In addition to field study, sample of some species were sampled from Behiçbey Nursery (AOÇ) and sample of several missing species were obtained from Ankara University Herbarium (ANK), Hacettepe University Herbarium (HUB), Gazi University Herbarium (GAZI) and Istanbul University Forest Department Herbarium (ISTO).

Collection of fresh willow leaves were extremely important, therefore the period of field studies were limited to spring and mid-summer. At least three samples, which are 200m away from each other, for every species from different locations, are collected. To utilize the identification of the species, one sample was preferably obtained with catkins and pressed. These samples were selected not only from the

locations of Flora of Turkey records for *Salix* L., but also nursery and new locations from various part of Turkey. Each sample with fresh leaves was well preserved by including silica gel in the packages. Dry leaf materials were taken from deposited herbarium specimens. All species sampled with their subgenus, section, and codes are listed in Appendix A.

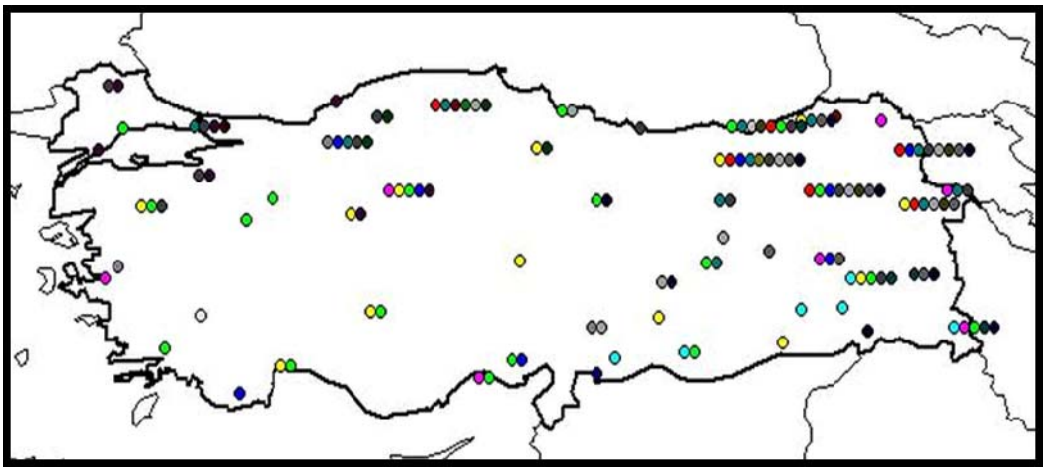

Figure 2.1. The natural distribution of *Salix* species in Turkey
(http://turkherb.ibu.edu.tr/index.php?sayfa=karsilastir)

Figure 2.2. The Locations of Sampled Turkish *Salix* species
(https://www.google.com/earth)

2.2. Morphological Analysis

During the field studies, other than the samples for molecular analysis, one individual for each species was preserved as a press material. Identification problems were solved by the help of Skvortsov (1999) transects method in which all *Salix* specimens are sampled from a selected region. The specimens were mainly identified according to the rules established in *Flora of Turkey* (Davis, 1965). Other relevant datasets such as TUBITAK checklist and Willow of Russia and Adjacent Countries (Skvortsov, 1999) were used for identification. The specimens were also compared with the type specimens in Ankara University Herbarium (ANK), Hacettepe University Herbarium (HUB), Gazi University Herbarium (GAZI).

2.3. Numerical Taxonomic Methods

The morphometric approach made infrageneric relationships of *Salix* species possible. Numerical taxonomic method, Principal Component Analysis (PCA), is applied for this purpose. Several morphological characters were chosen for analysis such as, life form, brunch habit, and bud scale, leaf and twig characters. Each scored specimen was designated and OTU measurements were made on them by the use of Leica MZ16 Fluorescence Stereomicroscope and Leica microscope camera. Principal Component (PCA) analysis was used as implemented in NTSYS pc software and viewed in MATLAB (MATLAB and Statistics Toolbox, 2012).

The data matrix formed with 9 characters (Table 2.1) belonging to *Salix* taxa was standardized with binary coding. Two more characters, bud angle (Ba) and Petiole length (Pl) were generated by taking photos and by measuring under stereomicroscopes (Figure 2.3). The petiole length is measured three times and averages were calculated. Finding out every generative part of samples is uneasy, especially for herbarium samples. Thus, only selected vegetative characters are used to discriminate morphologically. The morphological characters are chosen with respect to the discriminative features of *Salix* genus key in *Flora of Turkey* (Davis, 1965). Configuration of PCA accounted for 214 samples and all Operational Taxonomic Unit (OTU) used in the cluster analysis.

It is assumed that the selected characters may reflect taxonomic relationships quite well. Cluster analysis using the variables as numerical analysis allows determining the relationships among *Salix* species of Turkey.

Table 2.1. Selected morphological characters scored for *Salix* sp.

No	***Character***	***Data information***	***Units***
1	Life form (Lf)	Shrub or not	Binary; yes=1, no= 0
2	Bud scale (Bs)	Glabrous or not	Binary; yes=1, no= 0
3	Bud angle (Ba)	-	degree
4	Brunch habit (Bh)	Dropping or not	Binary; yes=1, no= 0
5	Petiole length (Pl)	-	mm
6	Bark type (Bt)	Fissured or smooth	Binary; yes=1, no= 0
7	Stipule persistence (Sp)	Persist or not	Binary; yes=1, no= 0
8	Decorticated wood (Dw)	Smooth or not	Binary; yes=1, no= 0
9	Leaf shape (Ls)	Lanceolate or not	Binary; yes=1, no= 0
10	Leaf colour (Lc)	Dark green above or not	Binary; yes=1, no= 0
11	Twig slender (St)	Slender or not	Binary; yes=1, no= 0

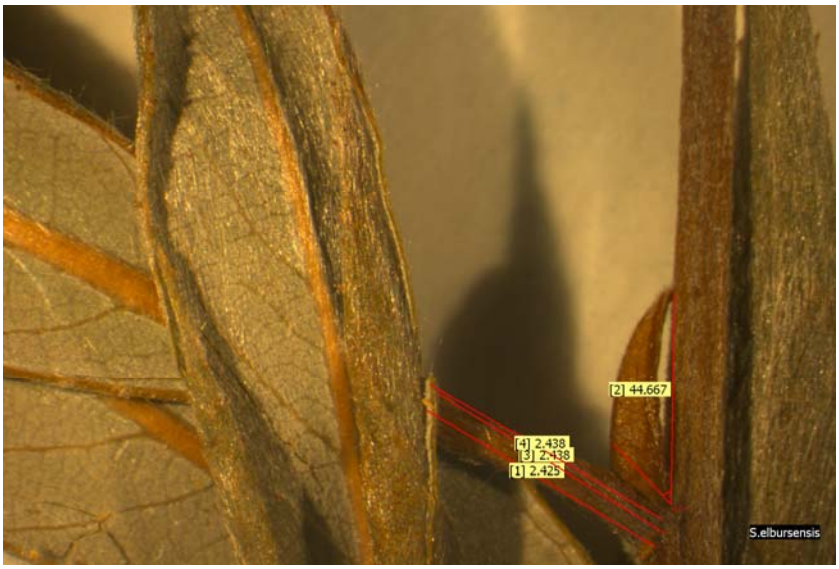

Figure 2.3. The images taken by Digital Firewire Color Camera System (Leica DFC320)

2.4. Total Genomic DNA Isolation from *Salix* species

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.

2.5. DNA Qualification

The presence and quality of the DNA were assessed by Biodrop (Biodrop µLite 7141 V.1.0.4, Biological Sciences Department, METU) which makes micro-volume measurements. The amount of every individual was adjusted to 10 ng/µl with best nucleic acid isolation visualization curve to use in the long run. The diluted DNAs were stored at +4ºC for a short period, whereas stock DNA samples were stored at -20ºC or -80ºC for the long term.

2.6. Studied regions; Chloroplast and Genomic DNA Regions

2.6.1. Chloroplast DNA (cpDNA) Regions

2.6.1.1. *trn*T-F

The region between the *trn*T (UGU) and *trn*F (GAA) genes in the large single-copy region is useful considering the succession of conserved cpDNA *trn*T-F genes. To amplify *trn*T-F regions, three primer pairs were used. The information on primer names, sequences and lengths of amplified regions were given Table 2.2 and illustrated in Figure 2.4.

Table 2.2. List of the universal primer pairs used to amplify the *trn*T-F region and their lengths as bp (Taberlet *et al.,* 1991)

Region name	Primer name	Sequence	Length of Region
*trn*T-L IGS	*trn*a(F) *trn*b(R)	5' CAT TAC AAA TGC GAT GCT CT 3' 5' TCT ACC GAT TTC GCC ATA TC 3'	~ 830 bp
*trn*L intron	*trn*c (F) *trn*d (R)	5' CGA AAT CGG TAG ACG CTA CG 3' 5' GGG GAT AGA GG ACT TGA AC 3'	~ 610 bp
*trn*L-F IGS	*trn*e (F) *trn*f (R)	5' GGT TCA AGT CCC TCT ATC CC 3' 5' ATT TGA ACT GGT GAC ACG AG 3'	~ 320 bp

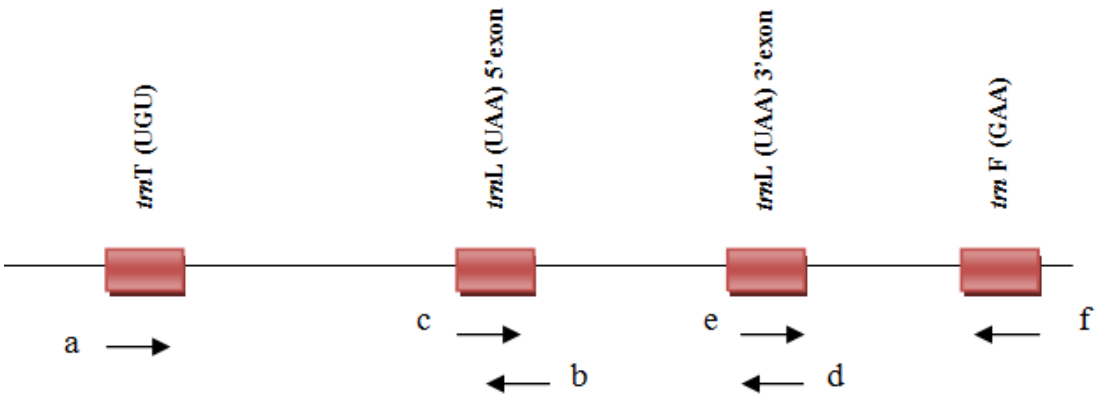

Figure 2.4. Positions and directions of universal primers (a-b, c-d and e-f) used to amplify non-coding regions of cpDNA (Taberlet *et al*., 1991)

(a, c and e = Forward primers, b, d and f =Reverse primers).

2.6.1.2. *mat*K (maturase Kinase)

Two primer pairs were used to amplify the whole *mat*K region. The gene is about 1500 bp in length and is embedded within introns that interrupt the two *trn*K exons (Li *et al*., 1997) (Figure 2.5). The primer names, their sequences, approximate length of the region were provided in Table 2.3.

Table 2.3. List of the primer pairs for amplifying *matK* region and their lengths as bp (Li *et al*., 1997)

Primer name	Sequence	Length of Region
***mat*K1** F1	5' ACT GTA TCG CAC TAT GTA TCA 3'	(yellow + green parts in Figure 2.5)
***mat*K1**R3	5' GAT CCG CTG TGA TAA TGA GA 3'	~ 1300 bp
***mat*K2**F5	5' TGGAGYCCTTCTTGAGCG 3'	(green + blue parts in Figure 2.5)
***mat*K2**R1	5' GAACTAGTCGGATGGAGTAG 3'	~ 1000 bp

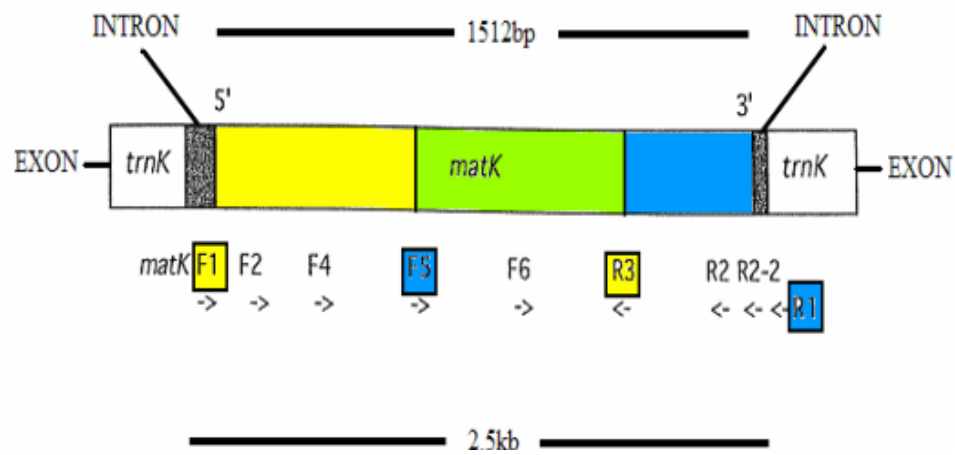

Figure 2.5. Relative positions of *mat*K primers (shaded areas are introns). Green part is amplified by both *mat*KF1, *mat*KR3 and *mat*KF5, *mat*KR1 (Li *et al*., 1997)

2.6.1.3. *rbc*L

The *rbc*L gene was amplified by the help of two primer pairs (Savolainen *et al*., 2000) (Figure 2.6). The first one includes 1F (Forward) and 724R (Reverse) called *rbc*L 1, while the second one comprises of 636 F (Forward), 1460 R (Reverse) called *rbc*L 2 (Table 2.4).

Table 2.4. List of the primer pairs for amplifying *rbc*L gene region and their lengths as bp (Savolainen *et al*., 2000).

Primer name	Sequence	Length of Region
1F (*rbc*L 1)	5'-ATGTCACCACAAACAGAAAC-3'	~ 660 bp
724R	5'- TCGCATGTACCTGCAGTAGC -3'	
636F (*rbc*L 2)	5'-GCGTTGGAGAGATCGTTTCT -3'	~ 840 bp
1460R	5'- TCCTTTTAGTAAAAGATTGGGCCGAG-3'	

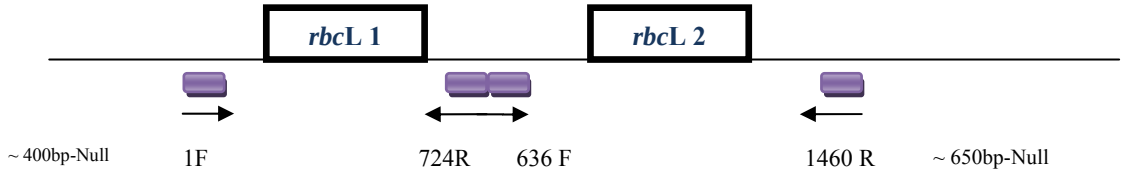

Figure 2.6. Primer used for amplification and sequencing of *rbc*L gene. The short solid lines show the primers (Savolainen *et al*., 2000)

2.6.2. Genomic DNA Region

2.6.2.1. *ITS* (Internal transcribed spacer) Region

To make comparison, other than the cpDNA regions, one genomic DNA region, *ITS*, was chosen in order to figure out the phylogenetic implications in *Salix* L. genus.

This region is highly repeated within the plant genome, but it evolves very quickly that promotes intragenomic uniformity (Leskin and Alström-Rapaport, 1999). *ITS*L (forward) and *ITS*4 (reverse) primers were used to amplify total region (*ITS*1+*5.8S*+*ITS*2) rather than using two primer pairs (*ITS*1, *ITS*2) (Hsiao *et al*., 1995) (Table 2.5). Illustration of *ITS* region was given in Figure 2.7.

Table 2.5. List of primers used for amplification of *ITS1+5.8S+ITS2* region and their lengths (Hsiao *et al*., 1995)

Primer name	Sequence	Length of Region (*ITS1-5.8S-ITS2*)
*ITS*L (F) *ITS*4 (R)	5' TCG TAA CAA GGT TTC CGT AGG TG3' 5' TCC TCC GCT TAT TGA TAT GC 3'	(~650bp)

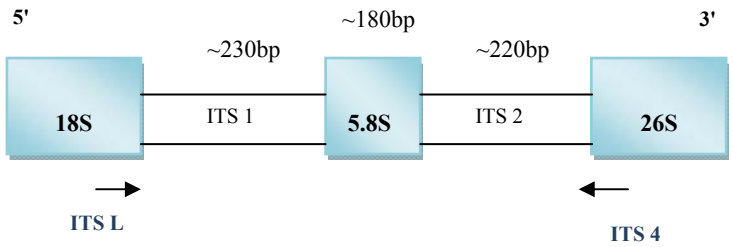

Figure 2.7. Primer used for amplification and sequencing of nrDNA *ITS* gene. The arrows show the primers (*ITS* L-4) (Hsiao *et al*., 1995)

2.7. Amplification; Polymerase Chain Reaction (PCR)

To obtain clear sequence data, the interested region need to be amplified. Even though there are many studies including optimized amplification and program conditions for each primer we used, the new optimized PCR conditions were determined for Turkish *Salix* L. species. Table 2.6 illustrates optimized PCR mixture conditions for each region or primer pairs. Besides, it is clearly seen the PCR

amplification program condition which were optimized for each primer pair were provided in Table 2.7.

Table 2.6. Tested PCR conditions of target regions for *Salix* L. species

Ingredients of PCR mixture	***trnT-L IGS*** *trn a-b*	***trnL intron/*** *trn c-d*	***trnL-F IGS*** *trn e-f*	***matK1*** *F1-F3*	***matK2*** *F5-R1*	***rbcL1*** *1F-724R*	***rbcL2/*** *636F - 1460R*	***ITS*** *ITSL-4*
H_2O	12.9 µl	13.2 µl	12.4 µl	12.2 µl	14.1 µl	13.3 µl	13.3 µl	13.3 µl
$MgCl_2$ (25 mM)	4 µl	3 µl	4 µl	4 µl	3 µl	3 µl	3 µl	3 µl
10X PCR Buffer	3 µl	2 µl	3 µl	4 µl	2 µl	3 µl	3 µl	3 µl
dNTP (10 mM)	0.6 µl	0.5 µl	0.5 µl	0.5 µl	0.6 µl	0.5 µl	0.5 µl	0.5 µl
Forward primer (10 µM)	0.5 µl	0.5 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.5 µl
Reverse primer (10 µM)	0.5 µl	0.5 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.5 µl
Taq DNA poly. (5u/µl)	0.5 µl	0.3 µl	0.3 µl	0.5 µl	0.5 µl	0.4 µl	0.4 µl	0.2 µl
DNA (10 ng/ µl)	3 µl	5 µl	4 µl	3 µl	4 µl	4 µl	4 µl	4 µl
Total	25 µl	25 µl	25 µl	25 µl	25 µl	25 µl	25 µl	25 µl

Table 2.7. Optimized thermal cycler program used for amplification of target region

Region (Primers)	Step	Temperature	Time	Cycle #	Purpose
trn*T-F *IGS (trn a-b)	1	95°C	4'	1	Denaturation
	2	94 °C	1'	35	Denaturation
		58°C	45"		Annealing
		72 °C	40"		Extension
	3	72 °C	10'	1	Final Extension
trn* L *Intron (trn c-d)	1	95°C	5'	1	Denaturation
	2	94 °C	30"	35	Denaturation
		58.3°C	30"		Annealing
		72 °C	1'		Extension
	3	72 °C	10'	1	Final Extension
trn*L-F *IGS (trn e-f)	1	95°C	4'	1	Denaturation
	2	94 °C	30"	35	Denaturation
		60.5°C	30"		Annealing
		72 °C	40"		Extension
	3	72 °C	10'	1	Final Extension
***mat*K1** (F1-R3)	1	94°C	5'	1	Denaturation
	2	94 °C	1'	35	Denaturation
		55.5°C	1'		Annealing
		72 °C	2'		Extension
	3	72 °C	10'	1	Final Extension
***mat*K2** (F5-R1)	1	95°C	5'	1	Denaturation
	2	94 °C	1'	35	Denaturation
		52°C	40"		Annealing
		72 °C	1'		Extension
	3	72 °C	8'	1	Final Extension
***rbc*L 1** (1F-724R)	1	95°C	5'	1	Denaturation
	2	95 °C	1'	35	Denaturation
		48°C	30"		Annealing
		72 °C	1'		Extension
	3	72 °C	7'	1	Final Extension
***rbc*L 2** (636F - 1460R)	1	95°C	5'	1	Denaturation
	2	95 °C	1'	35	Denaturation
		48°C	30"		Annealing
		72 °C	1'		Extension
	3	72 °C	7'	1	Final Extension
***ITS* 1and 2** (*ITS*L-4)	1	94°C	6'	1	Denaturation
	2	94 °C	70"	30	Denaturation
		57.5°C	45"		Annealing
		72 °C	90"		Extension
	3	72 °C	10'	1	Final Extension

2.8. Agarose Gel Electrophoresis

The suitable agarose gel was always prepared according to length of amplified region. In the current study, 0.8%, 1%, and 1.5% of agarose gels were preferably been used to run the samples. The gel was prepared by dissolving and boiling with 1X TBE (from 1 liter of 5X stock solution: 54 g of Tris base – 27.5 g of Boric acid –

20 ml of 0.5 M EDTA pH 8.0) buffer. After boiling by the help of microwave oven, the solution was stained with ethidium bromide. After that, the solution was poured into a horizontal gel tray including combs. Then, it was left for polymerization for 45 minutes. After polymerization, 1X TBE buffer was put into the electrophoresis apparatus and combs were removed. It is important not to damage the wells. 6X DNA loading dye (Fermentas) is mixed and loaded with samples into each well. A suitable ladder, a sample of DNA where the sizes of the bands are known, is generally used in first lane of the gel. After loading samples and ladder, the tank is connected to the power supply. Agarose gels were run at 80-100V for 30-60 minutes. The bands were visualized under UV light. GeneRuler™ 100 bp ladder, was used to comprehend the expected length of amplified bands in every gels. If the expected size of bands were clearly visible, they were ready to be used for sequencing.

2.9. Sequencing, Data Collection and Analysis of Sequence Data

The PCR products of amplification of target regions were stored at -20°C. The purification and sequencing procedure were done by REFGEN Biotechnology Company (Teknokent of Ankara University) and Genoks Molecular Biotechnology Company (Cinnah, Ankara). An ABI 310 Genetic Analyzer (PE applied Biosystem) and ABI3730XL 96 capillary automatic sequencer were used for sequencing of amplified DNA products. Both forward and reverse DNA sequences were amplified in order to get accurate sequences for chloroplast and nuclear regions. The CLUSTAL W software was used for multiple alignment if they are correctly read. Finch TV (Version 1.4.0) developed by the Geopiza Research Team was utilized to view the chromatogram data and to check base calls (Patterson *et al*., 2004-2006) (See Appendix C).

Phylogenetic and molecular evolutionary statistics such as total nucleotide length (bp), GC content (%), nucleotide deletion and insertion (indel), conserved and variable sites, parsimony informative sites, transition/ transversion (tr/tv) R ratio and nucleotide diversity of the sequences were estimated with the MEGA 7.0.2 Software (Tamura *et al*., 2012) for all cpDNA and genomic DNA regions.

2.10. The Genetic Distance between Taxa

Evolutionary distances among the studied taxa are significant for the study of molecular evolution and useful for phylogenetic reconstructions. Distance estimates the mean number of changes per site since 2 species (sequences) were split from each other. Thus, the number of nucleotide substitutions occurring between two or more DNA sequences designates the distance (Nei and Kumar, 2000). The genetic divergences between species or samples were calculated by using best fitting model in MEGA 7.0.2 Software.

2.11. Molecular Diversity and Construction of Phylogenetic Trees

The relevant sequences from 4 regions (*trn*T-F, *mat*K, *rbc*L and *ITS*) were used to figure out the differentiation among *Salix* L. species. It is fundamental to reveal evolutionary relations among various species and other groups of organisms having a common ancestor by constructing phylogenetic trees. During the construction of the phylogenetic tree, the outgroups' sequences were obtained from NCBI (National Center for Biotechnology Information Search database). To construct phylogenetic trees, MEGA 7.0.2 neighbor-joining (NJ) method (Saitou and Nei, 1987) was preferred. Only one representative individual for each species, code for every species and one species as an outgroup were taken into consideration.

Different from the main phylogenetic tree in every gene region, additional trees (subtrees) were also constructed for phylogenetic relations. In these subtrees, every species is labelled according to their sections. Deleted fractions and several substitutions were identified in the aligned DNA sequences of both cpDNA and nrDNA regions.

Felsenstein's (1985) bootstrap test commonly used to show reliability was also used in all constructed trees. The reliability of a given branch in a tree is proven by the frequency of existence of the branch in a large number of trees. The bootstrap value is the percentages of times each interior branch valued as 100 %. If the bootstrap value for a given interior branch is 95% or higher, so the topology at that branch is

examined as "correct". If the value is between 50 and 95 %, the topology is considered informative (Nei and Kumar, 2000).

2.12. Haplotype Frequency Analysis, Analyses of Molecular Variance (AMOVA), and Estimation of *F*st Values

In the current study, to make persuasive discrimination between individuals, haplotype (i.e. combination of alleles at one or more loci) analysis is applied. The frequency estimation of all possible haplotypes by maximum likelihood methods has been performed by using DnaSP v5 (Librado and Rozas, 2009). To calculate the amount of population genetic structure, Analyses of Molecular Variance (AMOVA) method of Arlequin 3.5.1.3 Software (Excoffier *et al*., 1992, 2005) were utilized. Additionally, the pairwise *F*st values were calculated as genetic distances to estimate the pairwise genetic distances among *Salix* taxa. It is used with the application of a transformation to linearize the distances with the population divergence time (Reynolds *et al*., 1983; Slatkin, 1995).

2.13. Molecular Clock Estimation

Zuckerkandl and Pauling (1965) were the first to define "the molecular evolutionary clock" term to show the time of the divergence of species by using nucleotide differences in the sequenced DNA region. Molecular clock states that DNA sequences evolve at a rate that is relatively constant over time and among different organisms. When two or more lineages evolved at the same, constant rate, the number of variations among two species or samples would be straightforward index of the time since they diverged from their common ancestor (Futuyma, 2005).

The number of parsimony informative sites in the sequenced DNA region was used for calculation. The following equation was used to estimate molecular clock for *Salix* L. genus.

$$\text{Molecular Clock} = \frac{k}{\text{mutation rate}}$$

Where k is equal to:

$$k = -\left(\frac{3}{4}\right)\ln\left(1 - \frac{4}{3}d\right)$$

The d in the above equation was calculated as:

$$d = \frac{\textit{Variable Site}}{\textit{Total Number of Base Pairs Sequenced}}$$

In the equation, *d* means the number of substitutions per base pair while *k* is the substitutions since divergence time. In the current study, DNA sequences of three regions (*trn*T-F*, mat*K*, rbc*L*)* in chloroplast DNA were evaluated. For each region, this value was estimated separately. As the mutation rate, 2.5×10^{-9} of a *Populus* sp. was used as a constant value (Levsen *et al*., 2012).

In the current study, not only DNA sequences from European, Chinese, Iranian and Russian samples (Old World *Salix*), but also those of American samples (New World *Salix*) were included to figure out approximate divergence times among *Salix* sp. in the world.

CHAPTER 3

RESULTS

3.1. Numerical Taxonomic Method

Principle component analysis (PCA) utilizing correlation matrices among the characters and similarity coefficients were calculated in the NTSYS-pc software package and viewed in MATLAB. The codes for each species are demonstrated in the box for every analysis given in a table (Table 3.1). By the help of the chosen morphological characters, two and three dimensional configuration of the principal component analysis showed two clusters (subg. *Salix* and *Vetrix*) for all Operational Taxonomic Unit (OTU) used in the cluster analysis (Figure 3.1). Two clusters are especially separated in two dimensional view of PCA (Figure 3.2). The members of subg. *Salix* are scattered to three dimensional PCA view (Figure 3.3). The hybrid species *S. alba x fragilis* with intermediate morphological character also positioned near to *S. fragilis* and *S. alba*. The hybrid is more close to *S. alba* than it is to *S. fragilis*. A few exceptional species deviate from the subg. *Vetrix* cluster; *S. wilhelmsiana*, *S. cinerea*, *S. caprea* and *S. myrsinifolia*. The number of deviated samples is represented by only one sample each for *S. cinerea* and *S. caprea*. The taxa, *S.amplexicaulis* is always located at the edge of the cluster of subg. *Vetrix* in both 2D and 3D views of PCA graphics (Figure 3.4). Furthermore, *S. myrsinifolia* and *S. wilhelmsiana* were herbarium materials and represented by very low sample size. Since the deviations are statistically significant due to limited number of sample for each species, it can be underestimated.

Even though the two quantitative data, Bud angle (Ba) and Petiole length (Pl), discriminate the species, bud scale (glabrous/not), life form (shrub/not) and leaf shape (lanceolate/not) are also the important characters in grouping. Especially these traits were useful in the members of subg. *Salix* which have mostly tree-like life form and lanceolate leaf shape. Pubescence nature of bud scale was also detected in members of subg. *Vetrix.*

Table 3.1. The codes and the number of samples used in PCA for each *Salix* species

The code	**Species**	**# Samples**
ALBA	*S. alba*	55
EXCE	*S. excelsa*	41
TRIA subsp tri	*S. triandra* subsp. *triandra*	29
TRIA subsp bor	*S. triandra* subsp. *bornmuelleri*	2
BABY	*S. babylonica*	22
PENT	*S. pentandroides*	6
CINE	*S. cinerea*	11
FRAG	*S. fragilis*	12
PSEUDO	*S. pseudomedemii*	2
AEGY	*S. aegyptiaca*	2
WILH*	*S. wilhelmsiana*	3
VIMI	*S. viminalis*	1
ALBxfra**	*S. alba x fragilis*	1
PEDI subsp pe	*S. pedicellata* subsp. *pedicellata*	3
ACMO	*S. acmophylla*	2
AMPL	*S. amplexicaulis*	3
ELBU	*S. elbursensis*	3
ARME	*S. armenorossica*	2
ELAE	*S. elaeagnos*	3
CAPR	*S. caprea*	3
CAUC	*S. caucasica*	3
APOD	*S. apoda*	1
PURP subsp leu	*S. purpurea* subsp. *leucodermis*	1
MYRS*	*S. myrsinifolia*	1
RIZE*	*S. rizeensis*	1
PSEUDEP*	*S. pseudodepressa*	1

*Herbarium materials.

**Hybrid species.

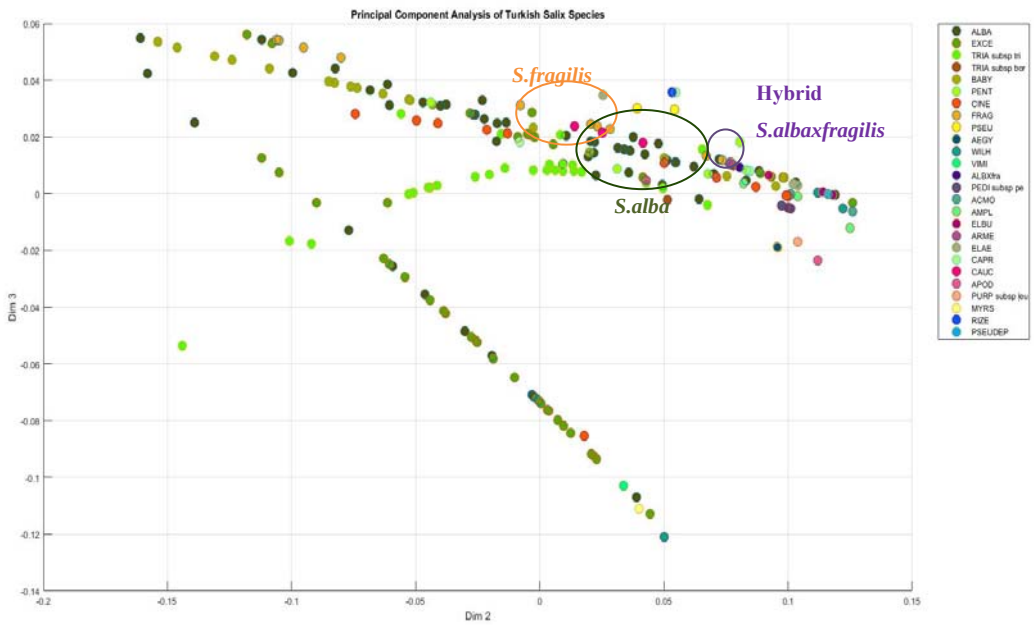

Figure 3.1. Two dimensional display of Turkish *Salix* species

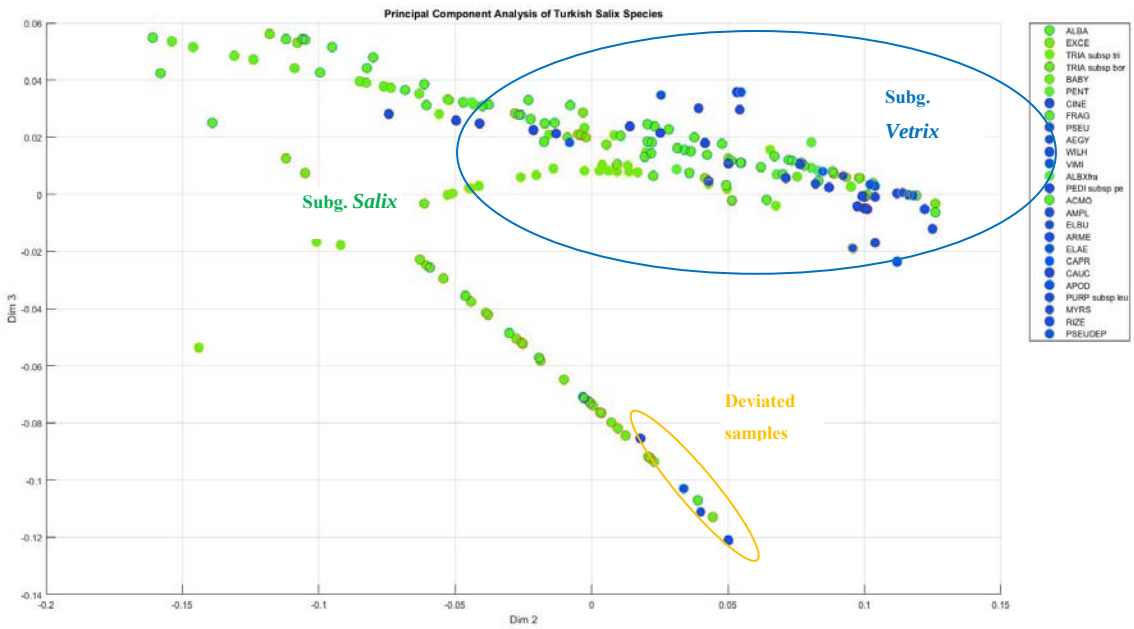

Figure 3.2. Two dimensional display of Subgenera of Turkish *Salix* species

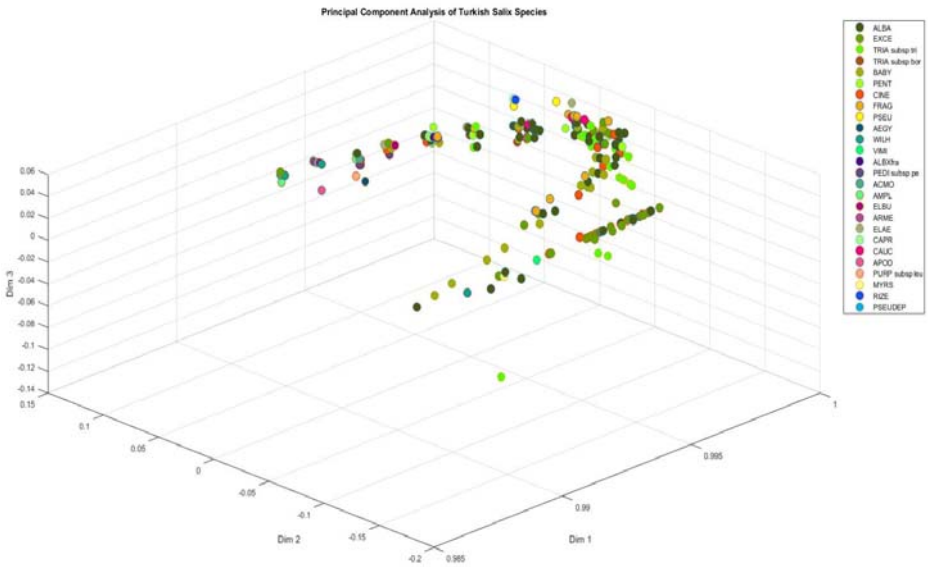

Figure 3.3. Three dimensional displays of PCA of Turkish *Salix* species

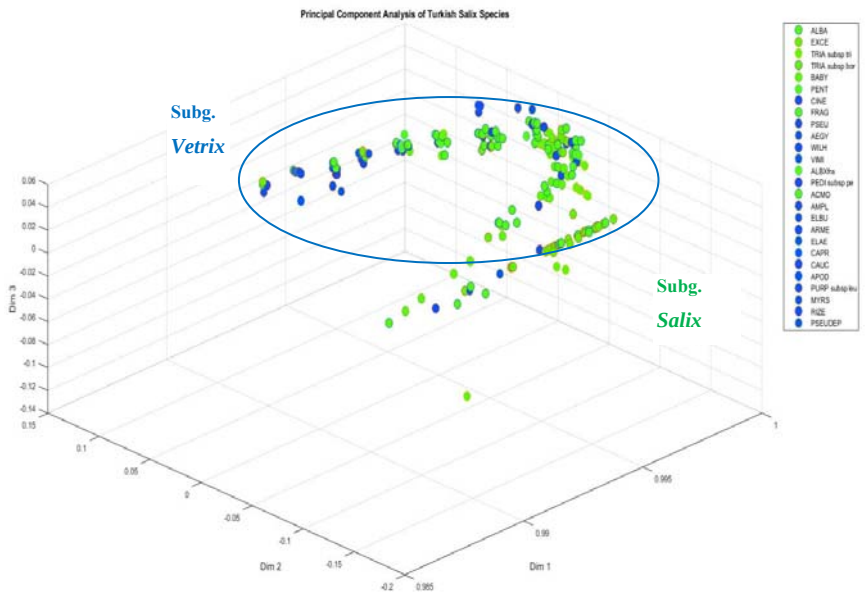

Figure 3.4. Three dimensional display of Subgenera of Turkish *Salix* species

3.2. Amplification of Studied Chloroplast and Nuclear Region

Both studied cpDNA and nrDNA regions were amplified and sequenced by using PCR mixture and cycle conditions (Table 2.6 and Table 2.7) for each of the 126 *Salix* individuals.

Different number of sequences from various locations were used for each species (Table 3.2). The location informations for every sample were listed in Appendix A in

detail. The length of target regions in proper concentrated agarose gel by running amplified DNAs of each species are given in Appendix B. The suitable DNA ladder, The visualization of sequences of primer pairs chromatograms were observed by the help of Finch TV Software (See Appendix C).

Table 3.2. The List of sequences and the locations used for every species (different province of Turkey)

Subgenus	Species name	Sequence number	Location (Town/Province)
SALIX	***S. acmophylla***	1	Asma KöprüSuçeken/Batman
	S .acmophylla	2	Asma KöprüSuçeken/Batman
	S. acmophylla	3	Birecik Halfeti/Şanlıurfa
	S. acmophylla	4	Birecik Halfeti/Şanlıurfa
	S. acmophylla	5	Birecik-Ziyaret Köyü/Şanlıurfa
	S. triandra* subsp. *triandra	1	ÇerkeşOrman Fidanlığı/Çankırı
	S. triandra* subsp. *triandra	2	Üçköy/Çorum
	S. triandra* subsp. *triandra	3	Ardıçlı-Niksar/Tokat
	S. triandra* subsp. *triandra	4	Afyon
	S. triandra* subsp. *triandra	5	Uzundere-Tortum/Erzurum
	S .triandra* subsp. *triandra	6	Uzungöl/Rize
	S. triandra* subsp. *triandra	7	Uzungöl/Rize
	S. triandra* subsp. *triandra	8	Tosya-Beşçam/Kastamonu
	S. triandra* subsp. *triandra	9	Tosya-Beşçam/Kastamonu
	S. triandra* subsp. *triandra	10	Ihlara Vadisi
	S. triandra* subsp. *triandra	11	Ihlara Vadisi
	S. triandra* subsp. *bornmuelleri	1	Çeltek-Tersakan/Amasya
	S. triandra* subsp. *bornmuelleri	2	Çeltek-Tersakan/Amasya
	S. pentandroides	1	Topulyurdu/Sivas
	S. pentandroides	2	Behiçbey Nursery/Ankara
	S. pentandroides	3	Çarşamba/Samsun
	S. pentandroides	4	Çoruh-Bağbaşı/Erzurum
	S. pentandroides	5	Çoruh-Bağbaşı/Erzurum
	S. pentandroides	6	Güleman-Ayıpınar/Elazığ
	S. pentandroides	7	Ladik/Amasya
	S. alba	1	Akyazı-Vakıf/Sakarya
	S. alba	2	Kızılcahamam-Güvem/Ankara
	S. alba	3	Konya
	S. alba	4	Kızılcahamam/Ankara
	S. alba	5	ÇoruhYavuzmah.

			Köprüsü/Artvin
	S. alba	6	Çoruh/Artvin
	S. alba	7	İspir/ Erzurum
	S. alba	8	Bor/Niğde
	S. alba	9	Ürgüp/Nevşehir
	S. alba	10	Uluırmak Köprüsü/ Aksaray
	S. alba	11	Kurtboğazıbarajıaltı/Ankara
	S. alba	12	Yabani-Ilgın-Çatak/Konya
	S. alba	13	Yabani-Ilgın-Çatak/Konya
	S. excelsa	1	Çarşamba-Çelikli/Samsun
	S. excelsa	2	Şenyurt/Artvin
VETRIX	***S. excelsa***	3	Kışlacık Köyü/Kırklareli
	S. excelsa	4	Kışlacık/Kırklareli
	S. excelsa	5	Ovacık Köyü/Sivas
	S. excelsa	6	Derekapı-Yusufeli/Artvin
	S. fragilis	1	Çay/Afyon
	S. fragilis	2	Behiçbey Nursery/Ankara
	S. fragilis	3	BeynamOrmanı /Ankara
	S. fragilis	4	Akşehir/Konya
	S. alba x fragilis	1	Behiçbey Nursery
	S. babylonica	1	Kalecik/Ankara
	S. babylonica	2	YaylacıkKöyü/Amasya
	S. babylonica	3	Tokat
	S. babylonica	4	ÇoruhYaylacıkÇıkışı/Artvin
	S. babylonica	5	Yusufeli-Çoruh/Artvin
	S. babylonica	6	Behiçbey Nursery/Ankara
	S. babylonica	7	IhlaraValley/Aksaray
	S. apoda	1	Ladik/Amasya
	S. apoda	2	Çamlıhemşin Şenköy çevresi/Artvin
	S. myrsinifolia	1	Ilgaz/Kastamonu
	S. myrsinifolia	2	Ilgaz/Kastamonu
	S. caucasica	1	Yokuşlu köyü/Artvin
	S. caucasica	2	Çoruh /Artvin
	S. caucasica	3	Çoruh-Sırakonaklar/Artvin
	S. caucasica	4	Çoruh-Sırakonaklar/Artvin
	S. caucasica	5	Ayder/Rize
	S. caucasica	6	Ayder/ Rize
	S. pedicellata* subsp. *pedicellata	1	Göksu-Ermenek/Karaman
	S. pedicellata* subsp. *pedicellata	2	Maraş
	S. caprea	1	Kızılcahamam Soğuksu Milli Parkı/Ankara
	S. caprea	2	Kastamonu-Çankırı il sınırı
	S. caprea	3	Kastamonu-Çankırı il sınırı
	S. caprea	4	Bostan/Kastamonu
	S. caprea	5	Kafkasör Yaylası/Artvin
	S. caprea	6	Kafkasör Yaylası/ Artvin
	S. caprea	7	Kafkasör Yaylası/Artvin
	S. aegyptiaca	1	Bahçesaray-yol kenarı/Van
	S. aegyptiaca	2	Tatvan/Bitlis
	S. aegyptiaca	3	Kars-Erzurum Yolu
	S. cinerea	1	Akyazı Gebeş/Sakarya
	S. cinerea	2	Çubuk-Karagöl/ Ankara
	S. cinerea	3	Çubuk-Karagöl/ Ankara
	S. cinerea	4	Çoruh-Bağbaşı/Erzurum
	S. cinerea	5	Çoruh-Bağbaşı/Erzurum
	S. cinerea	6	Çoruh-Bağbaşı/Erzurum

	S. cinerea	7	Çoruh-Bağbaşı/Erzurum
	S. pseudomedemii	1	Zile/Tokat
	S. pseudomedemii	2	Behiçbey Nursery/Ankara
	S. pseudomedemii	3	Beynam Forest/Ankara
	S. pseudomedemii	4	Beynam Forest/Ankara
	S. pseudomedemii	5	Beynam Forest/Ankara
	S. pseudodepressa	1	Sarıkamış/ Kars
	S. pseudodepressa	2	GümüşDamla Köyü/Bayburt
	S. viminalis	1	Nehir Başı/Erzurum
	S. viminalis	2	Erzurum
	S. armenorossica	1	Oltu/ Erzurum
	S. armenorossica	2	Oltu/ Erzurum
	S. armenorossica	3	Bağbaşı-Çoruh/Erzurum
	S. armenorossica	4	Bağbaşı-Çoruh/Erzurum
	S. armenorossica	5	Bağbaşı-Çoruh/Erzurum
	S. elaeagnos	1	Ilgaz/Kastamonu
	S. elaeagnos	2	Tosya-Beşçam/Kastamonu
	S. elaeagnos	3	Bostan/ Kastamonu
	S. elaeagnos	4	Cide/Kastamonu
	S. elaeagnos	5	Ağlı/Kastamnou
	S. elbursensis	1	Çoruh-Alanbaşı/Artvin
	S. elbursensis	2	Çoruh-Alanbaşı/Artvin
	S. elbursensis	3	Çoruh-Alanbaşı/Artvin
	S. elbursensis	4	Çoruh-Alanbaşı/Artvin
	S. amplexicaulis	1	Kızılcahamam-SoğuksuMilliParkı/ Ankara
	S. amplexicaulis	2	Çubuk-Karagöl/ Ankara
	S. amplexicaulis	3	Çubuk-Karagöl/ Ankara
	S. amplexicaulis	4	Kızılcahamam-SoğuksuMilliParkı/ Ankara
	S. amplexicaulis	5	Ilgaz/Kastamonu
	S. amplexicaulis	6	Ilgaz/Kastamonu
	S. amplexicaulis	7	Ilgaz/Kastamonu
	S. rizeensis	1	İkizdere/Rize
	S. rizeensis	2	İkizdere/Rize
	S. purpurea* subsp. *leucodermis	1	Köyceğiz-Dana deresi /Muğla
	S. purpurea* subsp. *leucodermis	2	Gökçeova Gölü/ Muğla
	S. purpurea* subsp. *leucodermis	3	Köyceğiz-Ağla/Muğla
	S. purpurea* subsp. *leucodermis	4	Köyceğiz-Dana deresi /Muğla
	S. purpurea* subsp. *leucodermis	5	Köyceğiz-Dana deresi /Muğla
	S. purpurea* subsp. *leucodermis	6	Köyceğiz-Dana deresi /Muğla
	S. wilhelmsiana	1	İkizdere/Rize
	S. wilhelmsiana	2	Tortum/Erzurum
	S. wilhelmsiana	3	Kars Erzurum Road

3.3. Molecular Diversity in *Salix* sp.

3.3.1. Molecular Diversity Statistics of *Salix* Species based on Each Region

3.3.1.1. *trn*T-F

The full-length of non-coding t-RNA (*trn*T-F) region of cpDNA was obtained with the help of three primers (*trn* a-b, c-d and e-f). The aligned sequences of three primers were assembled with the MEGA software. After alignment of all *Salix* samples, the length of the region was determined as 1339 bp (Table 3.3). The most variable part of *trn*T-F for *Salix* species is *trn*L intron when compared with other sub regions; *trn* T-L IGS and *trn*L-F IGS. No indels for the non-coding *trn* T-F region for Turkish *Salix* were observed (Table 3.3). The variable site is detected not only in subgenus level, but also observed in species level. A parsimony informative site means that at least two types of nucleotides with a minimum frequency of two must be located on the different DNA sequences of samples. Most of the variable sites for *trn*T-F region were parsimony informative (49 bp). GC content (%) in the sequences of subgenera and the total sample were almost the same which was approximately 30 %.

Considering transition and transversion sites, if a purine is substituted by another purine (Adenine vs Guanine) or a pyrimidine by another pyrimidine (Thymine vs Cytosine), it is called transition (si). However, if a purine is substituted by a pyrimidine or vice versa (Adenine vs. Thymine or Guanine vs. Cytosine), this situation is called tranversion (sv). The transition and transversion substution rates were calculated as 50.71 and 49.29, respectively. It appears that substitution rate is contributed equally in nucleotide composition. The rate (R) is calculated using the equation $R=[A \times G \times k1 + T \times C \times k2] / [(A + G) \times (T + C)]$ (k_1; transition/transversion rate (si/sv) ratios for purines and k_2; for pyrimidines) which is 0.91 based on *trn*T-F region for Turkish *Salix* species. Nucleotide diversity is a measure of polymorphism in a sample of gene sequences, and also, it is a summary statistic used to represent patterns of molecular diversity within a sample of gene copies (Nei, 1987). The nucleotide diversity of overall sequences was identified with

the number as high as 0.017. Nucleotide diversity was higher in subg. *Vetrix* than subg. *Salix*.

Table 3.3. The estimated molecular diversity parameters based on *trn*T-F of cpDNA region of Turkish *Salix* species data

	Subgenus *Salix*	Subgenus *Vetrix*	Over all (Subg. *Salix*, *Vetrix*)
Number of species	7+1hybrid*	16	24+1hybrid*
Number of total sequences	55	71	126
Total length (bp)	1339	1339	1339
GC content (%)	31	30.1	30.4
Conserved sites	1330	1314	1281
Variable sites	7	23	56
Parsimony informative sites	3	18	49
Transitional pairs	45.69	53.47	50.71
Transversional pairs	54.31	46.53	49.29
Transition/Transversion (tr/tv) (R) ratio	0.75	1.01	0.91
Number of indels	0	0	0
Nucleotide diversity	0.002	0.006	0.017

**S. alba x fragilis* was added to analyze as a hybrid species.

3.3.1.2. *mat*K (maturase Kinase)

The complete region of the *mat*K was sequenced. After alignment of *Salix* samples' sequences, the whole length of *mat*K region was determined as 1731 bp including some part of *trn*K intron with indels. The *mat*K sequence alignment from one herbarium species, *Salix myrsinifolia*, was not achieved and excluded from *mat*K gene analysis as an exception. Since the length of region is long, the variable sites of the region in both subgenera were high. Most of them are parsimony informative. The variation is higher at the sequence of 5' region (*mat*K1) than sequence adjacent to at the 3'(*mat*K2) ends. Besides, since the conserved 3'end has less but informative variable sites, it is useful in resolving phylogeny according to Hilu and Liang (1997). The most discriminative substitutions were illustrated for given Table 3.4. GC content (%) of each subgenera were nearly the same (about 35 %). The transition and transversion substitution rates were estimated as 60.78, 39.22, respectively. The rate (R) was calculated as 1.36 for total Turkish *Salix* species. Only one insertion (*S. alba*) cases was found in the current data of subg. *Salix* for cpDNA *mat*K region. As far as the nucleotide diversity is concerned, *mat*K region of *Salix* species revealed by

the least diversity (0.006) among other gene regions. Nucleotide diversity was detected higher in subg. *Vetrix* than subg. *Salix*.

Table 3.4. The estimated molecular diversity parameters based on *mat*K (maturase Kinase) cpDNA region of Turkish *Salix* species data

	Subgenus *Salix*	Subgenus *Vetrix*	Over all (Subg. *Salix*, *Vetrix*)
Number of species	7+1 hybrid*	16	23+1 hybrid*
Number of total sequences	53	60	113
Total length (bp)	1731	1731	1731
GC content (%)	35.5	35.5	32.4
Conserved sites	1715	1687	1683
Variable sites	13	40	45
Parsimony informative sites	12	27	33
Transitional pairs	43.22	65.1	60.78
Transversional pairs	56.78	34.9	39.22
Transition/Transversion (tr/tv) (R) ratio	0.67	1.64	1.36
Number of indels	1	0	1
Nucleotide diversity	0.003	0.004	0.006

**S. alba x fragilis* was added to analyze as a hybrid species.

3.3.1.3. *rbc*L

The length of the cpDNA *rbc*L region was 1480bp. Since it is really hard to amplify the long DNA gene region from *Salix* species, 2 primers were designed to analyze *rbc*L region with two small fragments (Savolainen *et al*., 2000). The molecular diversity parameters were estimated by using the combined sequence data from two fragments (Table 3.5). The low level of variable sites was observed in cpDNA *rbc*L region for Turkish *Salix* species (22 variable sites). Only 3 substitutions were observed among Subg. *Salix* members. Most of them were parsimony informative. No significant difference was found with respect to GC content among subgenera. The substitution rate was contributed by transitional in subg. *Vetrix* with 59.78. The R value was also high for this subgenus. Indel was absent in *rbc*L gene region of Turkish *Salix*. The nucleotide diversity was estimated as 0.005 which is lower than other cpDNA regions studied. However, nucleotide diversity was detected higher in subg. *Vetrix* than subg. *Salix*.

Table 3.5. The estimated molecular diversity parameters based on *rbc*L cpDNA region of Turkish *Salix* species data

	Subgenus *Salix*	Subgenus *Vetrix*	Over all (Subg. *Salix*, *Vetrix*)
Number of species	7+1hybrid*	17	24+1hybrid*
Number of total sequences	54	67	121
Total length (bp)	1480	1480	1480
GC content (%)	43.7	43.2	43.4
Conserved sites	1476	1460	1457
Variable sites	3	19	22
Parsimony informative sites	1	8	15
Transitional pairs	33.68	59.78	56.72
Transversional pairs	66.32	40.22	43.28
Transition/Transversion (tr/tv) (R) ratio	0.50	1.46	1.29
Number of indels	0	0	0
Nucleotide diversity	0.001	0.003	0.005

**S. alba x fragilis* was added to analyze as a hybrid species.

3.3.1.4. *ITS* (Internal transcribed spacer) Region

ITS nuclear region was amplified with only one primer pairs (*ITS*L-4). The total length of *ITS* region was found to be about 657 bp after alignment of all sequences from Turkish *Salix* species. The beginning and the end of the sequences were trimmed during the alignment process in order to exclude unreliable readings. Molecular diversity parameters were estimated for total and two subgenera of studied *Salix* species. GC content of *ITS* region was high 64.5%. High level of GC is the indicator of high genomic variation in DNA sequences (Serres Giardi *et al*., 2012). Thus, it can be said that *ITS* region was quite diverse and informative. As far as the number of variable sites are concerned, subg. *Vetrix* had higher diversity than subg. *Salix*. Most of the variable sites were responsible from divergence of two genuses, which are tabulated in Table 3.6. Particularly, the parsimony informative sites which range from 57 to 564 bp determines the separation of two subgenera. Transition and transversion rates as well as R value were quite high such as 74.88% for transition, 25.12% for transversion and 2.73 for transition/transversion bias (R). A remarkable transition number shows that variable sites in *ITS* are the results of purine or pyrimidine changes. One indel was observed in only subg. *Salix*. Total nucleotide diversity was estimated as 0.013 which was greatly influenced by nucleotide

variation in subg. *Vetrix*. Thus, it is clear that nucleotide diversity higher in subg. *Vetrix* than subg. *Salix*.

Table 3.6. The estimated molecular diversity parameters based on *ITS* nrDNA region of Turkish *Salix* species data

	Subgenus *Salix*	Subgenus *Vetrix*	Over all (Subg. *Salix*, *Vetrix*)
Number of species	7+1 hybrid*	17	24+1 hybrid*
Number of total sequences	53	64	117
Total length (bp)	657	657	657
GC content (%)	65.1	64.3	64.5
Conserved sites	654	642	640
Variable sites	2	15	17
Parsimony informative sites	-	9	14
Transitional pairs	52.46	76.68	74.88
Transversional pairs	47.54	23.32	25.12
Transition/Transversion (tr/tv) (R) ratio	1.0	3.02	2.73
Number of indels	1	0	1
Nucleotide diversity	0.001	0.006	0.013

**S. alba x fragilis* was added to analyze as a hybrid species.

3.4. Genetic Distance within and among Subgenera of *Salix* sp.

Genetic distance among species as well as among subgenera were evaluated by using sequences of three cpDNA regions (*trn*T-F, *mat*K and *rbc*L) and one nrDNA region (*ITS*). 126 nucleotide sequences for *trn*T-F, 113 nucleotide sequences for *mat*K, 121 nucleotide sequences for *rbc*L and 117 nucleotide sequences for *ITS* region were utilized to estimate an average evolutionary divergence within species/subgenus. Genetic divergence data among subgenera was given in Table 3.7. Overall genetic divergence and genetic divergence among subgenera were analyzed by selecting the best model. To get the best model, the model selection analysis was done for every region according to the lowest BIC (Bayesian information criterion) value. The evolutionary divergence rates among sites as gamma correction value (G) and certain fraction of sites as invariability correction value (I) were calculated in every model. These substitution models differ in regards to the parameters used to define the rates in which one nucleotide substituted for another during evolution. *HKY;*Hasegawa-Kishino-Yano was used for *trn*T-F region. The *HKY* is statistical method which is

preferred widely for estimating divergence dates of species from DNA sequence data by a molecular clock approach (Hasegawa, Kishino and Yano, 1980). T92; Tamura 3-parameters is a statistical method improved to estimate the number of substitutions per site between two DNA sequences (Tamura, 1992). The last model was used for the rest of them as a best fitting nucleotide substitution model.

Table 3.7. Genetic divergence among subgenera based on four studied gene regions (The Gama/Invariability corrections were obtained by model selection analysis)

	Subg. Salix	***Subg. Vetrix***	***Overall***	***Model***	***Gama/ Invariability Correction (I, G)***	***Regions***
Subg. *Salix*	0.002	0.031	***0.017***	***HKY***	0.05, 0.48	***trn* T-F**
Subg. *Vetrix*	0.006					
Subg. *Salix*	0.003	0.004	*0.006*	***T92***	0.05, 0.49	***mat*K**
Subg. *Vetrix*	0.009					
Subg. *Salix*	0.001	0.003	*0.005*	***T92***	0.05, 0.49	***rbc*L**
Subg. *Vetrix*	0.009					
Subg. *Salix*	0.001	0.006	***0.013***	***T92***	0.05, 0.49	***ITS***
Subg. *Vetrix*	0.023					

Genetic divergence among subgenera was provided by using number of substitutions per site from averaging over all sequence pairs. Besides, to evaluate overall divergence, the whole Turkish *Salix* data were used (Table 3.7). When overall divergence is considered, *trn*T-F (0.017) and *ITS* (0.013).were the most divergent regions. The highest genetic divergence value within the subgenera was in non-coding cpDNA region *trn*T-F as 0.031 whereas the lowest in coding *rbc*L region (0.003). As far as the divergence distance between the subgenera is considered, the largest value calculated for nrDNA *ITS* (0.023) and the lowest for cpDNA *trn*T-F region (0.006). Thus, *ITS* is the most discriminative gene region between two subgenera of *Salix* sp. Divergence within Subg. *Vetrix* was greater than that is in Subg. *Salix* for all regions.

The genetic distances among species in each subgenus were also estimated in order to understand the relationships (Table 3.8). The high divergence in all regions for Subgenus *Salix* is due to *S. acmophylla* and *S. triandra* subsp. *bornmuelleri* species. Especially, *S. acmophylla* was genetically diverged from other species in nrDNA *ITS* (0.003), cpDNA *trn* T-F (0.004) and *rbc*L (0.001) regions. In subg. *Vetrix* the species, *S. amplexicaulis* was determined as the most diverged taxa nearly for all

studied regions with the high values. Subg. *Vetrix* members; *S. wilhelmsiana* (cpDNA *trn*T-F) *S. rizeensis (*cpDNA *trn* T-F*), S. armenorossica (*cpDNA *trn* T-F*, rbc*L*) S. caprea* (cpDNA *rbc*L*,* nrDNA *ITS)* and *S. cinerea* (cpDNA *rbc*L*,* nrDNA *ITS), S. apoda (trn*T-L*)* species pairs were genetically the most distant combinations (Table 3.8).

Table 3.8. Genetic divergence of Turkish *Salix* sp. between species within each subgenus based on four studied regions

Subgenus	The most divergent species combinations within Subgenus	Distance	Region
Salix	*S. acmophylla-S. fragilis*	0.004	***trn*T-F**
	*S. acmophylla-S. alba x fragilis**	0.004	
	S. acmophylla- S. pentandroides	0.005	
Vetrix	*S. apoda-S. amplexicaulis*	0.010	
	S. armenorossica- S.wilhelmsiana	0.011	
	S. caucasica-S. rizeensis	0.011	
Salix	*S .alba-S. triandra* subsp. *bornmuelleri*	0.006	***mat*K**
	S. alba x fragilis- S. triandra subsp. *bornmuelleri*	0.005	
	S. alba x fragilis - S. triandra subsp. *bornmuelleri*	0.004	
Vetrix	*S. amplexicaulis-S. apoda*	0.011	
	S. amplexicaulis -S. pupurea subsp. *leucodermis*	0.010	
	S. amplexicaulis -S. elaeagnos	0.010	
Salix	*S. alba-S. acmophylla*	0.001	***rbc*L**
	*S. acmophylla-S. albaxfragilis**	0.001	
	S. acmophylla-S. triandra subsp. *triandra*	0.001	
Vetrix	*S. amplexicaulis -S. caprea*	0.010	
	S. amplexicaulis -S. cinerea	0.011	
	S. amplexicaulis -S. armenorossica	0.010	
Salix	*S. acmophylla-S. fragilis*	0.003	***ITS***
	S. acmophylla-S. babylonica	0.003	
	S. acmophylla-S. pentadroides	0.003	
Vetrix	*S. amplexicaulis-S. caprea*	0.012	
	S. amplexicaulis-S. cinerea	0.012	
	S. amplexicaulis-S. pseudomedemii	0.012	

*It was added to analysis as a hybrid species.

3.5. Haplotype Frequency Analysis, Analyses of Molecular Variance (AMOVA), Estimation of *F*st Value

The relative observed haplotype frequencies were obtained and provided as a haplotype tree using DNAsp software program (Librado and Rozas, 2009) (Figure

3.5 to Figure 3.7). Since the *trn*T IGS and *trn*F IGS regions for the world *Salix* sp. were not available in NCBI data base, they were excluded in comparative analysis. Furthermore, *trnL* intron sequence from New World *Salix* species could not be found, therefore they were not evaluated in this analysis.

There were 13 haplotypes based on *trn*L region analysis (Table 3.9). The haplotype number is the same as 13 when the *rbc*L region was analyzed. The results for *mat*K region indicated that there were 24 haplotypes in Turkish *Salix* sp. obtained from Turkey. It is found that not many Turkish *Salix* species shared the same haplotype in DNAsp analysis. Apparently, the haplotype number is so close to species number, especially for *mat*K region. All haplotype trees gave the similar results separating haplotypes into two clusters (Figure 3.5 to Figure 3.7). According DNAsp results, the high haplotype diversity were found in cpDNA *mat*K region (Table 3.9).

When Analysis of Molecular Variance (AMOVA) (Weir and Cockerham, 1984; Excoffier *et al*., 1992; Wier, 1996) analysis were performed, Turkish *Salix* species including both subgenera, Old World and New World *Salix* species were accepted as groups in the analysis.

When the comparison is done only for Turkish species, haplotype diversity among subgenera was higher than within subgenus based on all regions; *trn* L (82.15), *mat*K (63.71) and *rbc*L (77.70) (Table 3.11). Particularly, *trn*L results indicated that high level of diversity was found among Turkish subgenera rather than within subgenus. This was evident from a significant F_{st} value (0.821).

In the world analyses, data from *trn*L and *rbc*L regions indicated that Turkish *Salix* species were diverged from World *Salix* species with high haplotype diversity (74.58, 69.50). Both these cpDNA gene regions are discriminative in comparing Turkish and World willows. Besides that, the variation within the groups for encoded cpDNA *mat*K region was greater than among groups.

Pairwise F_{st} values were analyzed according to Tamura and Nei method (1993) (Table 3.12). All values based on 3 cpDNA regions were estimated within and among Turkish *Salix* Subgenera, Old World and New World *Salix* species. There

were high significant variations (p<0.001***) both within and among groups based on AMOVA analysis. Regarding the pairwise *F*st values, the respectively lower *F*st values were revealed between NWS-Subg. *Vetrix* (0.139) with respect to *trn*L region and OWS-Subg. *Salix* (0.051) based *mat*K gene region. The significant *F*st value (0.821) was estimated between Turkish subgenera (*Salix* and *Vetrix*) in *trn* L region. OWS are close to subg. *Salix* (*F*st; 0.051) whereas NWS close to subg. *Vetrix* (*F*st; 0.139) for *mat*K region. In the *rbc*L gene region, OWS were significantly different from subg. *Salix* (*F*st value as 0.800).

In cpDNA *mat*K and *trn*L gene regions, the high variation within groups and differences between subgenera were resulted by contribution of OWS and NWS among Turkish *Salix* species.

Tajima's D test statistics (Tajima, 1989) is the hypothesis that the frequency spectrum of mutation is selectively neutral. A significantly negative Tajima's D is expected when the data depart from neutral expectations. In the current data set, all Tajima's D results generated in Arlequin Software and all of them didn't reveal significant deviations (Table 3.10). High significance is the demonstration of an excess of low-frequency polymorphism which is indicative of a non-random process (Percy *et al*., 2014) which was not observed in the current data set.

Table 3.9. The haplotype number and diversity of Turkish *Salix* sp. according to studied cpDNA region

Region	Number of Haplotype	Haplotype diversity
***trn*L**	13	0.892
***mat*K**	24	0.996
***rbc*L**	13	0.855

Table 3.10. Results of Tajima's D test statistic (Tajima 1989) which test the hypothesis that the frequency spectrum of mutation is selectively neutral

Gene region	**Data Set**	**S** (number of polymorphic site)	**Pi** (nucleotide diversity)	**Tajima's D** (significance rejection of neutrality)
*trn*L	Subg. *Salix*	2	0.44	0.095 NS
	Subg. *Vetrix*	10	3.54	0.784 NS
	OWS	12	7.0	0.955 NS
*mat*K	Subg. *Salix*	2	0.6	0.292 NS
	Subg. *Vetrix*	16	2.5	0.016 NS
	NWS	16	5.2	0.326 NS
	OWS	2	0.8	0.095 NS
*rbc*L	Subg. *Salix*	3	1.0	0.379 NS
	Subg. *Vetrix*	16	3.02	0.067 NS
	NWS	9	6.0	1.000 NS
	OWS	16	5.3	0.016 NS

NS, P>0.10; *, P<0.05; **, P<0.01; ***, P<0.001.

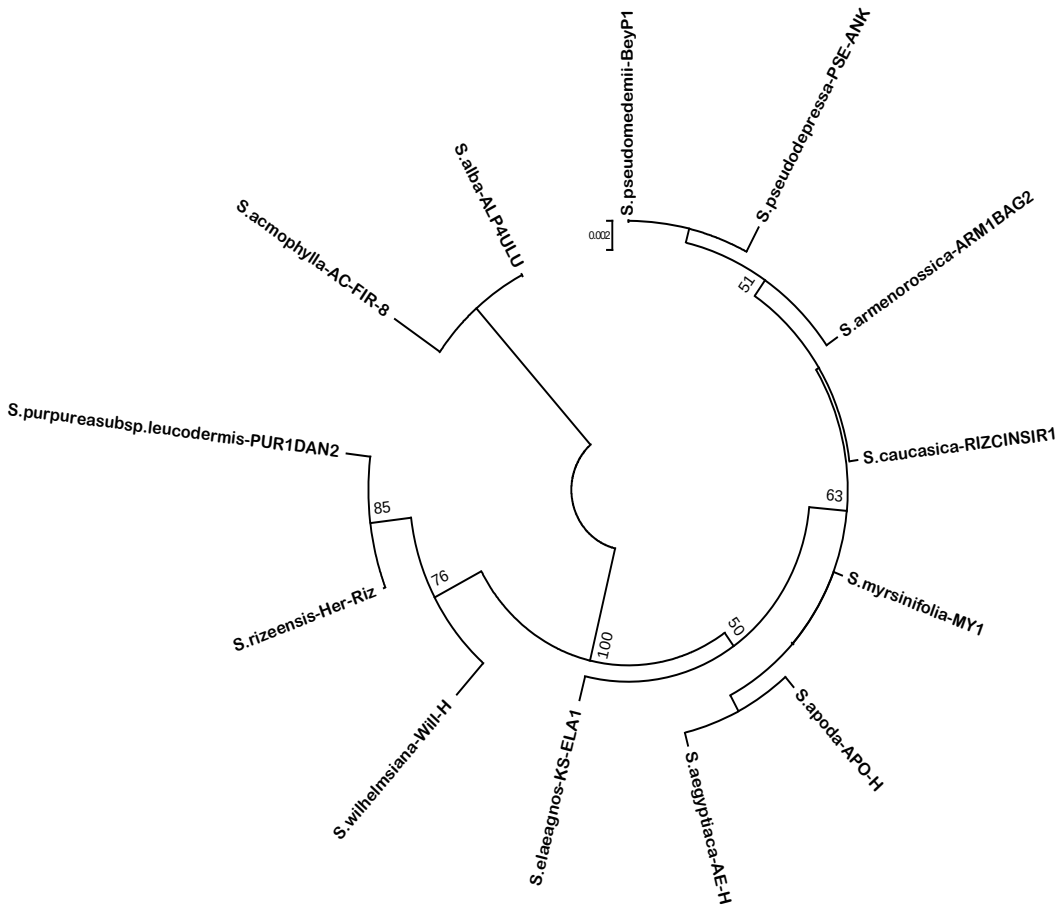

Figure 3.5. Haplotype tree showing the relationship of Turkish *Salix* sp. based on *trn*L region. Only bootstrap values greater than 50% were shown.

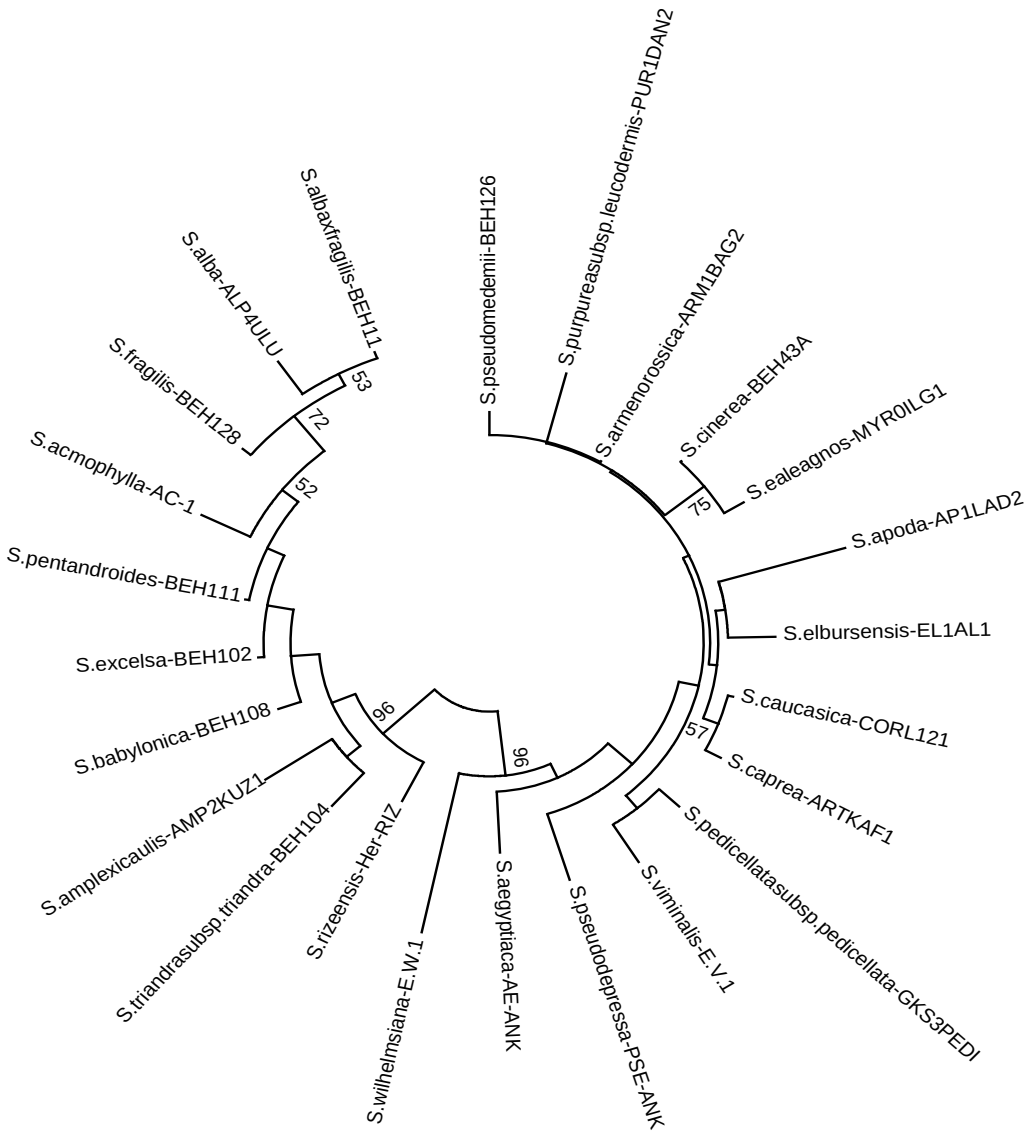

Figure 3.6. Haplotype tree showing the relationship of Turkish *Salix* sp. based on *mat*K region. Only bootstrap values greater than 50% were shown.

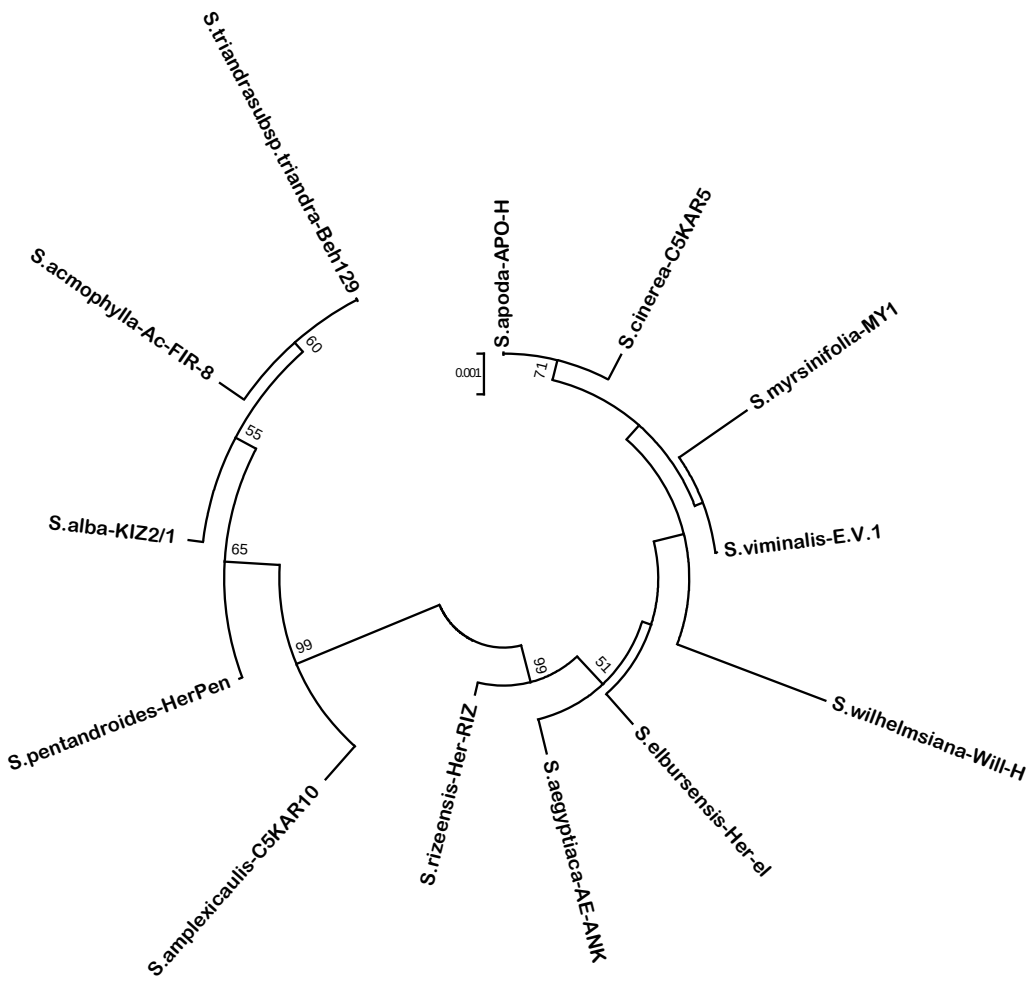

Figure 3.7. Haplotype tree showing the relationship of Turkish *Salix* sp. based on *rbc*L region. Only bootstrap values greater than 50% were shown.

Table 3.10. Analysis of Molecular Variance (AMOVA) based on three cpDNA regions of Turkish *Salix* sp. and the world *Salix* sp.

Source of Variation	Degrees of Freedom	Sum of Squares	Variance Components	Percentage of Variation	*F*st
Turkish *Salix* sp. for *trn* L region					
Among Subgenera	1	69.254	5.77765 Va	82.15	0.82149
Within Subgenus	24	30.131	1.25545 Vb	17.85	
Total	25	99.385	7.03309		
Turkish and World *Salix* sp. for *trn*L region					
Among Groups	2	115.385	6.09228 Va	74.58	0.74584
Within Groups	28	58.131	2.07610 Vb	25.42	
Total	30	173.516	8.16838		
Turkish *Salix* sp. for *mat*K region					
Among Subgenera	1	64.817	5.02870 Va	63.71	0.63715
Within Subgenus	23	65.868	2.86382 Vb	36.29	
Total	24	130.685	7.89253		
Turkish and World *Salix* sp. for *mat*K region					
Among Groups	3	31.469	0.99628 Va	43.19	0.43191
Within Groups	35	45.864	1.31040 Vb	56.81	
Total	38	77.333	2.30668		
Turkish *Salix* sp. for *rbc*L region					
Among Subgenera	1	61.174	5.07411 Va	77.70	0.77705
Within Subgenus	24	34.941	1.45588 Vb	22.30	
Total	25	96.115	6.52999		
Turkish and World *Salix* sp. for *rbc*L region					
Among Groups	3	85.517	3.49627 Va	69.50	0.69498
Within Groups	31	47.569	1.53447 Vb	30.50	
Total	34	133.086	5.03074		

Table 3.11. Genetic distance among samples among Turkish *Salix*, NWS and OWS species (pairwise *F*st values). Estimated by Tamura and Nei Distance Method (*F*st p values < 0.0050) (NWS: New World *Salix*, OWS:Old World *Salix*)

Subgenus *Salix*				
Subgenus *Vetrix*	0.82149 0.66187 0.77705			
NWS	* 0.28000 0.55372	* 0.13973 0.75170		
OWS	0.65572 0.05186 0.80000	0.72364 0.63428 0.37048	* 0.21134 0.55126	
	Subgenus *Salix*	**Subgenus *Vetrix***	**NWS**	**OWS**

***trn* L region**

***mat*K region**

***rbc*L region**

*Data was not available.

3.6. Construction of Phylogenetic Trees Based on Studied Regions

After the final analysis, the phylogenetic tree was constructed by using MEGA 7.0.2 Software for all Turkish *Salix* species and two subgenera. Deleted fractions and several substitutions were recognized in the aligned DNA sequences of both cpDNA and nrDNA regions. Tables 3.13 to Table 3.16 provide some symbols which were used to indicate indels and substitutions. In every table, nucleotide absence was demonstrated as a dot, inserted (deleted) base as green and substitution as red color. Numbers on the top of columns show the position of the corresponding nucleotide.

3.6.1. *trn*T-F

The constructed phylogenetic tree based on *trn*T-F gene region give rise to 2 clusters which were previously established morphologically subgeneric division; Subg. *Salix* and Subg. *Vetrix*. There were many substitutions that allowed the separation of two subgenera. There were 10 substitutions in *trn*T-L part, 18 substitutions in *trn*L part and 11 substitutions in *trn*L-F (Table 3.13). There wasn't any recorded indel in *trn*T-F gene region.

In first clade; Subg. *Salix* group, it was clear that the first subclade included members of *Salix* section; *S. alba, S. excelsa, S. fragilis, S. triandra* subsp. *triandra* and *S. triandra* subsp. *bornmuelleri* (Figure 3.8). The *S. acmophylla* located at the top of the subg. *Salix* tree in which a few substitution caused its location. The hybrid species (*S. alba x fragilis)* was positioned near to *S. alba* and *S. fragilis* as expected. Thus, Subg. *Salix* clade (Figure 3.9) consists of large polytomy with 7 taxa. Other species; *S. babylonica* and *S. pentandroides* joined to subclade externally as sister species.

In the second clade; Subg. *Vetrix* group; there is no uniformity in clustering (Figure 3.10). But there were still two main sub-clades supported with the moderate bootstrap value as 56 and 67. The first sub-clade of Subg. *Vetrix* consists of three groups; the first one including *S. pseudomedemii*, *S. armenorossica*, *S. pseudodepressa* and *S. viminalis*, the second one with *S. cinerea*, *S. caucasica* and *S. pedicellata* subsp. *pedicellata*, and the third group with *S. myrsinifolia*, *S. aegyptiaca*, *S. apoda* and *S. caprea*. The most close taxa were *S. apoda* and *S. caprea* based on substitutions of the 434, 827 and 978th bp positions. In the second sub-clade of subg. *Vetrix*, *S. wilhelmsiana*, *S. rizeensis* and *S. purpurea* subsp. *leucodermis* were located. *S. rizeensis* and *S. purpurea* subsp. *leucodermis* based on substitutions of 851, 874 and 877th bp positions were the closest taxa. *S. elaeagnos*, *S. elbursensis* were connected to these two sub-clades of Subg. *Vetrix* from external as a small clade. *S. amplexicaulis* was the most distant taxa to subg. *Vetrix*.

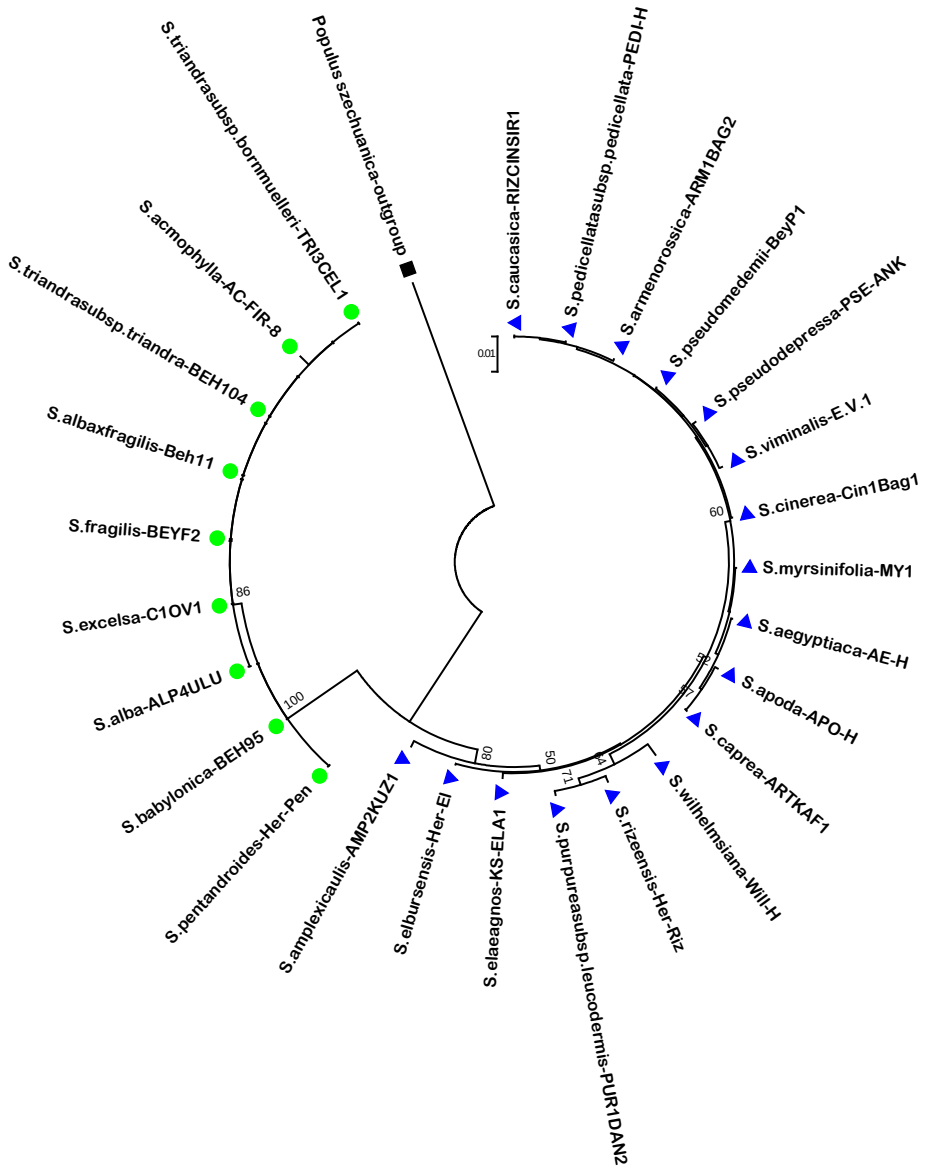

Figure 3.8. Phylogenetic tree, constructed by using the sequences from *trn*T-F region of cpDNA (Numbers on tree branches: bootstrap values, ●: Subgenus *Salix*, ◀: Subgenus *Vetrix,* ■; Outgroup).

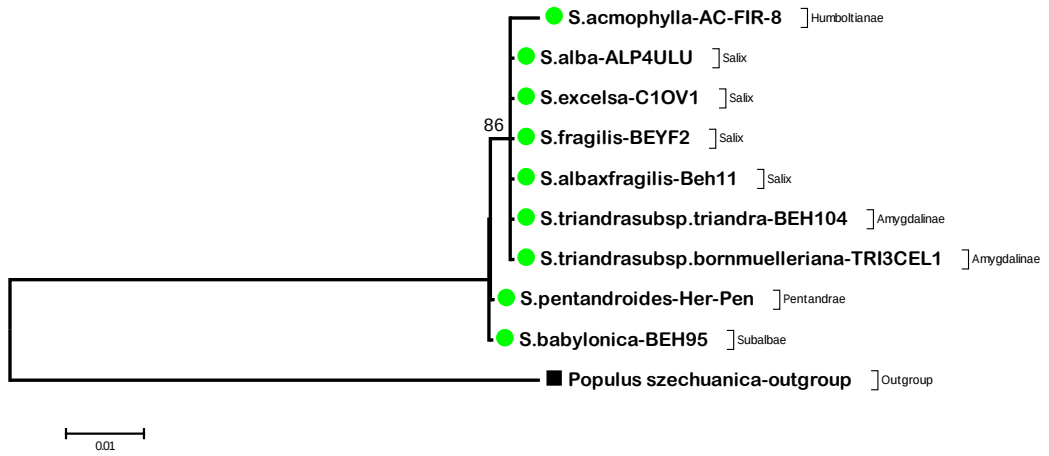

Figure 3.9. Phylogenetic tree, constructed by using the sequences from *trn*T-F region of cpDNA, showing detail relationships among the members of Subg. *Salix* and the section of species (Numbers on tree branches: bootstrap values, ●: Subg. *Salix*, ■; Outgroup, Parenthesis: section).

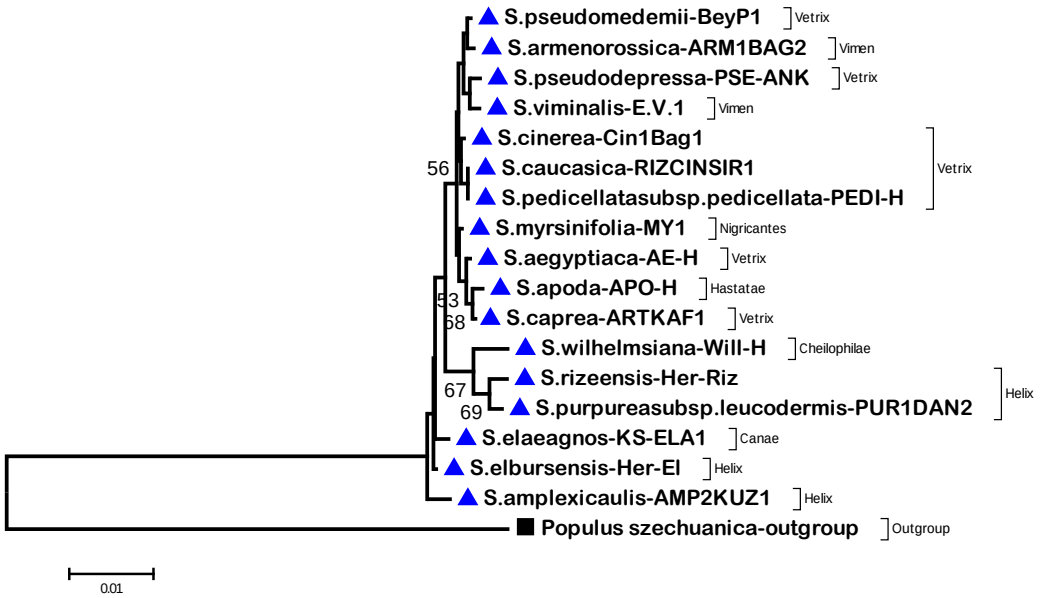

Figure 3.10. Phylogenetic tree, constructed by using the sequences from *trn*T-F region of cpDNA, showing detail relationships among the members of Subg. *Vetrix* and the section of species (Numbers on tree branches: bootstrap values, ◀: Subg. *Vetrix,* ■; Outgroup, Parenthesis: section).

Table 3.12. The regions of indels and nucleotide substitutions in the DNA sequences of *trn*T-F of cpDNA region

Subgenus	Sections	Species name	*Nucleotide Positions in the Aligned DNA sequences of trn*T-F *Region*																																														
			trn T IGS														*trn* L intron																															
			2	142	193	196	226	243	293	309	320	320	434	472	481	482	527	641	667	696	715	729	739	768	782	787	790	824	827	831	833	851	874	877	880	978	1010	1034	1079	1093	1149	1165	1172	1183
	Humboltianae	*S. acmophylla*	C	C	T	G	A	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	A	A	T	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
	Amygdalinae	*S. triandra* subsp. *triandra*	C	C	T	G	T	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	G	A	A	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
	Amygdalinae	*S. triandra* subsp. *bornmuelleri*	C	C	T	G	T	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	G	A	A	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
SALIX	**Pentandrae**	*S. pentandroides*	C	C	T	G	T	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	G	A	A	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
	Salix	*S. alba*	C	C	T	G	T	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	G	A	A	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
	Salix	*S. excelsa*	C	C	T	G	T	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	G	A	A	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
	Salix	*S. fragilis*	C	C	T	G	T	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	G	A	A	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
	Salix	*S. alba x fragilis*	C	C	T	G	T	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	G	A	A	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
	Subalbae	*S. babylonica*	C	C	T	G	T	T	C	T	T	C	G	T	C	T	G	C	C	C	A	T	G	A	A	C	A	C	A	C	T	T	T	G	T	A	G	T	G	G	G	A	A	A
	Hastatae	*S. apoda*	T	A	C	T	T	C	T	G	G	A	A	T	T	T	G	T	T	T	C	G	G	A	A	T	C	T	T	A	G	T	T	T	T	T	G	C	T	G	T	G	G	G
	Nigricantes	*S. myrsinifolia*	T	A	C	T	T	C	T	G	G	A	A	T	T	T	G	T	T	T	C	G	G	A	A	T	C	T	T	T	G	T	T	T	T	A	A	C	T	G	T	G	G	G
	Vetrix	*S. caucasica*	T	A	C	T	T	C	T	G	G	A	G	T	T	T	G	T	T	T	C	G	G	A	A	T	C	T	T	A	G	T	T	T	T	A	A	C	T	G	T	G	G	G
	Vetrix	*S. pedicellata* subsp. *pedicellata*	T	A	C	T	T	C	T	G	G	A	G	T	T	T	G	T	T	T	C	G	G	A	A	T	C	T	T	A	G	T	T	T	T	A	A	C	T	G	T	G	G	G
	Vetrix	*S. caprea*	T	A	C	T	T	C	T	G	G	A	A	T	T	T	G	T	T	T	C	G	G	A	A	T	C	T	T	A	G	T	T	T	T	T	A	C	T	G	T	G	G	G
	Vetrix	*S. aegyptiaca*	T	A	C	T	T	C	T	G	G	A	A	T	T	T	G	T	T	T	C	G	G	A	A	T	C	T	T	T	G	T	T	T	T	T	A	C	T	G	T	G	G	A
	Vetrix	*S. cinerea*	T	A	C	T	T	C	T	G	G	A	G	T	T	T	G	T	T	T	C	G	G	A	A	T	C	T	T	T	G	T	T	T	T	A	A	C	T	G	T	G	G	G
	Vetrix	*S. pseudomedemii*	T	A	C	T	T	C	T	G	G	A	G	T	T	T	G	T	T	T	C	G	G	C	A	T	C	T	T	T	G	T	T	T	T	A	A	C	T	G	T	G	G	G
VETRIX	**Vetrix**	*S. pseudodepressa*	T	A	C	T	T	C	C	G	G	A	A	T	T	T	G	T	T	T	C	G	G	C	A	T	C	T	T	T	G	A	T	T	T	A	A	C	T	G	T	G	G	G
	Vimen	*S. viminalis*	T	A	C	T	T	C	C	G	G	A	A	T	C	T	G	T	T	T	C	G	G	C	A	T	C	T	T	T	G	T	T	T	T	A	A	C	T	A	T	G	G	A
	Vimen	*S. armenorossica*	T	A	C	T	T	C	T	G	G	A	G	T	T	T	G	T	T	T	C	G	G	C	A	T	C	T	T	A	G	T	T	T	T	A	A	C	T	G	T	G	G	G
	Canae	*S. elaeagnos*	T	A	C	T	T	C	T	G	G	A	A	A	T	T	G	T	T	T	A	G	G	A	A	T	A	T	T	A	G	T	T	T	T	A	A	C	T	G	T	G	G	A
	Helix	*S. elbursensis*	T	C	C	T	T	C	T	G	G	A	G	T	C	T	G	T	T	T	A	G	G	A	A	T	A	T	T	A	G	T	T	T	T	A	A	C	T	G	T	G	G	G
	Helix	*S. amplexicaulis*	T	C	C	T	T	C	C	G	G	A	G	T	T	G	G	T	T	T	A	G	G	A	A	T	A	T	T	A	G	T	T	T	T	A	A	T	G	A	T	G	G	A
	Helix	*S. rizeensis*	T	C	C	T	T	C	C	G	G	A	G	T	C	T	G	T	T	T	A	G	G	A	A	T	A	T	T	T	G	T	A	A	A	T	A	T	T	A	T	G	G	A
	Helix	*S. purpurea* subsp. *leucodermis*	T	A	C	T	T	C	C	G	G	A	G	T	T	T	G	T	T	T	A	G	G	A	A	T	A	T	T	T	G	A	A	A	A	T	A	C	T	G	T	G	G	G
	Cheilophilae	*S. wilhelmsiana*	T	A	C	T	T	C	C	G	G	A	A	T	C	T	A	T	T	T	C	G	G	A	A	T	A	T	T	T	G	T	A	A	A	A	A	C	G	A	T	G	G	A

Table 3.13. Continued.

Subgenus	Sections	Species name	*trn*L-F *IGS*													
			1188	1233	1238	1248	1269	1272	1276	1279	1280	1287	1289	1294	1297	1310
SALIX	**Humboltianae**	*S. acmophylla*	G	A	G	A	C	A	G	C	G	C	C	G	G	G
	Amygdalinae	*S. triandra* subsp. *triandra*	G	A	G	A	C	A	G	C	C	C	C	G	G	G
	Amygdalinae	*S. triandra* subsp. *bornmuelleri*	G	A	G	A	C	A	G	C	C	C	C	G	G	G
	Pentandrae	*S. pentandroides*	G	A	A	A	C	G	G	C	C	C	C	G	G	G
	Salix	*S. alba*	G	A	G	A	C	A	G	C	C	C	C	G	G	G
	Salix	*S. excelsa*	G	A	G	A	C	A	G	C	C	C	C	G	G	G
	Salix	*S. fragilis*	G	C	G	A	C	A	G	C	C	C	C	G	G	G
	Salix	*S. alba x fragilis*	G	C	G	A	C	A	G	C	C	C	C	G	G	G
	Subalbae	*S. babylonica*	G	A	A	A	C	G	G	C	C	C	C	G	G	G
VETRIX	**Hastatae**	*S. apoda*	G	A	A	T	T	G	T	T	C	G	T	A	T	T
	Nigricantes	*S. myrsinifolia*	G	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vetrix	*S. caucasica*	G	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vetrix	*S. pedicellata* subsp. *pedicellata*	G	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vetrix	*S. caprea*	G	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vetrix	*S. aegyptiaca*	G	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vetrix	*S. cinerea*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vetrix	*S. pseudomedemii*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vetrix	*S. pseudodepressa*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vimen	*S. viminalis*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Vimen	*S. armenorossica*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Canae	*S. elaeagnos*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Helix	*S. elbursensis*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Helix	*S. amplexicaulis*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Helix	*S. rizeensis*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Helix	*S. purpurea* subsp. *leucodermis*	T	A	A	T	T	G	T	T	C	G	T	A	T	T
	Cheilophilae	*S. wilhelmsiana*	T	A	A	T	T	G	T	T	C	G	T	G	T	T

3.6.2. *mat*K

The *mat*K gene region based phylogenetic tree supported two main clusters; Subg. *Salix* and Subg. *Vetrix* (Figure 3.11). Two members of Subg. *Vetrix* ; *S. rizeensis* and *S. amplexicaulis* were placed in the middle of the first cluster (Subg. *Salix* cluster). The substitutions at 370, 533, 780 and 1634th bp positions were responsible for changes in their locations (Table 3.14). Different from these, it is obvious that Subg. *Salix* comprised of two sub-clades (Figure 3.12). The first sub-clade includes the species; *S. alba, S. fragilis*, *S. alba x fragilis*, *S. acmophylla, S. excelsa, S. babylonica* and *S. pentandroides*. Section *Salix* members are together except *S. excelsa* species. The second sub-clade involved *S. amplexicaulis, S. triandra* subsp. *triandra and S. triandra* subsp. *bornmuelleri.* In the first sub-clade *S. alba, S. fragilis* and *S. alba x fragilis* show close relation whereas in the second, the close relation are seen between *S. triandra* subsp. *triandra* and *S. triandra* subsp. *bornmuelleri.* All members of subg. *Salix* species differ as a result of the substitutions at 509, 943, 1446, 1650, 1654, 1661 and 1666th bp positions of *mat*K1-2. *S. rizeensis* was attached to these clades externally and was located at the basal position of subg. *Salix* (Table 3.14).

Among subg. *Vetrix*, there was no distinct main sub-clades so most of them were nested a polytomy (Figure 3.13). In spite of that, there were weakly formed two sub-clades. S. *caprea, S. caucasica, S. apoda, S. elbursensis* and *S. psedomedemii* are in the first*, S. cinerea, S. elaeagnos, S. armenorossica, S. pseudodepressa* and *S. purpurea* subsp. *leucodermis* were located in the second sub-clade. *S. pedicellata* subsp. *pedicellata*, *S. wilhelmasiana*, *S. viminalis* and *S. aegyptiaca* generates externally connected to these sub-clades.

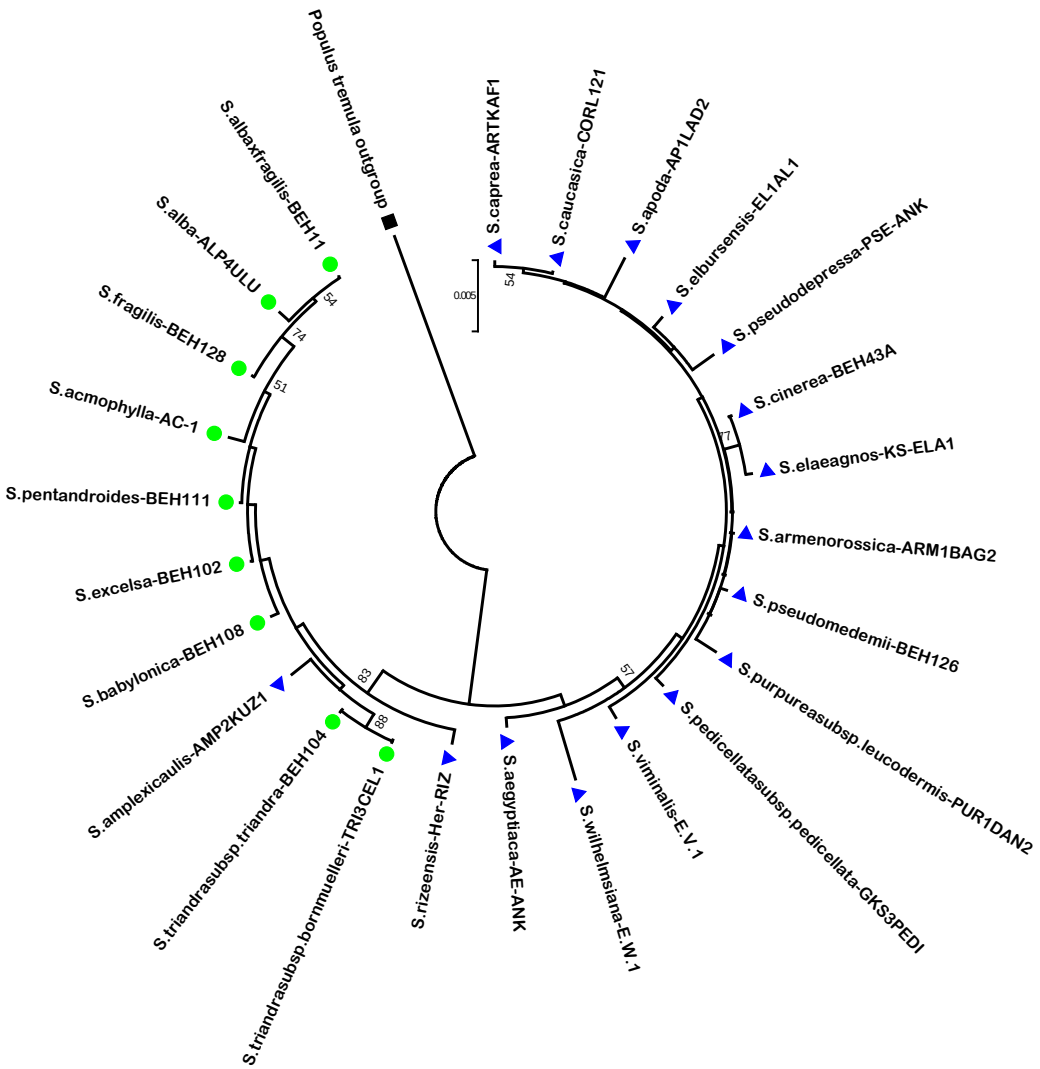

Figure 3.11. Phylogenetic tree, constructed using the sequences of *mat*K region of cpDNA (Numbers on tree branches: bootstrap values, ●: Subgenus *Salix*, ◀: Subgenus *Vetrix*, ■; Outgroup)

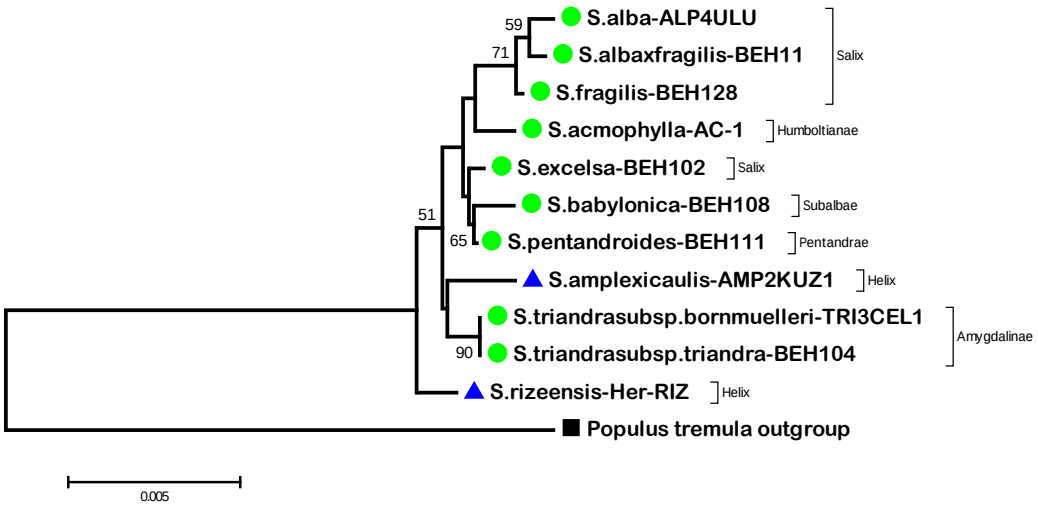

Figure 3.12. Phylogenetic tree, constructed using the sequences of *mat*K region of cpDNA, showing detail relationships among the members of Subgenus *Salix* the section of species (Numbers on tree branches: bootstrap values, ●: Subg. *Salix*, ◀:Subg. *Vetrix*, ■;Outgroup, Parenthesis: name of each Section)

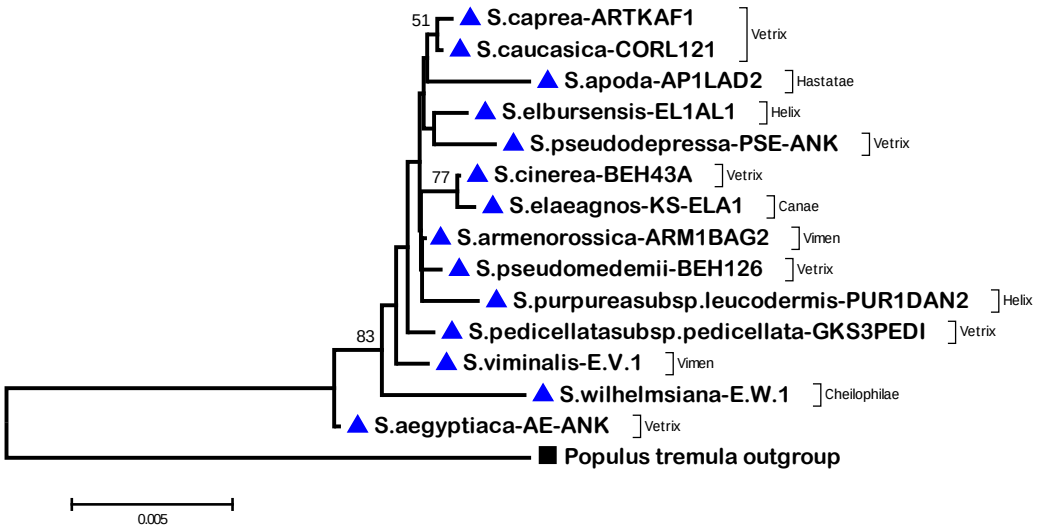

Figure 3.13. Phylogenetic tree, constructed using the sequences of *mat*K region of cpDNA, showing detail relationships among the members of Subg. *Vetrix* and the section of species (Numbers on tree branches: bootstrap values, ◀: Subg. *Vetrix,* ■; Outgroup, Parenthesis: name of each Section)

Table 3.14. The regions of indels and nucleotide substitutions in the DNA sequences of *mat*K cpDNA region.

Subgenus	Sections	Species name	12	16	138	169	183	282	276	370	440	449	450	463	509	533	573	658	679	728	780	796	797	827	877	916	943	978	1025	1040	1177	1192	1212	1226	1376	1438	1446	1474	1505	1572	1629	1634	1650
			Nucleotide Positions in the Aligned DNA sequences of matK Region																																								
			***mat*K 1**																													***mat*K 2**											
SALIX	**Humboltianae**	*S. acmophylla*	C	A	T	C	T	T	A	T	G	A	T	T	A	C	G	G	T	C	A	C	C	A	G	A	T	A	C	G	T	C	G	T	A	C	A	-	A	A	A	G	A
	Amygdalinae	*S. triandra* subsp. *triandra*	C	A	T	C	T	T	A	T	G	A	T	T	C	A	G	G	T	C	A	C	C	A	G	A	T	A	C	G	T	C	G	T	A	C	A	-	A	A	A	G	A
	Amygdalinae	*S. triandra* subsp. *bornmuelleri*	C	A	T	C	T	T	A	T	G	A	T	T	C	A	G	G	T	C	A	C	C	A	G	A	T	A	C	G	T	C	G	T	A	C	A	-	A	A	A	G	A
	Pentandrae	*S. pentandroides*	C	A	T	C	T	T	A	T	G	A	T	T	C	C	G	G	T	C	A	C	C	A	G	A	T	A	C	G	T	C	G	T	A	C	A	-	A	A	A	G	A
	Salix	*S. alba*	C	A	T	C	T	T	A	T	G	T	A	T	A	C	G	G	T	C	A	C	C	A	G	A	T	A	C	G	T	C	G	T	A	C	T	G	A	A	A	G	A
	Salix	*S. excelsa*	C	A	T	C	T	T	A	T	G	A	T	T	C	C	G	G	T	C	A	C	C	A	G	A	T	A	C	G	T	C	G	T	A	C	A	-	A	A	A	G	C
	Salix	*S. fragilis*	C	A	T	C	T	T	A	T	G	T	A	T	C	C	G	G	T	C	A	C	C	A	G	A	T	A	C	G	T	C	G	T	A	C	A	-	A	A	A	G	A
	Salix	*S. alba x fragilis*	C	A	T	C	T	T	A	T	G	T	A	T	A	C	G	G	T	C	A	C	C	A	G	A	C	A	C	G	T	C	G	T	A	C	A	-	A	A	A	G	A
	Subalbae	*S. babylonica*	C	A	T	C	T	T	A	T	G	A	T	T	C	C	G	G	T	C	A	C	C	A	G	A	T	A	C	G	T	C	G	T	G	C	A	-	A	A	A	A	A
VETRIX	**Hastatae**	*S. apoda*	C	A	T	C	T	A	A	T	A	A	T	C	C	C	T	G	T	C	A	A	C	C	G	G	C	A	C	G	A	C	G	T	G	T	A	-	A	A	G	A	C
	Vetrix	*S. caucasica*	C	A	T	C	T	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	T	C	G	A	C	A	C	G	T	C	G	T	G	T	A	-	A	G	A	A	C
	Vetrix	*S. pedicellata* subsp. *pedicellata*	C	A	T	T	T	T	A	T	A	A	T	C	C	C	T	G	G	C	A	A	T	C	G	A	C	A	C	G	T	C	G	T	G	T	A	-	A	A	A	A	C
	Vetrix	*S. caprea*	C	A	T	C	T	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	C	C	G	A	C	A	C	G	T	C	G	T	G	T	A	-	A	G	A	A	C
	Vetrix	*S. aegyptiaca*	C	A	T	C	T	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	T	A	G	A	C	A	C	G	T	C	G	T	G	C	A	-	A	A	A	A	C
	Vetrix	*S. cinerea*	C	G	T	C	T	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	T	C	G	A	C	A	C	G	T	A	G	T	G	T	A	-	A	A	A	A	C
	Vetrix	*S. pseudomedemii*	C	A	T	C	T	T	A	T	A	A	T	C	C	C	T	G	T	G	A	A	T	C	G	A	C	A	C	G	T	C	G	T	G	T	A	-	A	A	A	A	C
	Vetrix	*S. pseudodepressa*	C	A	T	C	T	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	T	C	G	A	C	A	G	G	T	C	A	T	A	C	A	-	A	A	A	A	C
	Vimen	*S. viminalis*	C	A	T	T	T	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	T	C	G	A	C	A	C	T	T	C	G	T	G	T	A	-	A	A	A	A	C
	Vimen	*S. armenorossica*	C	A	T	C	T	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	T	C	G	A	C	A	C	G	T	C	G	T	G	T	A	-	A	A	A	A	C
	Canae	*S. elaeagnos*	C	G	T	C	A	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	T	C	G	A	C	A	C	G	T	A	G	T	G	T	A	-	A	A	A	A	C
	Helix	*S. elbursensis*	C	A	T	C	A	T	A	T	A	A	T	C	C	C	T	G	T	C	A	A	C	C	G	A	C	A	C	G	T	C	G	T	A	T	A	-	A	A	A	A	C
	Helix	*S. amplexicaulis*	T	A	T	C	T	T	A	T	G	A	T	T	A	C	G	G	T	C	A	C	C	A	G	A	C	A	C	G	T	C	G	T	A	C	A	-	A	A	A	G	A
	Helix	*S. rizeensis*	C	A	T	C	T	T	A	T	G	A	T	C	C	C	G	G	T	C	A	C	C	A	G	A	T	G	C	G	T	C	G	T	A	C	A	-	A	A	A	G	C
	Helix	*S. purpurea* subsp. *leucodermis*	C	A	G	C	T	T	C	T	A	A	T	C	C	C	T	G	T	C	G	A	T	C	G	A	C	A	C	G	T	C	G	T	G	T	A	-	A	A	A	A	C
	Cheilophilae	*S. wilhelmsiana*	C	A	T	T	T	T	A	A	G	A	T	T	C	C	T	A	T	C	A	A	T	C	A	A	C	A	C	G	T	C	G	G	G	C	A	-	G	A	A	A	C

Table 3.14. continued

Subgenus	Sections	Species name	*mat*K2 1654	1661	1666	1674	1684
			Nucleotide Positions in the Aligned DNA sequences of matK Region				
	Humboltianae	*S. acmophylla*	A	G	A	A	G
	Amygdalinae	*S. triandra* subsp. *triandra*	C	A	G	A	C
	Amygdalinae	*S. triandra* subsp. *bornmuelleri*	C	A	G	A	C
SALIX	**Pentandrae**	*S. pentandroides*	C	G	A	A	C
	Salix	*S. alba*	T	G	A	A	C
	Salix	*S. excelsa*	C	G	A	A	C
	Salix	*S. fragilis*	T	G	A	A	C
	Salix	*S. alba x fragilis*	T	G	A	A	C
	Subalbae	*S. babylonica*	C	G	A	A	C
	Hastatae	*S. apoda*	C	G	G	G	C
	Vetrix	*S. caucasica*	C	A	G	G	C
	Vetrix	*S. pedicellata* subsp. *pedicellata*	C	A	G	G	C
	Vetrix	*S. caprea*	C	A	G	G	C
	Vetrix	*S. aegyptiaca*	C	G	G	A	G
	Vetrix	*S. cinerea*	C	A	G	G	C
VETRIX	**Vetrix**	*S. pseudomedemii*	C	A	G	G	C
	Vetrix	*S. pseudodepressa*	C	A	G	G	C
	Vimen	*S. viminalis*	C	A	G	A	C
	Vimen	*S. armenorossica*	C	A	G	G	C
	Canae	*S. elaeagnos*	C	A	G	G	C
	Helix	*S. elbursensis*	C	A	G	G	C
	Helix	*S. amplexicaulis*	T	A	G	A	C
	Helix	*S. rizeensis*	C	A	G	A	C
	Helix	*S. purpurea* subsp. *leucodermis*	C	A	G	G	C
	Cheilophilae	*S. wilhelmsiana*	C	A	G	A	C

3.6.3. *rbc*L

The sequence data from *rbc*L gene region of Turkish *Salix* species supported two clusters in the phylogenetic tree (Figure 3.14); Subg. *Salix* and Subg. *Vetrix* same as revealed by the data from previous regions. The changes at 114, 264, 342, 603, 671, 690, 772, 952, 1120, 1156, 1243, 1261, 1457th bp positions formed the two clusters (Table 3.15). Among first clade, Subg. *Salix* obviously showed polytomy. The first main sub-clade involves *S. babylonica, S. alba x fragilis*, *S. fragilis*, *S. excelsa S. alba* with a 65 bootstrap value. Especially, *S. babylonica* got the top position of this sub-clade. Also, *S. amplexicaulis* which was accepted as the member Subg. *Vetrix* was grouped in this sub-clade. Externally joined ones were *S. pentandroides*, *S. triandra* subsp. *triandra*, *S. triandra* subsp. *bornmuelleri* (Figure 3.15). *S. acmophylla* was also joined to the sub-clade outside due to presence at 361st position. Species *S. amplexicaulis* showed high similarities with the other subg. *Salix* members based on specific substitution locations at 114 and 625th positions.

The species in second cluster of which Subg. *Vetrix* revealed dispersed allocation. However, there still has one main clade. *S. cinerea*, *S. pseudomedemii, S. pseudodepressa, S. caprea, S. pedicellata* subsp. *pedicellata*, *S. caucasica*, *S. apoda, S. armenorossica, S. elaeagnos* and *S. purpurea* subsp. *leucodermis* were lined more closely (Figure 3.16). *S. myrsinifolia-S. viminalis* pair was the sister clade to this polytomic clade with the high bootstrap value. Besides, *S. rizeensis, S. aegyptiaca* and *S. elbursensis* were the species which joined externally with the low bootstrap values under 50. *S. wilhelmsiana* showed clear separation from other members of subg. *Vetrix* on with specific substitutions at 57, 193, 360, 370 and 772th bp positions.

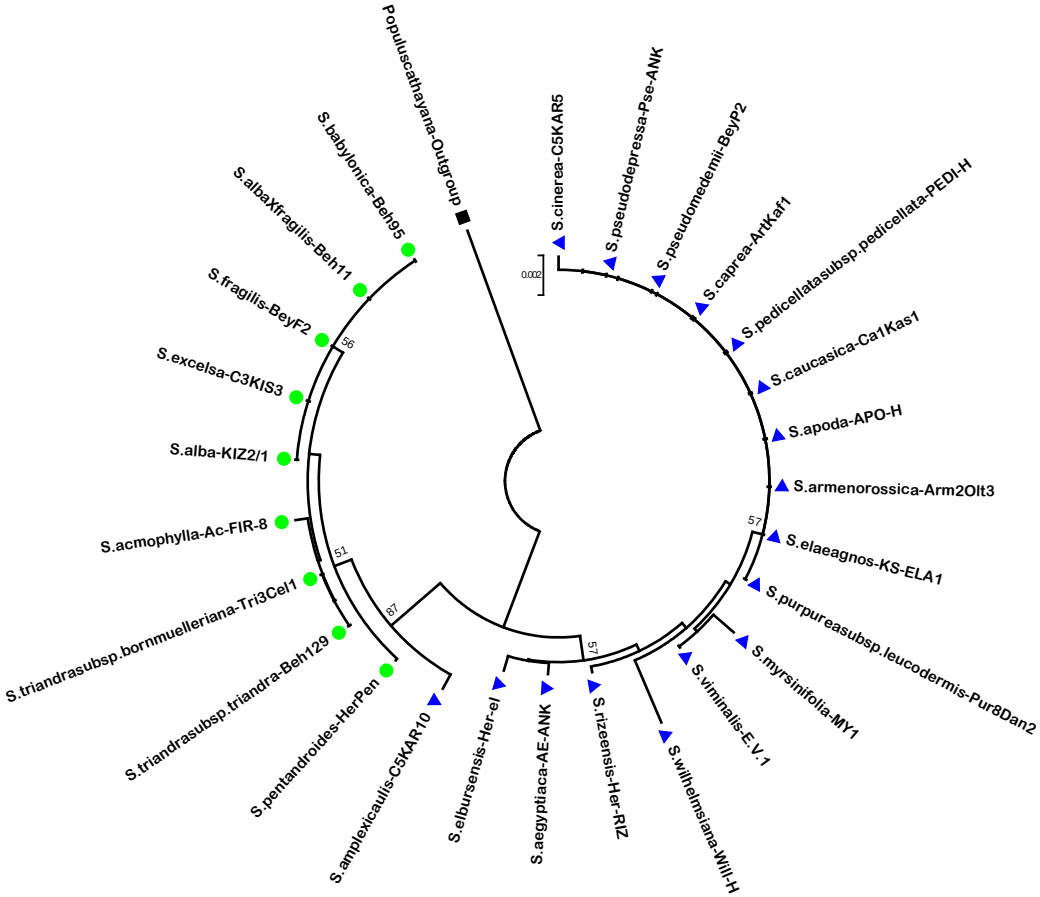

Figure 3.14. Phylogenetic tree, constructed using the sequences of *rbc*L region of cpDNA (Numbers on tree branches: bootstrap values, ●: Subg. *Salix*, ◀: Subg. *Vetrix*, ■; Outgroup)

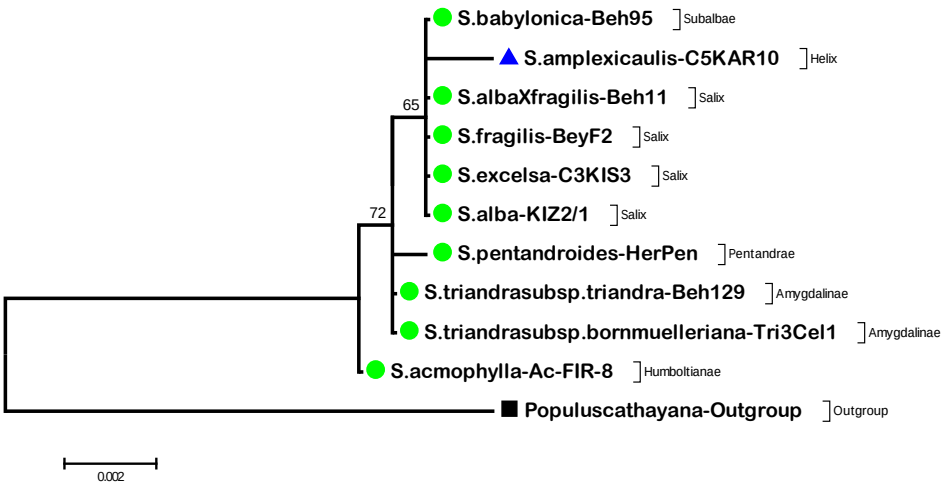

Figure 3.15. Phylogenetic tree, constructed using the sequences of *rbc*L region of cpDNA, showing detail relationships among the members of Subg. *Salix* and the section of species (Numbers on tree branches: bootstrap values, ●: Subg. *Salix*, ◀: Subg. *Vetrix*, ■; Outgroup, Parenthesis: Sections)

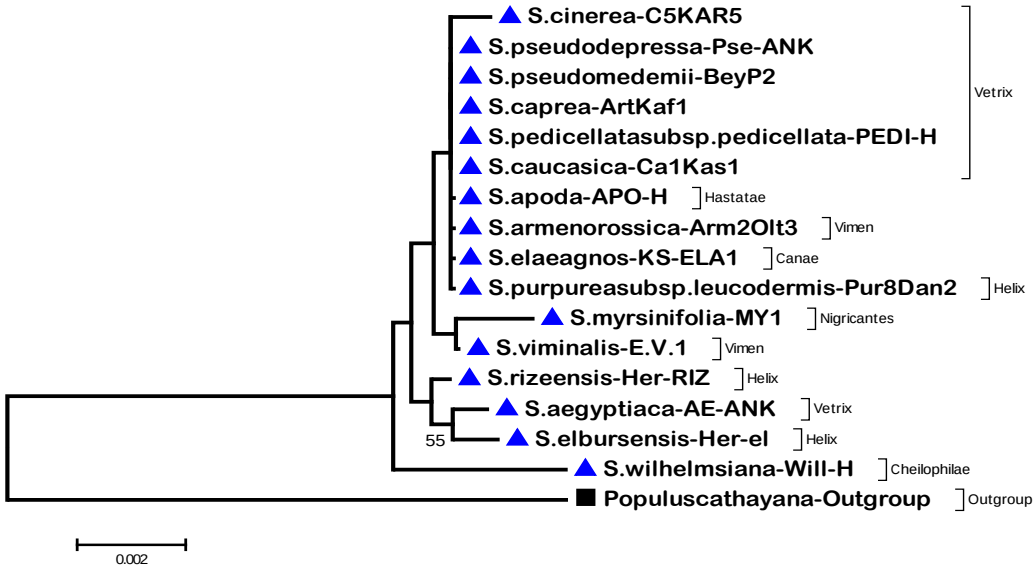

Figure 3.16. Phylogenetic tree, constructed using the sequences of *rbc*L region of cpDNA, showing detail relationships among the members of Subgenus *Vetrix* and the section of species (Numbers on tree branches: bootstrap values, ◀: Subg. *Vetrix,* ■; Outgroup, Parenthesis: Sections)

Table 3.15. The regions of indels and nucleotide substitutions in the DNA sequences of *rbc*L cpDNA region.

			Nucleotide Positions in the Aligned DNA sequences of rbcL Region																					
			***rbc*L 1**													***rbc*L 2**								
Subgenus	**Sections**	**Species name**	**4**	**57**	**114**	**193**	**217**	**264**	**342**	**360**	**361**	**370**	**603**	**625**	**671**	**690**	**772**	**866**	**952**	**1120**	**1156**	**1243**	**1261**	**1457**
	Humboltianae	*S. acmophylla*	A	C	G	G	G	G	G	T	C	G	C	T	G	C	A	G	G	C	G	T	G	G
	Amygdalinae	*S. triandra* subsp. *triandra*	A	C	G	G	G	G	G	T	A	G	C	T	G	C	A	G	G	C	G	T	G	G
	Amygdalinae	*S. triandra* subsp. *bornmuelleri*	A	C	G	G	G	G	G	T	A	G	C	T	G	C	A	G	G	C	G	T	G	G
SALIX	**Pentandrae**	*S. pentandroides*	A	C	G	G	G	G	G	T	A	G	C	T	G	C	A	G	G	C	G	C	G	G
	Salix	*S. alba*	A	C	G	G	G	G	G	T	A	G	C	T	G	C	A	T	G	C	G	T	G	G
	Salix	*S. excelsa*	A	C	G	G	G	G	G	T	A	G	C	T	G	C	A	T	G	C	G	T	G	G
	Salix	*S. fragilis*	A	C	G	G	G	G	G	T	A	G	C	T	G	C	A	T	G	C	G	T	G	G
	Salix	*S. alba x fragilis*	A	C	G	G	G	G	G	T	A	G	C	T	G	C	A	T	G	C	G	T	G	G
	Subalbae	*S. babylonica*	A	C	G	G	G	G	G	T	A	G	C	T	G	C	A	T	G	C	G	T	G	G
	Hastatae	*S. apoda*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Nigricantes	*S. myrsinifolia*	G	C	T	G	G	A	A	T	A	G	T	G	G	C	T	T	A	T	A	C	G	T
	Vetrix	*S. caucasica*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Vetrix	*S. pedicellata* subsp. *pedicellata*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Vetrix	*S. caprea*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Vetrix	*S. aegyptiaca*	A	C	T	G	G	A	G	T	A	G	T	T	T	T	T	G	A	T	A	C	G	T
	Vetrix	*S. cinerea*	A	C	T	G	A	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
VETRIX	**Vetrix**	*S. pseudomedemii*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Vetrix	*S. pseudodepressa*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Vimen	*S. viminalis*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	T	A	T	A	C	A	T
	Vimen	*S. armenorossica*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Canae	*S. elaeagnos*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Helix	*S. elbursensis*	A	C	G	G	G	G	G	T	A	G	T	T	T	C	T	T	G	C	G	T	G	G
	Helix	*S. amplexicaulis*	A	C	G	G	G	A	G	T	A	G	T	T	G	T	T	G	A	T	A	C	A	T
	Helix	*S. rizeensis*	A	C	T	G	G	A	G	T	A	G	T	G	G	C	T	G	A	T	A	C	G	G
	Helix	*S. purpurea* subsp. *leucodermis*	A	C	T	G	G	A	A	T	A	G	T	G	G	T	T	G	A	T	A	C	A	T
	Cheilophilae	*S. wilhelmsiana*	A	T	T	C	G	A	G	C	A	A	T	G	T	T	T	T	A	T	A	C	A	T

3.6.4. *ITS* (Internal Transcribed Spacer)

Constructed phylogenetic tree based on *ITS* region of nuclear DNA, depicted similar results with the trees constructed by using non-coding cpDNA sequences of *trn* T-F region. Members of Subg. *Vetrix* replaced in Subg. *Salix* clade. One main cluster was formed with subg. *Salix* and subg. *Vetrix* members which were externally attached to it and formed separate one (Figure 3.17). The responsible substitution for this separation were observed at 57, 99, 125, 128, 143, 181, 188, 246, 391, 421, 501 and 564 th base positions and one deletion at 576th base position (Table 3.16). Although *ITS* is one of the shortest aligned region among the studied regions, it had many informative sites when subgenera clustering was considered. The most relative species were *S. alba*, *S. excelsa, S. fragilis, S. alba x fragilis, S. triandra* subsp. *triandra, S. triandra* subsp. *bornmuelleri* and *S. babylonica* with high bootstrap value (65) nested together in Subg. *Salix* clade (Figure 3.18). S. *babylonica* was located at the bottom of Subg. *Salix* the cluster. *S. acmophylla* was the most distant species as a result of specific substitution at 54th base (Table 3.16).

In subg. *Vetrix* cluster, there was not any real sub-clade formation. Complex relations among species were observed (Figure 3.19). However, some species were phylogenetically closer to each other. For example, *S. caprea-S. cinerea* were close because of the presence of substitution at 605th bp position. Also, *S. aegyptiaca-S. pseudomedemii*, *S. apoda-S. caucasica-S. pedicellata* subsp. *pedicellata* were attached to the main cluster as sister clades. Thus, most of the section *Vetrix* members were nested together in subg. *Vetrix* clade. Besides, *S. pseudodepressa*, *S. elbursensis* and *S. amplexicaulis* were the most distantly related species. *S. amplexicaulis* was located very close to members of Subg. *Salix* cluster as a result of substitutions at 57, 99, 145, 421, 501 and 564th base positions.

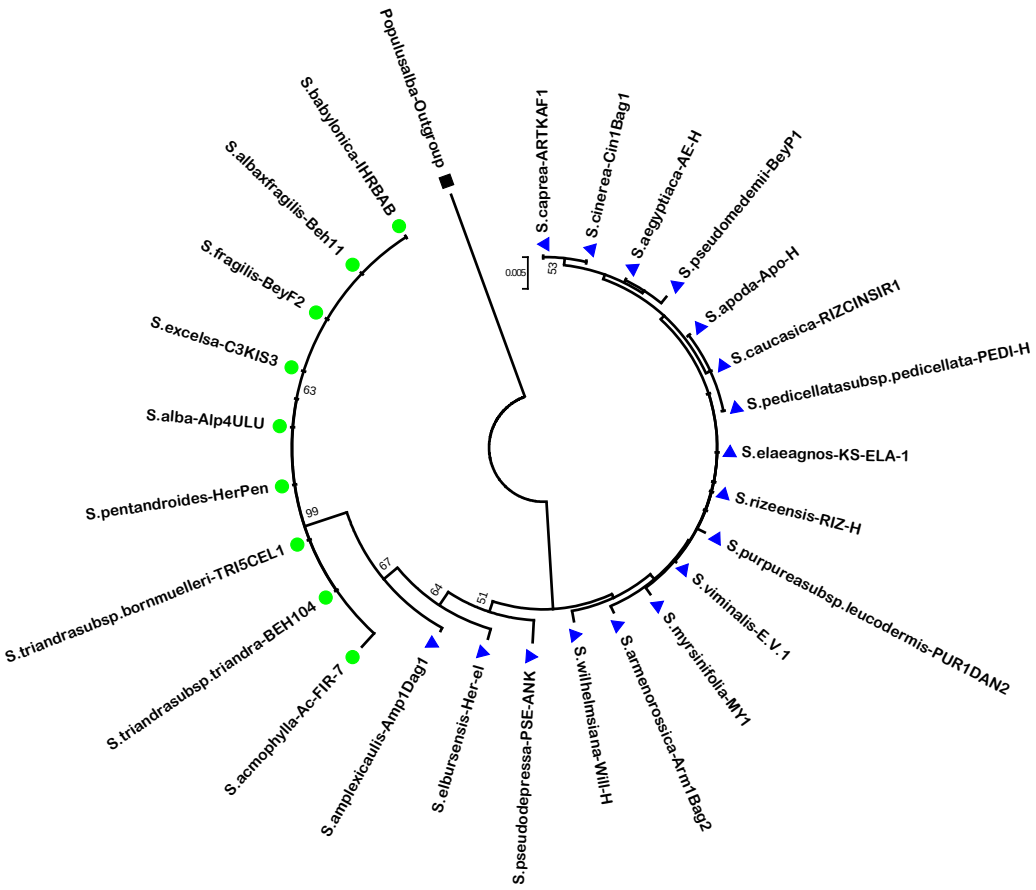

Figure 3.17. Phylogenetic tree, constructed by using the sequences from *ITS* region of nrDNA (Numbers on tree branches: bootstrap values, ●: Subg. *Salix*, ◀: Subg. *Vetrix*, ■; Outgroup).

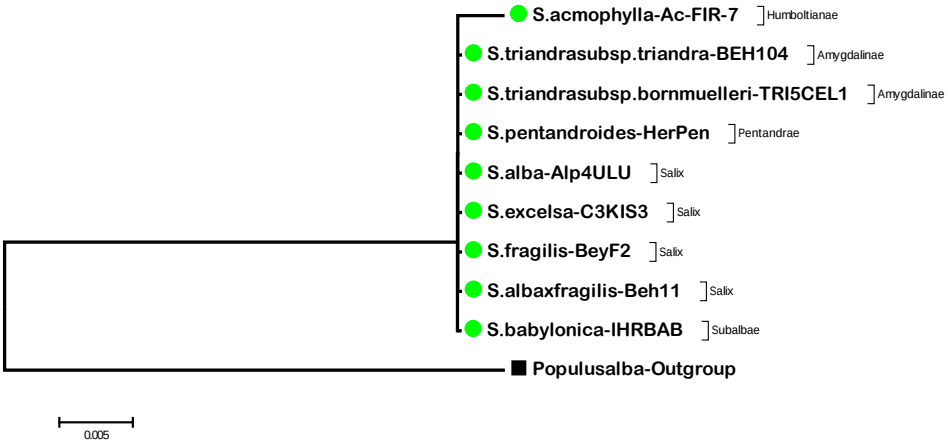

Figure 3.18. Phylogenetic tree, constructed by using the sequences from *ITS* region of nrDNA, showing detail relationships among the members of Subg. *Salix* and the section of species. (Numbers on tree branches: bootstrap values, ●: Subg. *Salix*, ■; Outgroup, Parenthesis: section).

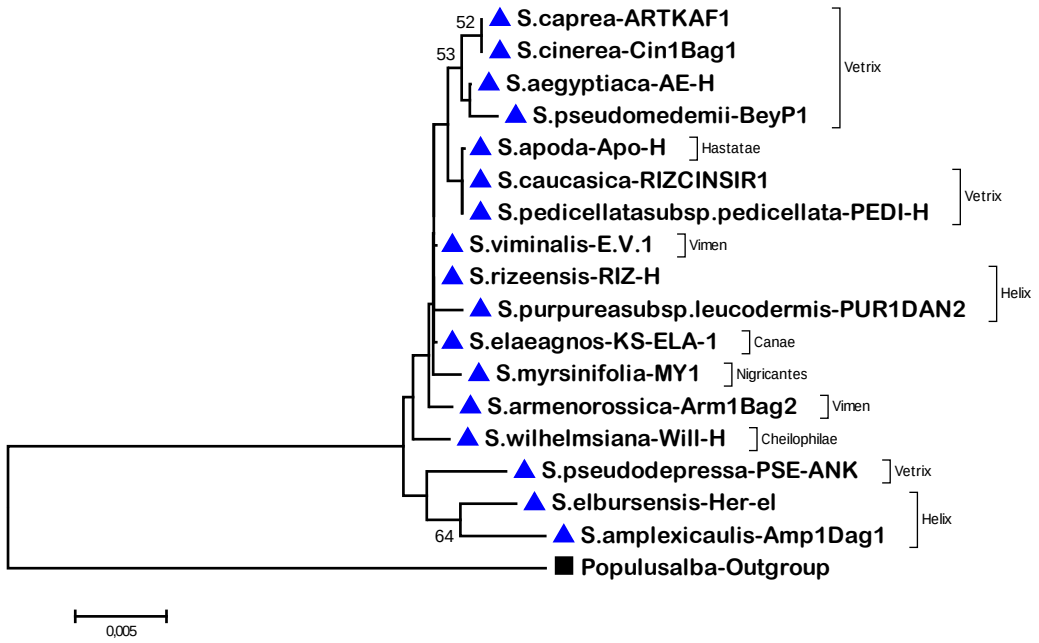

Figure 3.19. Phylogenetic tree, constructed using the sequences of *ITS* region of nrDNA, showing detail relationships among the members of Subg. *Vetrix* and the section of species (Numbers on tree branches: bootstrap values, ◀: Subg. *Vetrix,* ■; Outgroup, Parenthesis: Sections)

Table 3.16. The regions of indels and nucleotide substitutions in the DNA sequences of *ITS* nrDNA region

			Nucleotide Positions in the Aligned DNA sequences of ITS nrDNA Region																	
			ITS (ITS1- 5.8 -ITS2 region)																	
Subgenus	**Sections**	**Species name**	**54**	**57**	**99**	**125**	**128**	**143**	**181**	**188**	**246**	**391**	**421**	**501**	**549**	**564**	**576**	**605**	**610**	**652**
	Humboltianae	*S. acmophylla*	G	T	C	G	C	C	G	C	T	T	G	C	T	C	-	C	A	G
	Amygdalinae	*S. triandra* subsp. *triandra*	A	T	C	G	C	C	G	C	T	T	G	C	G	C	-	C	A	G
	Amygdalinae	*S. triandra* subsp. *bornmuelleri*	A	T	C	G	C	C	G	C	T	T	G	C	G	C	-	C	A	G
	Pentandrae	*S. pentandroides*	A	T	C	G	C	C	G	C	T	T	G	C	G	C	-	C	A	G
	Salix	*S. alba*	A	T	C	G	C	C	G	C	T	T	G	C	G	C	-	C	A	G
	Salix	*S. excelsa*	A	T	C	G	C	C	G	C	T	T	G	C	G	C	-	C	A	G
SALIX	**Salix**	*S. fragilis*	A	T	C	G	C	C	G	C	T	T	G	C	G	C	-	C	A	G
	Salix	*S. albaxfragilis*	A	T	C	G	C	C	G	C	T	T	G	C	G	C	-	C	A	G
	Subalbae	*S. babylonica*	A	T	C	G	C	C	G	C	T	T	G	C	G	C	-	C	A	G
	Hastatae	*S. apoda*	A	C	T	A	T	T	T	T	G	C	A	T	G	T	G	C	A	G
	Nigricantes	*S. myrsinifolia*	A	C	T	A	T	T	T	C	G	T	A	T	G	T	G	C	A	G
	Vetrix	*S. caucasica*	A	C	T	A	T	T	T	T	G	C	A	T	G	T	G	C	A	G
	Vetrix	*S. pedicellata* subsp. *pedicellata*	A	C	T	A	T	T	T	T	G	C	A	T	G	T	G	C	A	G
	Vetrix	*S. caprea*	A	C	T	A	T	T	T	T	G	C	A	T	G	T	G	A	A	G
	Vetrix	*S. aegyptiaca*	A	C	T	A	T	T	T	T	G	T	A	T	G	T	G	A	A	G
	Vetrix	*S. cinerea*	A	C	T	A	T	T	T	T	G	C	A	T	G	T	G	A	A	G
	Vetrix	*S. pseudomedemii*	A	C	T	A	T	T	T	T	T	T	A	T	G	T	G	A	A	G
VETRIX	**Vetrix**	*S. pseudodepressa*	A	C	C	G	T	T	G	T	G	T	G	T	G	T	G	C	A	C
	Vimen	*S. viminalis*	A	C	T	A	T	T	T	T	G	T	A	T	G	T	G	C	A	G
	Vimen	*S. armenorossica*	A	C	C	A	T	T	T	T	G	T	A	T	G	T	G	C	A	G
	Canae	*S. elaeagnos*	A	C	T	A	T	T	T	T	G	T	A	T	G	T	G	C	A	G
	Helix	*S. elbursensis*	A	T	T	A	T	C	G	C	G	T	G	T	G	T	G	C	A	G
	Helix	*S. amplexicaulis*	A	T	C	A	T	C	T	T	G	T	G	C	G	C	G	C	A	G
	Helix	*S. rizeensis*	A	C	T	A	T	T	T	T	G	T	A	T	G	T	G	C	A	G
	Helix	*S. purpurea* subsp. *leucodermis*	A	C	T	A	T	T	T	T	G	T	A	T	G	T	G	C	G	G
	Cheilophilae	*S. wilhelmsiana*	G	C	T	G	T	T	T	T	G	T	A	T	G	T	G	C	A	G

3.7. Comparison of Turkish *Salix* Species with those are Native to Old and New World Species

To comprehend the evolutionary relationship and divergence between Turkish and world *Salix* species, the sequences from four studied molecular regions were obtained for the world *Salix* species from NCBI data base. There was no available DNA sequence for *trn*T-L and *trn*L-F IGS regions of any *Salix* species in the NCBI database. Also, there was no aligned *trn*L sequence from New World *Salix* species. Thus, the genetic divergence and molecular diversity parameters were evaluated by using data from only *trn*L region for Turkey vs Old World *Salix* sp. (Table 3.17).

For molecular diversity parameter estimations, the total length of *trn*L region was decreased to 510bp to make reliable comparisons. There were 34 variable sites observed. Most of them were parsimony informative (Table 3.18). It is clear to say that rather than *trn*T-L IGS and *trn*L-F IGS regions, *trn*L region is not conserved very much when Turkish species and OWS were compared. Also, there were 14 indel sites in OWS. Even though it is a small region, *trn*L region had the highest nucleotide diversity (0.018) among other studied gene regions. Transversional pairs (61,23) were almost doubled the transitional pairs because of Turkish sequence data distribution (38,77). The moderate GC content (31.0%) can be explained by the fact that *trn*L is a non-coding region. As far as genetic divergence is concerned, overall divergence was quite high (0.027). The genetic distance is reached to high values in OWS. Because of this, Old World *Salix* sp. made distinct clade in phylogenetic tree based on *trn*L region. Turkish *Salix* species were grouped as previous which made two clusters (Figure 3.20). But interestingly, some OWS members which belong to subg. *Salix* were genetically close to members of Turkish subg. *Salix*. The other OWS species from subg. *Vetrix* were grouped with Turkish *Vetrix* based on *trn*L region. Thus, OWS species were scattered to Turkish *Salix* subgenera regardless of geographic distribution.

To make good comparisons with world *Salix* sp., *mat*K sequence data was trimmed to 780 base pairs. After alignment of *mat*K gene region of Turkish and World *Salix* species, it was found that there were 32 variable sites. Only 7 of them were

parsimony informative sites. High variations were the consequence of variations in both Turkish *Salix* and NWS. In total evaluation, the transversional pairs value was high and the R ratio was the same as it was in *trn*L region. Over all genetic distance was determined as the lowest value among other regions (0.005). Divergence between New World and Turkish *Salix* species was not high compared to Old World. Different from *trn* L region phylogenetic tree, World *Salix* species were not clustered together based on *mat*K gene region. They were dispersed to relate the places according to which subgenus the species belongs to. Thus, selected Old and New world *Salix* species were nested among Turkish subgenera rather than clustering together (Figure 3.21). The NWS species were especially positioned close to subg. *Vetrix.*

The *rbc*L gene region length was about 1324 bp long when the additional *Salix* species from the world were included. This region had the lowest variable sites, (28 bp) and almost half of them were parsimony informative. GC content of *rbc*L region is quite high (44.2). Even though it was the longest region to others, it had low nucleotide diversity. Another lowest value was calculated for overall divergence (0.006). The genetic divergence between subg. *Vetrix* and others for *rbc*L gene region was high (Figure 3.22). Subg. *Vetrix* members were separated from NWS species (0.011) and from OWS species (0.003) with high divergence values. In *rbc*L based phylogenetic tree, OWS and NWS species were scattered throughout the Turkish *Salix* species according to taxonomically related subgenera which they belong to. Selected NWS species were close to subg. *Salix* whereas selected OWS species were nested in subg. *Vetrix* (Figure 3.22).

Finally, nrDNA *ITS* is the most studied region in the world so high number of sequences were added to analysis from the world. This region had the richest GC content in the molecular diversity parameters (64.7). Besides, the highest variable sites were observed in this region. Among 38 variable sites, 21 sites were parsimony informative sites. Transitional pairs analysis was interestingly resulted in a value as high as 79.64. As a consequence of this, the R increased to 3.57. Genetically, Subg. *Salix* was diverged from all other groups; Subg. *Vetrix*, Old World species and New World species with high genetic distance value. Overall divergence value was quite high (0.010). Old and New World *Salix* species stayed very close to each other with a

value of 0.006. From the phylogenetic tree constructed by nrDNA *ITS* gene region, it was observed that Old World and New World *Salix* species were scattered among to Turkish *Salix* species, especially they were positioned among the members of Subg. *Vetrix* (Figure 3.23). Turkish Subg. *Salix* made a distinct clade like other cpDNA phylogenetic trees.

Table 3.17. Genetic divergence among *Salix* species throughout the World. Distance values were obtained by a bootstrap procedure (1000 replicates) and appropriate models for every region. (OWS: Old World *Salix* and NWS: New World *Salix*). The high values are marked in bold.

Subgenera/ OWS/NWS	Subg. *Salix*	Subg. *Vetrix*	NWS	OWS	Over All	Studied cpDNA and nrDNA Regions
Subg. *Salix*	0.001					
Subg. *Vetrix*	0.043	0.008			0.027	***trn*L**
OWS	0.018	0.045	-	**0.020**		
Subg. *Salix*	0.001					
Subg. *Vetrix*	0.007	0.004			0.005	***mat*K**
NWS	0.005	0.006	0.007			
OWS	0.001	0.007	0.006	0.001		
Subg. *Salix*	0.001					
Subg. *Vetrix*	0.010	0.002			0.006	***rbc*L**
NWS	0.003	0.011	0.005			
OWS	0.010	0.003	0.011	0.005		
Subg. *Salix*	0.001					
Subg. *Vetrix*	0.017	0.005			0.010	***ITS***
NWS	0.018	0.003	0.001			
OWS	0.017	0.008	0.006	**0.010**		

*There is no data to evaluate divergence in NWS.

Table 3.18. The estimated molecular diversity parameters based on three cpDNA and one nrDNA gene regions in *Salix* genus for Turkey

	*trn*L			*mat*K				*rbc*L				*ITS*			
	Turkey	*OWS*	*Total*	*Turkey*	*OWS*	*NWS*	*Total*	*Turkey*	*OWS*	*NWS*	*Total*	*Turkey*	*OWS*	*NWS*	*Total*
Number of species	24+1*	5	**26**	23+1*	5	9	**36**	24+1*	6	3	**31**	24+1*	18	5	**42**
Number of sequence	26	5	**31**	25	2	9	**39**	26	6	3	**35**	26	18	5	**49**
Total length (bp)**	510	510	**510**	780	780	780	**780**	1324	1324	1324	**1324**	598	598	598	598
GC content (%)	30.9	31.5	**31.0**	33.4	33.6	33.5	**33.4**	44.2	44.2	44.4	**44.2**	64.8	64.7	64.3	**64.7**
Conserved sites	489	496	**483**	760	777	763	**747**	1304	1307	1314	**1296**	581	567	595	**559**
Variable sites	21	12	**34**	19	2	16	**32**	19	16	9	**28**	16	30	2	**38**
Parsimony informative sites	18	11	**24**	7	-	5	**7**	12	-	-	**16**	14	14	-	**21**
Transitional pairs	25.89	44.23	**38.77**	42.84	0.00	40.24	**39.1**	68.36	75.36	67.06	**67.08**	77.74	84.12	52.16	**79.64**
Transversional pairs	74.11	55.77	**61.23**	57.16	100.0	59.76	**60.9**	31.64	24.64	32.94	**32.92**	22.26	15.88	47.84	**20.36**
Transition/Transversion (tr/tv) (R) ratio	0.31	0.72	**0.57**	0.67	0.00	0.60	**0.57**	2.13	3.02	2.01	**2.01**	3.19	4.84	1.00	**3.57**
Number of Indels	-	14	**14**					0	1	1	**1**	1	7	0	**7**
Nucleotide diversity	0.016	0.014	**0.018**	0.005	0.001	0.007	**0.005**	0.005	0.004	0.005	**0.005**	0.011	0.010	0.001	**0.010**

OWS (Old World *Salix*), NWS (New World *Salix*).*One hybrid was added to analysis. **The total length was estimated after all sequence are aligned and trimmed.

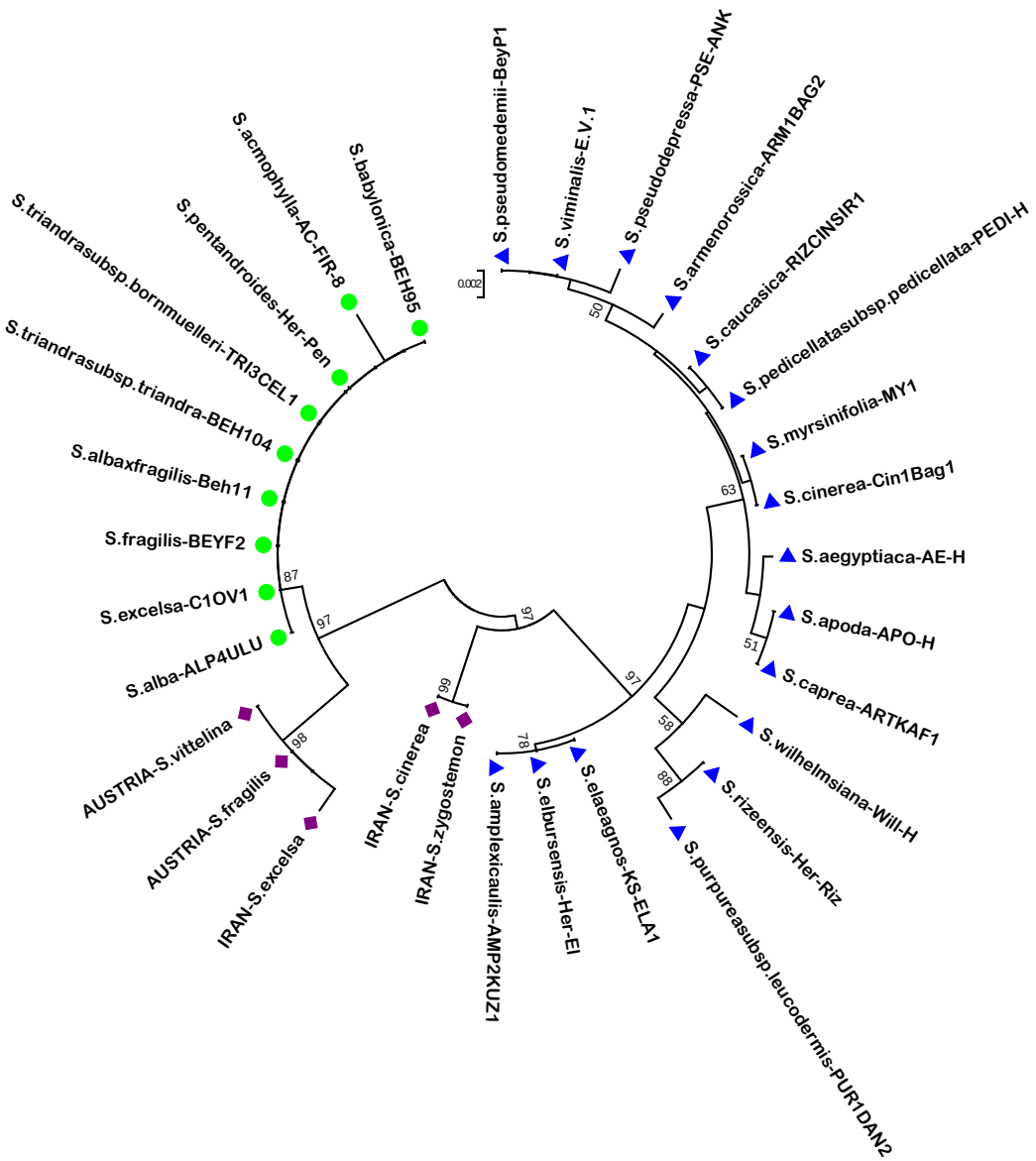

Figure 3.20. Phylogenetic tree (*trn*L) showing relations of Turkish *Salix* species with ones from the world (◆; OWS; Old World *Salix*, ▲; Subg. *Vetrix*, ●;Subg. *Salix*)

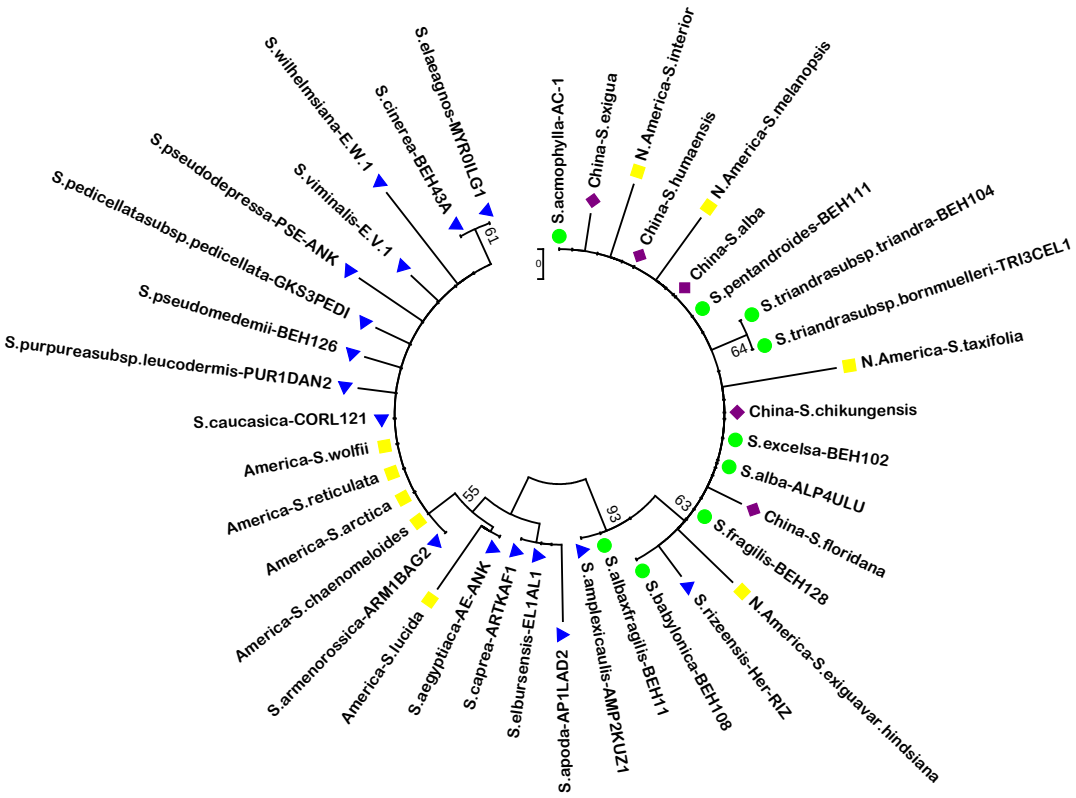

Figure 3.21. Phylogenetic tree (*mat*K) showing relations of Turkish *Salix* species with ones from the world (◆ ; OWS;Old World *Salix*, ■ ; NWS; New World *Salix*, ▲; Subg. *Vetrix*, ●;Subg. *Salix*)

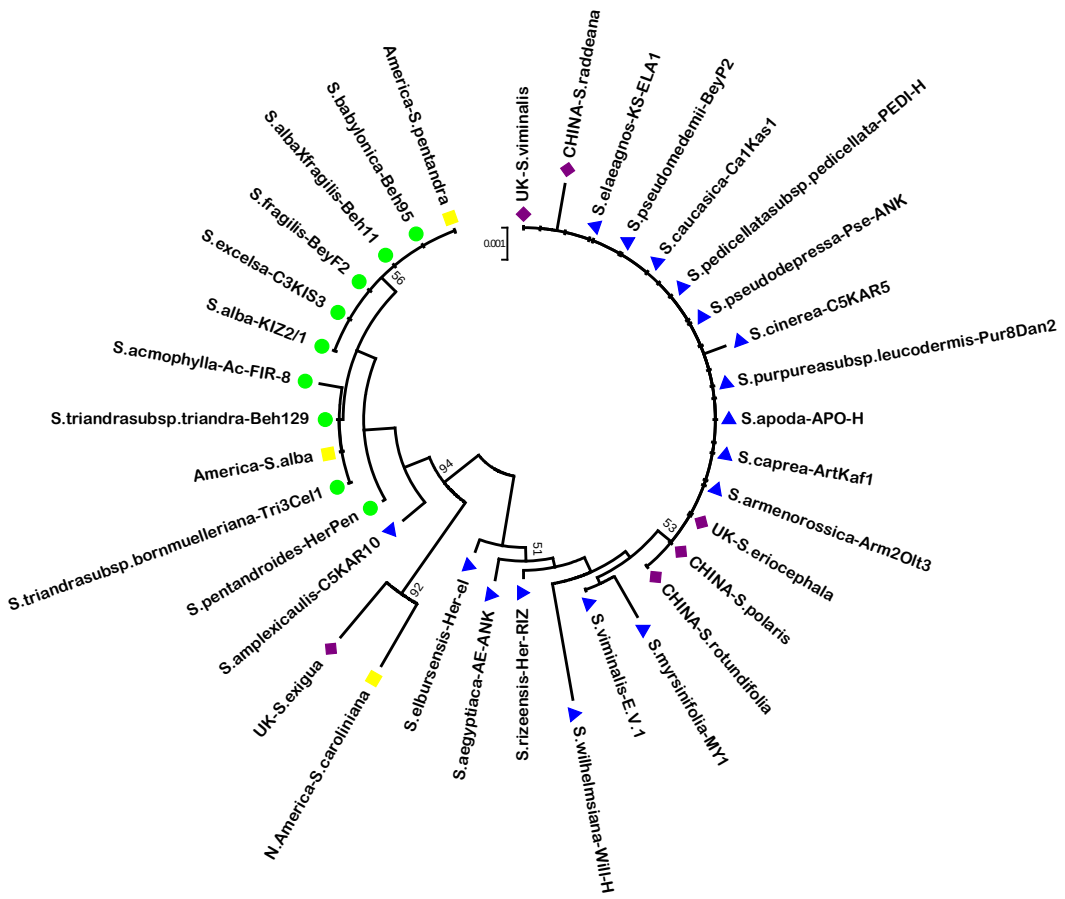

Figure 3.22. Phylogenetic tree (*rbc*L) showing relations of Turkish *Salix* species with ones from the world (◆; OWS; Old World *Salix*, ■ ; NWS; New World *Salix*, ▲; Subg. *Vetrix*, ●; Subg. *Salix*)

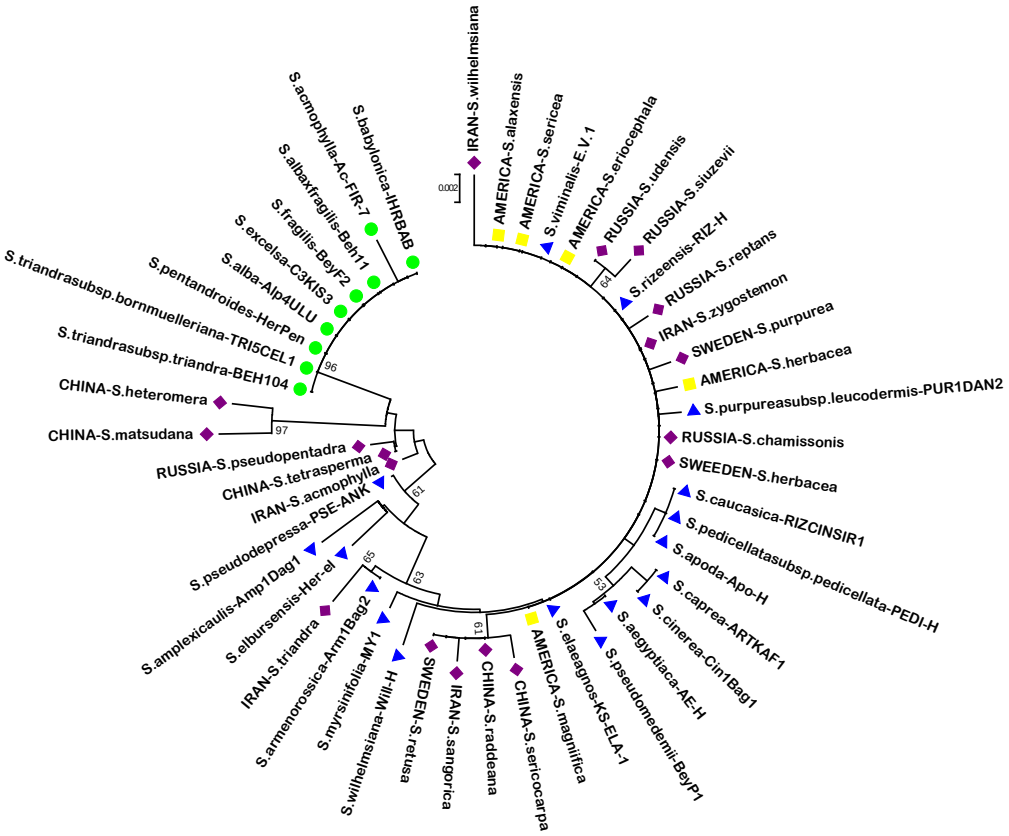

Figure 3.23. Phylogenetic tree (*ITS*) showing relations of Turkish *Salix* species with ones from the world (◆; OWS; Old World *Salix*, ■; NWS; New World *Salix*, ▲; Subg. *Vetrix*, ●; Subg. *Salix*)

3.8. Molecular Clock Estimation

Sequences from NCBI database were used to measure the evolutionary divergence time of *Salix* L. based on *rbc*L and *mat*K cpDNA regions. The *trn*L cpDNA region was used to estimate molecular clock for only OWS and subgenera of Turkish *Salix*. The number of parsimony informative sites, total length of the region, d and k values, and molecular clock times are tabulated in detail in Table 3.19. The synonymous mutation rate per base pair per year was assumed to be 2.5×10^{-9} which was used in *Populus* sp. by Levsen *et al*. (2012). Also, the aligned sequences for every region were used to calculate molecular clock as total length.

The oldest molecular divergence time was observed in *trn*L region between Turkish and OWS backed to 19.4 million years ago (Table 3.19). When the Turkish subgenera are considered, the number was decreased to 14.5 Mya. High variation in Iranian *Salix* gave rise to divergence from Turkish *Salix* species. Since the *trn*L

region was the non-coding part of cpDNA, the divergence times dated back much earlier times. The estimated age for diversification of two main clades of Turkish *Salix* was recently when compared with OWS from Turkey.

General divergence times of both *rbc*L and *mat*K gene seem to be relatively recent when compared with *trn*L gene region. The molecular clock estimation evaluated as total Turkish *Salix* vs NWS or OWS. To illustrate, the divergence times of OWS and NWS from Turkish *Salix* was calculated as 3.6 Mya (Table 3.19). Similarly, the subgenera of Turkish *Salix* divergence was 3.6 Mya when DNA sequences of *rbc*L and *mat*K regions were used. Very recent divergence of OWS-subg. *Salix* (*trn*L and *mat*K) and NWS-Subg. *Vetrix* (*mat*K) were the evidence of close relations for those pairs. The calculation of molecular clock of NWS vs OWS based on *mat*K gene region was contrary with the dates based on *rbc*L gene region. The most recent divergence time was found between NWS and OWS based on *mat*K region (2.5 Mya). On contrast, *rbc*L data estimated the divergence time of OWS from NWS dated back to much earlier (4.5 million years ago). For the convenient time scale, 4.5 Mya (*rbc*L) was used in the current evaluation.

Table 3.19. Molecular Clock Estimations for NWS-OWS and Turkish *Salix* species based on cpDNA region.

Salix Groups	Regions	# of Parsimony Informative Sites	Length of the regions (bp)	d	k	MCE (Mya)*
NWS-OWS	***rbc*L**	15	1324	0.011	0.011	**4.5**
	***mat*K**	5	780	0.006	0.006	**2.5**
Turkey (Subg. *Salix*-Subg. *Vetrix*)	***trn*L**	18	510	0.035	0.036	**14.5**
	***rbc*L**	12	1324	0.009	0.009	**3.6**
	***mat*K**	7	780	0.009	0.009	**3.6**
NWS-Turkey	***rbc*L**	12	1324	0.009	0.009	**3.6**
	***mat*K**	7	780	0.009	0.009	**3.6**
NWS-Subg. *Salix*	***rbc*L**	1	1324	0.001	0.001	**0.3**
	***mat*K**	6	780	0.008	0.008	**3**
NWS-Subg. *Vetrix*	***rbc*L**	12	1324	0.009	0.009	**3.6**
	***mat*K**	0	780	0	0	**-**
OWS-Turkey	***trn* L**	24	510	0.047	0.049	**19.4**
	***rbc*L**	12	1324	0.009	0.009	**3.6**
	***mat*K**	7	780	0.009	0.009	**3.6**
OWS*-Subg. *Salix	***trn* L**	13	510	0.025	0.026	**10.3**
	***rbc*L**	12	1324	0.009	0.009	**3.6**
	***mat*K**	1	780	0.001	0.001	**0.5**
OWS-Subg. *Vetrix*	***trn* L**	24	510	0.047	0.049	**19.4**
	***rbc*L**	5	1324	0.005	0.005	**1.5**
	***mat*K**	6	780	0.008	0.008	**3**

* Molecular Clock Estimation (Million Years Ago)

CHAPTER 4

DISCUSSION

4.1. Numerical Taxonomical Method

Vegetative characters were used to find similarities between *Salix* sp. Since they have reduced flowers which is different than other angiosperms (Azuma *et al*., 2000). Turkish *Salix* species are divided into two subgenera with respect to eleven selected morphological characters, with some exceptions, in line with Skvortsov (1968). However, the limited representative number of samples for *S. caprea* and *S. cinerea* and for *S. myrsinifolia* and *S. wilhelmsiana* (as herbarium materials) have deviated from other members of subg. *Vetrix*. Particularly, the deviation of *S. wilhelmsiana* was clearly observed from subg. *Vetrix* cluster in morphological PCA analysis. It can be the result of deformation of herbarium materials. According to morphological data, there is not any support for positioning of *S. rizeensis* and *S. amplexicaulis* in subg. *Salix.* Only, the taxa *S. amplexicaulis* located at the edge of Subg. *Vetrix* clusters may support the molecular data. So this result is harmony with the molecular data which provided replacement of taxa position of *S. amplexicaulis*. The hybrid species *S. alba x fragilis,* determined according to intermediate morphological characters, exhibits intersected variable characters as stated by Percy *et al.* (2014) for hybrids.

The most discriminative morphological features based on taxonomic classification were petiole length (Pl) and bud angle (Ba) in the current study. Since only two characters did not reveal the relation of Turkish *Salix* sp. which was also noticed with previous study (Chen *et al*., 2010). Nine more traits (life form, bud scale, brunch

habit, bark type, stipule persistence, decorticated wood, leaf shape, leaf color, twig slender) were added to analysis. The most reliable binary data were bud scale (glabrous/not), life form (shrub/not) and leaf shape (lanceolate/not). The glabrous bud scale has the ability to be used as morphology based taxonomic classification of *Salix* genus which was also prove by Wu *et al.* (2015). Pubescence bud scale nature of Turkish subg. *Vetrix* can be explained by the adaptation to high and cooler habitat conditions (Wu *et al*., 2010; Lauren-Moreau, *et al*., 2015). Since most of subg. *Vetrix* members are naturally found in Black Sea region of northern Anatolia, bud scale (pubescence) can be an advantage in high altitude and wind-cooling habitat (northern Anatolia) rather than warmer climatic condition (inner Anatolia) to conserve water and reflect solar radiation (Ehleringer and Björkman, 1978). Most of the subg. *Salix* individuals were revealed as tree-like life form and lanceolate leaf shape. Appearance of distinct lanceolate leaf form in Subg. *Salix* is in parallel with taxonomist's morphological classification (Davis, 1965; Skvortsov, 1999). The findings are congruence with the Skvortsov(1999)'s statements that subg. *Salix* is a natural and ancient group, similar to *Populus* and exhibiting "primitive" features.

The remaining binary data results in our PCA data show variability from sample to sample rather than making good correlation. Thus, Azuma *et al*. (2000) and this study reached at similar conclusion that morphology based systematic treatment of *Salix* varies widely, and needs careful consideration.

4.2. Molecular Diversity in *Salix* sp.

The three cpDNA and one nrDNA used in our study are the regions selected for the barcoding of Turkish *Salix* sp. Samples collected from many locations in Turkey were aligned 1339bp, in *trn*T-F, 1731bp in *mat*K, 1480bp in *rbc*L and 657bp in *ITS*.

The obtained variable sites of *trn*T-F region in Turkish *Salix* sp. is coherent with the previous statement that non-coding regions of chloroplast genome have a potential of highest frequency of mutations (Taberlet *et al.,* 1991). This is also seen in European willows in which highlighted polymorphism was detected for non-coding region of cpDNA (Hamza-Babiker *et al.,* 2009). The most variable part of this region was found in *trn*L intron which is also supported by results of Bakker *et al*. (2000). On

the other hand, two flanking intergenic spacers (IGS); *trn*T-L and *trn*L-F are highly conserved. In the present study, the highest nucleotide diversity was observed in non-coding cpDNA *trn*T-F region (0.017) that expresses the degree of polymorphism in a population at the nucleotide level (Nei and Li, 1979).

The length of *mat*K region varies from 797 to 1565 in previous studies (Hardig *et al*., 1998; Barkalov and Kozyrenko; 2014a*,* Percy *et al*., 2014; Lauren Moreau *et al*., 2015; Wu *et al*., 2015). The least number of variable sites for *mat*K region was found in Russian willows as 12 variable sites and only 6 of them were informative (Table 4.1). The variable sites of Subg. *Salix* in the world species were about 54 (Wu *et al.,* 2015) whereas in Turkey Subg. *Salix* involves 13 variable sites. This may be due to low number of Turkish Subg. *Salix* in the study. As far as Turkish *Salix* species are concerned, the highest parsimony informative sites (33) in variable sites (45) for *mat*K gene region was detected in Turkish species than those in world *Salix* studies. Particularly, the results are similar to variable sites of North American *Salix* species with highest values (50) (Percy *et al*., 2014). This situation can be clarified by the long aligned sequence of *mat*K gene region including some part of *trn*K intron. The variation was higher at the sequence of 5' region (*mat*K1) than sequence adjacent to at the 3' (*mat*K2) ends. Thus, the conserved 3' end with informative sites is beneficial in resolving phylogeny (Hilu and Liang, 1997). The moderate rate of variation in *mat*K region, supported that it evolves moderately (Hilu and Liang 1997; Kelchner, 2002).

The various scientists extensively used *rbc*L region to depict phylogenetic relationships among species of *Salix* (Azuma *et al*., 2000; Chen *et al*., 2010; Barkalov and Kozyrenko*,* 2014a*;* Percy *et al*., 2014; Lauren Moreau *et al*., 2015; Wu *et al*., 2015; Liu *et al*., 2016). The chloroplast encoded *rbc*L was found as much conserved in European and Russian *Salix* sp. (Hamza-Babiker *et al.,* 2009; Barkalov and Kozyrenko, 2014b). As opposed to that, Chen *et al.* (2010) found many meaningful (53) changes in all substitutions (117) for amplified cpDNA *rbc*L region (1484bp). In the current *rbc*L data set, the most of the varied sequence of *rbc*L gene of *Salix* sp. were informative (19). However, the *rbc*L gene alone is still too conserved and not resolved relations in lower taxonomic level which was also argued by Gielly and Taberlet (1994). The efficiency of this gene region is increased by

combination of an other cpDNA region such as *mat*K (The Consortium for Barcode of Life (CBOL) Plant working group, 2009).

Based on *ITS* region molecular diversity, data from the current study showed correlation with moderate level of variable sites in *ITS* gene region with previous studies. Particularly, moderate level of variation is fitted with European willows by Leskinen and Alström-Rapaport (1999) and Russian willows by Barkalov and Kozyrenko (2014b). On the contrary, the studies represented with high number of *Salix* species (Wu *et al*., 2010; Lauren *et al*., 2015) detected significant variable sites in nrDNA *ITS* region. The GC content of the *ITS* regions in *Salix* sp. is over 65%, where the results are similar to those given in other studies (Leskinen and Alström-Rapaport, 1999; Wu *et al*., 2015). High level of GC is the indicator of high genomic variation in DNA sequence and effectively used parts of DNA (Serres Giardi *et al*., 2012; Mishra *et al*., 2009). Moreover, the region had the one of the highest nucleotide diversity value as 0.013 which indicates high polymorphism. In this way, it can be said that *ITS* region within the nuclear genome had many informative sites with respect to the total studied regions of cpDNA (*trn*T-F*, mat*K*, rbc*L) which was also verified by the Wu *et al*. (2015). Thus, the current cpDNA and nrDNA data are in parallel direction but not in totally coherent. Since the chloroplast sharing and hybrid speciation are known among *Salix* species (Brunsfeld *et al*., 1991). Thus, the evaluation of nuclear DNA is useful to get elucidative results.

In all regions, the variable sites are present in high number in Subg. *Vetrix* while the observed indel numbers were high in Subg. *Salix*. Only cpDNA *mat*K and nrDNA *ITS* region possess one indel and the other cpDNA *trn* T-F and *rbc*L had none in Turkish *Salix* sp., indicating the important function of selected regions in evolution and conservation (Hilu and Liang, 1997). The coding region of cpDNA (*mat*K and *rbc*L) was found to be as more conserved and having slow evolution rate than non-coding cpDNA(*trn*T-F) and nrDNA *ITS.* Besides that, nucleotide diversity was found higher in subg. *Vetrix* than subg. *Salix* for all studied regions.

In order to make comperative analysis of world sequences, both cpDNA and nrDNA gene regions were trimmed to 510bp in *trn*L, 780 bp in *mat*K, 1324 bp in *rbc*L and 559 bp in *ITS* with selected OWS and NWS species sequences from the world. The

highest variable site was+ observed in nrDNA *ITS* gene region (38) and most of them were informative. Despite longest cpDNA gene region aligned was *rbc*L, it has the lowest variable site (28) in comparison of world due to the large subunit of ribulose 1, 5 bisphosphate carboxylase/oxygenase. The R value in *rbc*L gene region is estimated as 3.57 due to high number of transitional pairs. *ITS* gene region has another highest score with the GC content value as 64.7 among other studied regions. The least parsimony informative site was found in *mat*K gene region (7). The other important molecular parameter, nucleotide diversity which express the degree of polymorphism in a population at the nucleotide level (Nei and Li, 1979) were found as highest in non-coding cpDNA *trn*L (0.018) and nrDNA *ITS* (0.010) gene regions.

Table 4.1. Comparison of the Molecular Diversity Parameters of *Salix* species in Previous Studies

Author	# species	Studied Area	Gene Region	# sequenced sites (bp)	# variable sites(bp)	# parsimony informative sites(bp)	# indels	GC content %
Leskinen and Alström-Rapaport (1999)	***13***	***Europe***	***ITS***	596	25	11	*	65-69
Azuma *et al.* (2000)	***23***	***Japan***	***rbc*L**	1398	*	*	*	*
Chen *et al.* (2010)	***46***	***Asia***	***rbc*L**	1484	117	53	*	*
Hardig *et al.* (2010)	***25***	***North America***	***mat*K**	1564	*	*	*	*
			ITS	665				
Abdollahzadeh *et al.* (2011)	***57***	***Iran***	***trn*L-F**	848	*	*	*	*
			ITS	608		49		
Barkalov and Kozyrenko (2014a)	***68***	***Far Eastern of Russia***	***mat*K**	835	12	6	*	*
			***rbc*L**	597	9	5		
			ITS	620	32	22		
Percy *et al.* (2014)	***73***	***North America***	***mat*K**	888	*	28	*	*
			***rbc*L**	628		10		
Lauron-Moreau *et al.* (2015)	***123***	***North America***	***mat*K**	874	50	*	*	*
			***rbc*L**	553	20			
			ITS	608	95			
Wu *et al.* (2015)	***128***	***All world (subg. Salix)***	***mat*K**	787	54	20	-	33.5
			***rbc*L**	1119	77	34	16	43.6
			ITS	682	157	57	209	64.5
Liu *et al.* (2016)	***18***	***China***	***mat*K *rbc*L combination**	2134	215	134	-	-
The current data set (2017)	***24***	***Turkey***	***trn*T-F**	1339	56	49	-	30.4
			***mat*K**	1731	45	33	1	32.4
			***rbc*L**	1480	22	19	-	43.4
			ITS	657	17	14	1	64.5

*Data is not represented in the related paper.

4.3. Genetic Divergence of *Salix* sp.

In the current study, the highest overall genetic divergence value recorded in the *trn*T-F (0.017) and *ITS* regions (0.013) in which the species were the most diverged. The highest genetic divergence value within the subgenus was in non-coding cpDNA region *trn*T-F (Subg.*Vetrix*-0.031) while the lowest in coding *rbc*L (subg. *Salix*-0.001) region. When the distance between the subgenera is considered, the largest value calculated from nrDNA *ITS* (0.023). Thus, nrDNA ITS is the most discriminative gene region in Turkish *Salix* subgenera separation. According to distance between each taxon, the closest relationship has been identified in subg. *Salix* than subg. *Vetrix* for all regions.

According to Azuma *et al*. (2000), *rbc*L gene sequence divergences for *Salix* were very low (0.007-0.005) among the studied species. This is totally concordant with our analyses. This divergence is widely decreased for North American willow within sect. *Longifolia* species based on nrDNA *ITS* dataset as 0.0012 (Brunsfeld *et al.,* 1991). Recently Barkalov and Kozyrenko (2014b) revealed the level of species divergence as 0.027 from *ITS* data in Far Eastern willow species which is a little more than the current study.

The estimated highest genetic distance among species was found as a result of presence of species; *S. acmophylla* in subg. *Salix* members and *S. amplexicaulis* in subg. *Vetrix* members. Natural distribution of the distant taxa, *S. acmophylla,* is generally located on mostly arid continental areas of Asia including South East (Skvortsov 1968). Anatolian diagonal separating South Eastern part of Turkey can be significant barrier to differ *S. acmophylla* from other subg. *Salix* member. Another explanation was suggested by Abdollahzadeh *et al*. (2011) for *S. acmophylla* in which he underlined the hybrid origin of it with several polymorphic sites in *ITS* gene region. Hardig *et al*. (1998) and Barkalov and Kozyrenko (2014b) confirmed our results that *S. acmophylla* is distant from other subg. *Salix* species with well supported clade in *ITS* phylogenetic tree. The subg. *Vetrix* had high divergence level among the species for all studied cpDNA and nrDNA regions as a result of high level of substitution. At species level, *S. wilhelmsiana (trn* T-F*), S. rizeensis (trn* T-F*), S. armenorossica (trn* T-F*, rbc*L*), S. caprea* (*rbc*L*, ITS)* and *S. cinerea* (*rbc*L*, ITS)* were

distantly related species within subg. *Vetrix* based on both nrDNA and cpDNA. But, the species *S. amplexicaulis* was particularly determined as the most distant taxa based on data from all studied regions with the high values in subg. *Vetrix* as it was observed in morphological data. At the end, it appeared that rather than subg. *Salix*, the members of subg. *Vetrix* had complex relations among taxa.

The overall divergence in *trn*L region is also reached to a high level (0.363) for world *Salix* comparisons. This might be explained with being *trn*T-F as non-coding and conserved part of cpDNA (Taberlet *et al.,* 1991).

In world *Salix* species, there is no genetic divergence between OWS-NWS and Turkish *Salix*. Thus, it was revealed that nearly all OWS and NWS scattered throughout the Turkish *Salix* species in phylogeny based on all cpDNA and nrDNA regions. OWS genetically has close relations with subg. *Salix* members whereas NWS are always near to Subg. *Vetrix* in regards to low genetic divergence level.

4.4. Haplotype frequency analysis, analyses of molecular variance (AMOVA), Estimation of *F*st value and Divergence times

Haplotype diversity in three cpDNA regions for Turkish *Salix* species was high. The most diverse haplotype in *mat*K region (23 haplotype) are identified in Turkish *Salix* species. Haplotype sharing was increased for both *trn*L and *rbc*L gene regions, but they were still high (13). These haplotype sequences are fixed in the *Salix* species (Futuyma, 2011). Two scenarios for high haplotype diversity can be suggested, one is the incomplete lineage sorting with many different haplotype lines in where genealogical event occurs in gene pool of individual but not informative about relations among species (Wayne and Knowles, 2006). The other one is chloroplast introgression (chloroplast capture) in which this organelle chloroplast can be laterally transmitted between species thus changing haplotype composition (Stegemann *et al.,* 2012). Haplotype sharing among infrageneric level due to incomplete lineage sorting and chloroplast introgression always appear in biogeographically close and local areas (Percy *et al*., 2014). Anatolia's biogeography is so variable having wetlands, steppes, high mountains in where appropriate climatic condition may facilitate the lateral haplotype transfer in Turkish *Salix* sp. Thus, this situation support that the

closely related geographic habitats facilitate the chloroplast capturing among *Salix* species (Stegemann *et al.,* 2012) .

Haplotype sharing (chloroplast capturing) can be common within the geographically close members of subgenus (*Salix* or *Vetrix*), but not in between two subgenera of *Salix* as results of Analysis of Molecular Variance (AMOVA) for Turkish subgenera. Non-coding cpDNA *trn*L region is the most discriminative one with high variation between subg. *Salix* and subg. *Vetrix*. The *F*st values is also significant between two subgenera of Turkey based on *trn*L region. Because the *trn*L region is highly conserved for members of subg. *Salix*, limited gene transfer between subgenera may explained by the phenomena that ancestral nature of subg. *Salix.* It was also provided by Wu *et al.* (2015) who found low haplotype sharing between subg. *Salix* and subg. *Vetrix*. Azuma *et al*., (2000) also supported this idea with the statement that the hybridization events generally occurred not between but within subgenera of *Salix*. The geographic patterns of divergence can be explained that Subg. *Vetrix* members geographically located in specific habitat and high elevation whereas subg. *Salix* members have wide distribution throughout Turkey. The current data are consistent with Skvortsov's (1999) interpretations that various members of subg. *Vetrix* underwent diversification at specific habitats such as high latitudes so that actual chloroplast transfers may be responsible for the presence of current haplotype composition of Turkish subgenera of *Salix* sp. (Hardig *et al.,* 2010).

In the World AMOVA comparison, Turkish and World willows showed high value of differention with significant *F*st value based on regions of cpDNA (*trn*L and *rbc*L) which support the haplotype sharing in local areas. However, not only intergeneric level, but also infrageneric level sharing was detected in *mat*K gene region and within each groups (Turkey, NWS and OWS).

Pairwise *F*st values showed that there were no significant relations between subg. *Salix*-OWS and subg. *Vetrix*- NWS for both *trn*L and *mat*K gene regions. On the other hand, significant *F*st values were observed between Turkish subgenera in *trn* L region. Thus, pairwise *F*st values and AMOVA analysis revealed that for cpDNA *mat*K and *trn*L gene regions, the high variation within groups and differences between Turkish subgenera are a result of contribution of OWS and NWS. Thus,

according to pairwise *F*st values rather than geographical distribution, taxonomical relations are important between subg. *Salix*-OWS and subg. *Vetrix*- NWS.

4.5. Molecular Clock

Molecular clock is estimated by the ages of haplotypes based on cpDNA barcode region; *trn*L, *mat*K, *rbc*L. The oldest molecular divergence time was dated back to 19.4 Mya between OWS and Turkey (*trn* L). The non-coding region of cpDNA *trn*L gene region molecular clock analyses showed that the estimated divergence time of Old World *Salix* species, especially Iranian *Salix*, from Turkish species assessed as ancient . The obtained divergence date for Turkish subgenera based on *trn*L region is recently (14.5 Mya).

The *Salix* genus divergence from other Salicaceae family members based on *mat*K and *rbc*L genes goes back to very old times which was dedicated by Percy *et al.,* (2015) as 35 Mya (*Populus*) and 20 Mya (from other Salicaceae members).In contrast to non-coding *trn* L gene region, for other encoded cpDNA gene regions (*mat*K and *rbc*L), the divergence times were dated back to very recent times. Since both are exonic encoded cpDNA regions, the parsimony informative sites are very low compared to non-coding cpDNA region (*trn* L). So it is show that much slower evolution occurred in *mat*K and *rbc*L regions than that is in *trn* L region. NWS diverged from Turkish subgenera 3.6 Mya (*mat*K and *rbc*L) and OWS divergence time dated back to 4.5Mya (*rbc*L). The only incongruence for cpDNA *mat*K and *rbc*L was observed in divergence of NWS from OWS with different divergence time. This can be clarified by the different species sequence used from the world for every region which changes molecular clock estimation.

The divergence of species in the Russian willow (Barkalov and Kozyrenko, 2014) was very close to today as result of high similarity of changes in plastid genome which may marked that recent divergence of species likewise Turkish willows based on *mat*K and *rbc*L gene result. Wu *et al.* (2015) obtained molecular dating time as 33.99 (The New and Old World) and 23.76 (subg. *Salix* and subg. *Vetrix*) which are about 4.5 (*rbc*L) and 3.6 respectively in the current data set. As in our *mat*K and *rbc*L molecular dating results, very low parsimony informative sites were found in

Salicaceae which can be an indicator that the species are undergoing recent evolution are also changes shown by Leskinen and Alström-Rapaport study (1999).

Two main clades of Turkish *Salix* species even if they shared the same biogeography are differentiated in the divergence times for coding (3.6 Mya) and non-coding cpDNA gene regions (14.5 Mya). When we consider the oldest fossil of subg. *Salix* from North America, it dates back to Eocene (33.9-56 Mya) , and subg. *Vetrix* in Alaska Oligocene back to 23-33.9 Mya (Wolfe, 1988). The occurrence time of European and Russian *Salix* fossil also dated back to Oligocene (Collinson, 1992). These time frames, in the current *mat*K and *rbc*L region data set are not so informative but it is in harmony with our world phylogenetic analysis in which refers close taxonomical relations of subg. *Salix*-OWS and subg. *Vetrix*-NWS combinations. Kasaplıgil (2015) provides information about of *Salix* fossil record. The age of *Salix* fossil was determined as Pliocene (2.6-5.3Mya) which confirmed the divergence time of Turkish *Salix* subgenera (3.6Mya). Thus, the last glacial period starting from 2.6 Mya (late Pliocene) has an effect on shaping composition of *Salix* sp. subgenera (Ledig, 1998).

To conclude, the data from molecular clock and phylogenetic analysis supported that presence of OWS and Turkish subg. *Salix* members in Anatolia dating back to ancient times whereas occurrence of NWS and Turkish subg. *Vetrix* was more recently. Thus, the timing of the origin of *Salix* species remains unclear but the current data supported the migration of *Salix* species from Old World countries to New World areas. Key element of this migration system might be Bearing Land Bridge which is a way from Asia to North America (Wu *et al*., 2015) and causing *Salix* transportation. Additionally, Pliocene origin of Turkish subgenera confirmed by the molecular data. Turkish subgenera diverged recently (late Pliocene) and undergone recent evolution.

4.6. Construction of Phylogenetic Trees

There are many classification systems exist for *Salix* at generic and infrageneric levels (Fang, 1987; Argus, 1997; Hardig, 2014). In the current study, the system of Skvortsov (1999) is followed where Turkey *Salix* L. has two subgenera (*Salix* and

Vetrix). Our cpDNA and nrDNA data separated the two well supported clade as Subg. *Salix* and Subg. *Vetrix*. These two well supported clades are based on *rbc*L gene tree, is also observed in Japanese *Salix* (Azuma *et al.,* 2000) and Chinese *Salix* sp. (Chen *et al*., 2010).

The *S. acmophylla* (*mat*K and *ITS*), *S. babylonica* (*trn*T-F, *rbc*L and *ITS*), *S. triandra* subsp. *triandra*, *S. triandra* subsp. *bornmuelleri* (*rbc*L, *mat*K) were found to be distant from other members of subg. *Salix*. The similar placement of the *S. acmophylla* in cladograms (Hardig *et al*., 2010) based on *mat*K, *ITS* and *trn*L-F (Abdollahzadeh *et al*., 2011) was explained by either introgression or incomplete lineage sorting (Hardig *et al.,* 2010). *S. acmophylla* may be the hybrid originating from one parent definitely as *S. alba* due to the presence of polymorphic sites (Abdollahzadeh *et al.,* 2011; Barkalov and Kozyrenko, 2014a). *S. acmophylla* is naturally found in Eastern part of Turkey, was well allied far from a clade of Subg. *Salix*. Thus, the other scenario for distant position of *S. acmophylla* can be the effect of the Anatolian Diagonal which may be more important geographic barrier in causing differentiation of taxa at the species and subspecies level (Bilgin, 2011). Because of the exotic origin, *S. babylonica* might be differentiated from others (Azuma *et al*., 2000). Finally, the presence of polymorphic sites and polyploidy characterization, the indicators of hybridization, can explain the conflicting distant placement of *S. triandra* in subg. *Salix* (Leskinen and Alström-Rapaport, 1999; Abdollahzadeh *et al*., 2011, Barkalov and Kozyrenko, 2014). Hybrids between *S. alba* and *S. fragilis* naturally occur in Europe (Hamza-Babiker *et al*., 2009). The morphologically identified hybrid species, *S. alba x fragilis* located in all trees closer to *S. alba* and *S. fragilis* that it was supported by the intermediate morphology data and molecular data. Subg. *Salix* including *S. alba*, *S. fragilis*, *S. alba x fragilis*, *S. excelsa*, *S. babylonica* made one cluster in cpDNA tree cladogram of Hamza-Babiker *et al*. (2009), and also in cladogram of Abdollahzadeh *et al.* (2011) *ITS* and *trn*L-F tree. The natural group, Subg. *Salix* with many primitive ancestral characters are dispersed in warm temperatures of the world (Azuma *et al.*, 2000; Hardig *et al.,* 2010; Abdollahzadeh *et al.,* 2011). Except for the clade Triandrae involving *S. triandra*, studies revealed the monophyletic nature of the subg. *Salix* (Chen *et al.,* 2010; Abdollhazadeh *et al.,* 2011). The more slowly evolving gene provides better information on more ancient divergence (Futuyma, 2011). The coding region of

cpDNA (*mat*K and *rbc*L) are found as more conserved and have a slow evolutionary rate than non-coding cpDNA (*trn*T-F) and nrDNA *ITS.* Thus, it could be said that the phylogeny with more branches in *mat*K and *rbc*L gene tree gave more information about ancient form of subg. *Salix.*

Subg. *Vetrix* members including *S. amplexicaulis (mat*K*, rbc*L*)* and *S. rizeensis (mat*K*)* are placed at subg. *Salix* rather than Subg. *Vetrix* in phylogenetic trees constructed with coding region of cpDNA. Also, *S. amplexicaulis* got the basal position of subg. *Vetrix* clade based on *trn*T-F and *ITS* sequence. It is observed that *S. amplexicaulis* can make mixed populations in warmer habitats with other members of Subg. *Salix*, which may facilitate the hybridization between them. Since *S. rizeensis* is endemic to Northern part of Turkey, widespread hybridization is not the case for it as in the case of *S. amplexicaulis*. Thus, we suggest that it will be meaningful to transfer *S. amplexicaulis* to subg. *Salix*. For the replacement of *S. rizeensis*, it needs more supportive data. An other common pattern in subg. *Vetrix* clade is it's externally attached species which are *S. elbursensis* (*trn*T-F, *mat*K), *S. viminalis* (*rbc*L, *mat*K), *S. aegyptiaca* (*rbc*L, *mat*K), *S. wilhelmsiana* (*rbc*L, *mat*K). All of the species in subg. *Vetrix* in all related cpDNA analysis came out in polytomy like similar results reported by Hardig *et al.* (2010); Chen *et al.* (2010). On the other hand, subg. *Vetrix* members didn't form a real clade in nrDNA *ITS* tree. The high observed polymorphism as a result of high nucleotide diversity based on all studied regions in subg. *Vetrix* is resulted in high degree of polytomy. The naturalness of subg. *Vetrix* is uncertain as stated by Hardig *et al.* (2010), and also suggested that polyphyly in this subg. *Vetrix* is the result of hybridization. Most of subg. *Vetrix* members are naturally found in cooler habitats like northern parts of Turkey. Thus, the reticulated relation in subg. *Vetrix* may also be resolved by hybridization in specific habitat conditions such as cooler climates (Lauren-Moreau *et al.,* 2015) or high altitude (Skvortsov, 1999) which facilitates the chloroplast transfers. Besides, the common polytomy structure of subg. *Vetrix* was observed in all cpDNA and nrDNA phylogeny support of the newly evolving nature of this subgenera. The most recent and rapid diversification of Subg. *Vetrix* among other subgenera (*Chamaetia, Chosenia, Longifoliae, Protitea* and *Salix*) were also put forward by Lauren-Moreau *et al*. (2015). Thus, all selected genes (cpDNA and nrDNA) showed polytomy for subg. *Vetrix* members supported the younger nature of subg. *Vetrix*.

The phylogenetic trees constructed for comparisons of world *Salix* are evaluated for every region. Using molecular clock results and the constructed phylogenetic trees based on each studied region, New world *Salix* and Old World *Salix* species did not form a distinct group in all phylogenetic trees. For the comparison of aligned *trn*L region, the OWS dispersed throughout Turkish species regardless of geographic distribution in the tree, but they were still together among subg. *Salix* and *Vetrix.* Different from that, for other cpDNA regions (encoded *mat*K and *rbc*L), every NWS and OWS species are widely scattered throughout the Turkish *Salix* according to taxonomically related subgenera (*Salix* or *Vetrix*) which they are in. In other words, some of the OWS and NWS members which are belonging to subg. *Salix* are genetically close to members of Turkish subg. *Salix* and other OWS and NWS from subg. *Vetrix* are grouped with Turkish *Vetrix* based on encoded cpDNA regions (*mat*K and *rbc*L). In *ITS* phylogenetic tree show the same composition for world species likewise other cpDNA (*mat*K and *rbc*L) trees in which NWS and OWS dispersed along the subg. *Salix* and. *Vetrix.* In short, Turkish Subg. *Salix* members made a genetically distinct clade and the other NWS and OWS species nested among the subg. *Vetrix* members in the tree based on *ITS* data set. This incongruence between chloroplast and nuclear gene data sequence phylogenies can be explained by haplotype sharing as result of chloroplast capturing (Stegemann *et al*., 2012)

For the subg. *Salix,* close relation with OWS and NWS data set in cpDNA phylogenetic tree can be described by wide distribution of subg. *Salix* members and also, ancient haplotype capturing events (Hardig *et al.,* 2010, Wu *et al.*, 2015). Wu *et al.* (2015) underlined the wide range distribution of subg. *Salix* in the Old World and New World as result of climatic cooling. It can also be explained by the chromosomal data that most of the subg. *Salix* are diploid (Hamza-Babiker *et al.,* 2009). Conversely, the case showed itself up in different way for nrDNA *ITS* tree that subg. *Salix* made distinct cluster rather than having close relations with world species. NWS and OWS are nested among subg. *Vetrix* members based on *ITS* dataset. The most recent study (Lauron-Moreau *et al.,* 2015) have concluded similarly that Subg. *Salix* and Subg. *Vetrix* separated from each other and the NWS are nested between American subgenera members based on their *rbc*L and *ITS* molecular datasets.

The Salicaceae members (*Populus* and *Salix*) have common ancestors and they are monophyletic (Liu *et al*., 2015). According to Skvortsov (1999) subg. *Salix* should be considered as a natural (not polyphyletic) group. Turkish *Salix* species phylogenetic tree for all studies region including cpDNA showed monophyly with having two well supported clades as subg. *Salix* and *Vetrix*. However, in nrDNA *ITS* phylogenetic tree is potentially monophyletic since subg. *Salix* formed one main clade and the subg.*Vetrix* members were attached externally. Previous studies on nrDNA *ITS*, *rbc*L and the combined *atp*B-*rbc*L-*trn*D-T sequences (Brunsfeld *et al.,* 1991; Azuma *et al.,* 2000) showed that the traditionally recognized subgenera *Salix* and *Vetrix* are not monophyletic. But recently Leskinen and Alström-Rapaport (1999), Chen *et al.* (2010) and Lauren-Moreau *et al.*(2015) supported the monophyletic origin of *Salix* sp. Furthermore, Hardig *et al.* (2010) studying, all subgenera including *Salix* and *Vetrix* stated that monophyletic or potentially monophyletic nature of the genus. Thus, for Turkish *Salix* sp., Subg. *Salix* was always formed one main clade with a high bootstrap value. On the other hand, polytomy is found in Subg. *Vetrix* for all related regions. In Barkalov and Kozyrenko's study (2014), subg. *Vetrix* group formed a clade in phylogenetic willow tree based on cpDNA and *ITS* combination that is polyphyletic whereas subg. *Salix* is monophyletic, which is very similar to our study. Leskinen *et al.,* (1999), Hardig *et al*. (2010) and Abdollahazadeh *et al.* (2011), reported similar results having extensive polytomy in subg. *Vetrix.*

To sum up, Turkish Subgenera are potentially monophyletic, but polyphyly can be formed within each subgenus (*Salix* and *Vetrix*) respecting the gene considered due to both cpDNA and nrDNA introgression. Two distinct subgenera were formed in every cpDNA phylogenetic trees and almost in nrDNA *ITS* tree for Turkish *Salix* sp. The clear separation of two genera can be result of the effect of different biogeographic patterns on he related species (Lauren-Moreau *et al.,* 2015). Subg. *Salix* prefer warmer habitats like most parts of Turkey whereas subg. *Vetrix* species are found cooler climates such as in Northern Turkey.

Many groups of plants can be perfectly identified by using DNA barcodes. However others such as *Salix* is difficult to identify (Hollingsworth *et al.,* 2011). The current data set did not provide much information in species system of Turkish *Salix* to make

taxonomical classification. The sectional classification of *Salix* is also not clear in the world. A main reason this may be the fact that most sectional willow systems are only based on localized flora (Wu *et al.,* 2015). High seed dispensability and introgressive hybridization resulted in chloroplast capturing events in local areas. In this sense, DNA barcoding system, especially plastid markers are not enough to identifying *Salix* species due to reticulate relations (Percy *et al.,* 2014). None of the previous molecular phylogenetic studies has covered the complex Turkish *Salix* sp. in detail as much as the current study.

CHAPTER 5

CONCLUSION

The aim of the current study was to reveal interspecific morphological, phylogenetic and evolutionary relationships of Turkish *Salix* L. by using DNA sequences of 24 species and one hybrid from both nuclear and chloroplast DNA regions.

The selected morphological characters, especially as petiole length (Pl) and bud angle (Ba) seem to be useful to make subgeneric level identification of Turkish *Salix* genus. The pubescence on bud scale character is discriminative for subg. *Vetrix* members adapted to cooler climatic condition whereas tree-like life form and lanceolate leaf shape distinguished subg. *Salix* members from others. The morphological and the molecular data results agree with the traditional taxonomic concepts in clustering Turkish willows into two subgenera. The morphological data supported the close position of the taxa, *S.amplexicaulis* to subg. *Vetrix* rather than subg. *Salix*.

The highest variable sites and nucleotide diversity were found in non-coding cpDNA *trn*T-F, especially in *trn*L intronic region of *Salix* sp. The *rbc*L gene alone is too conserved and effectiveness can be raised by combining with cpDNA *mat*K region. The *ITS* region as a nuclear gene region and had the richest GC content and informative sites with respect to all the studied regions of cpDNA (*trn* T-F, *mat*K, *rbc*L) regions. Most of *ITS* gene substitutions were informative in discriminating two subgenera of Turkish *Salix* sp. The variable sites were higher in Subg. *Vetrix* than Subg. *Salix* for all gene regions. While two coding cpDNA gene regions (*mat*K and

*rbc*L) were found as conserved and non-coding cpDNA (*trn*T-F) and nrDNA *ITS* evolved rapidly for Turkish *Salix* sp.

When overall genetic distance among the studied species is compared, the highest was in *trn*T-F region due to non-coding nature of this region. *S. acmophylla* from subg. *Salix* and *S. amplexicaulis* from subg. *Vetrix* were categorized as the most distant taxa. As far as the species divergence level is concerned, the closest relationships were identified among subg. *Salix* species whereas the members of subg. *Vetrix* had distant and complex relations among the taxa.

The constructed phylogenetic tree for Turkish willow species showed monophyly with having two well supported clades in cpDNA regions and potentially monophyletic in nrDNA *ITS*. The main mechanism that causes the conflicts in all phylogenetic trees based on cpDNA and nrDNA, are introgressive hybridization and incomplete lineage sorting in Turkish willows. Combined sequence data from (*trn*T-F, *mat*K, *rbcL* and *ITS*) enabled a reliable subgenus-level classification (*Salix* and *Vetrix*) of Turkish willows. Additionally, subg. *Vetrix* group made a clade that is polyphyletic whereas subg. *Salix* is monophyletic among the main phylogenetic trees. This division of Turkish *Salix* subgenera based on molecular data is explained by that the subg. *Vetrix* members having morphologically adaptive characters prefer cooler habitat of Turkey (northern Anatolia) while Subg. *Salix* prefer warmer habitat (inner Anatolia). Intense hybridization resulted the high level of observed polymorphisms in specific habitats are the reasons of polytomy for subg. *Vetrix* clade. As far as the taxa positions in the phylogenetic trees are considered, *S. acmophylla*, *S. triandra* and *S. babylonica* had the distant position to Subg. *Salix*. The distant position of *S. acmophylla* can be determined either by biogeographically isolated position or by hybrid origin of it. The morphologically identified hybrid species, *S. alba x fragilis,* located in all trees closer to *S. alba* and *S. fragilis.* The distant appearance of the *S. amplexicaulis* in the phylogenetic tree, with the representatives belonging to the subg. *Vetrix*, can be clarified by the hybridization with the species from the subg. *Salix* in shared habitats. Therefore, it would be better to move *S. amplexicaulis* from subg. *Vetrix* to subg. *Salix*.

The diverse haplotype was detected in *mat*K region (23) and haplotype sharing was increased for both *trn*L and *rbc*L gene regions, but they are still high (13). Those haplotype sequences have been fixed in the related species whereas others are eliminated. Incomplete lineage sorting with many different haplotype and introgressive hybridization are the phenomena that likely determine the haplotype composition of Turkish *Salix* sp. Haplotype sharing (chloroplast capturing) can be common within the geographically close members of subgenus (*Salix* or *Vetrix*) but not between two subgenera of *Salix*. Anatolia's varied regional and specific climate seem to create suitable environments for haplotype sharing within subg. *Vetrix* species (higher altitude and cooler climates) and within subg. *Salix* species (warmer climate).

To understand evolutionary relationships between Turkish willows and the species from the world, the DNA sequences of studied regions of cpDNA were gathered from NCBI database. The constructed phylogenetic trees based on each studied region revealed that New world *Salix* and Old World *Salix* species are scattered throughout the Turkish subgenera in all phylogenetic trees. Especially, in encoded cpDNA regions (*mat*K and *rbc*L) based phylogenetic trees, NWS and OWS nested among the Turkish *Salix* according to subgenera (*Salix* or *Vetrix*) which taxonomically they are in. However, subg. *Salix* made a distinct clade in nrDNA *ITS* tree. Wide distribution of subg. *Salix* members and ancient haplotype capturing events may have caused for distinct separation. The incongruence between chloroplast and nuclear gene data sequence phylogenies can also be explained by chloroplast capturing.

The AMOVA analysis with all *Salix* species revealed that according to cpDNA *mat*K and *trn*L data sets, the high variation within groups and differences between Turkish subgenera are a result of contribution of OWS and NWS sequneces among Turkish *Salix* species. Furthermore, no significant pairwise F_{st} values were detected between subg. *Salix*-OWS and subg. *Vetrix*- NWS (*trn*L and *rbc*L). Thus, rather than geographical distribution, taxonomical relations may have determined the molecular relations between subg. *Salix*-OWS and subg. *Vetrix*-NWS.

Evolutionary divergence times for Turkish *Salix* genera were estimated on the average as 14.5 Mya for *trn*L and 3.6 Mya for *mat*K and *rbc*L gene regions. Much slower evolution was indicated in *mat*K and *rbc*L region than *trn*L regions. Furthermore, divergence time of NWS and OWS was about 4.5 Mya (*rbc*L). In molecular dating, to make direct comparison with the numbers, the combination of cpDNA data will be more informative for Turkish willows. It is concluded that occurrence of ancient OWS and subg. *Salix* members dated back much earlier whereas existence of NWS and Turkish subg. *Vetrix* dates back much recently. NWS species and Turkish subg. *Vetrix* members which are in close relations diversified in Anatolia recently. The timing of origin of *Salix* species remains unclear, however, the molecular data set supported the migration route of *Salix* species from Old World to New World. Intercontinental dispersal occurred from Asia to America for *Salix* genus. Turkish subgenera diverged recently (late Pliocene) and undergone a recent evolution.

Morphology based systematic treatment of *Salix* varied widely, and need more attention. Additional taxa and sequences from other nrDNA and cpDNA regions may provide further insights to understand phylogenetic relationships among Turkish *Salix* species at both sectional and species levels.

REFERENCES

Abdollahzadeh, A., Kazempour, O.S., and Maassoumi A. A., Molecular Phylogeny Of the Genus *Salix* (Salicaceae) with an Emphasis to its Species in Iran, Molecular phylogeny of the genus *Salix* (Salicaceae) with an emphasis to its species in Iran. -*Iran. J. Bot.* 17(2): 244-253, 2011.

Alvarez, I., and Wendel, J.F., Ribosomal *ITS* sequences and plant phylogenetic inference, *Molecular Phylogenetics and Evolution* 29:417–434, 2003.

Anderson, E. and Stebbin, G.L., Hybridization as Evolutionary Stimulus. *Evolution*. 8(4): 1954.

Andersson N.J., *Salix.*—In: A. P. De Candolle (ed.). *Prodromus systematis naturalis regni vegetabilis*, 16 (2): 190–323. Genève, 1868.

Aravanopoulos, F.A., Kimb, K.H. and Zsuffab, L., Genetic diversity of superior Salix clones selected for intensive forestry plantations, *Biomass and Bioenergy, 16*, 249-255, 1999.

Argus, G. W., *The genus Salix* (Salicaceae) *in the southeastern United States*. Syst. *Bot. Monogr.* 9: 1–170, 1986.

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.

Arıhan, O. and Güvenç, A., Ankara çevresinde yetişen 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 (Turk J Bot),* 35, 535-551, 2011.

Avcı, M., and Zielinski, J., *Salix myrisinifolia* Salisb. (Salicaceae)-A new species for Flora of Turkey.*Karaca Arboretum Magazine,* 4(2), 49-54, 1997.

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

Azuma, T., Kajita, T., Yokoyama, J. and Ohashi, H., Phylogenetic relationships of *Salix* (Salicaceae) based on rbcL sequence data*, American Journal of Botany* 87(1): 67-75, 2000.

Babayi F., Pakravan, M., Maassoumi, A. A. and Tavasoli, A., Palynological study of *Salix* L. (Salicaceae) in Iran, *Iran. J. Bot.*,18 (1): 118-126, 2012.

Bakker, F.T., Culham, A., Gomez-Martinez, R., Carvalho, J., Compton, J., Dawtrey, R. and Gibby, M., Patterns of nucleotide substitution in angiosperm cpDNA *trn*L (UAA)-*trn*F (GAA) regions, *Molec. Biol. Evol.* 17:1146-1155, 2000.

Baldwin, B., Sanderson, M.J. and Porter, M., The *ITS* region of nuclear ribosomal DNA a valuable source of evidence on angiosperm phylogeny, *Ann. Missouri Bot. Garden*, 82: 247-277, 1995.

Baldwin, B.G. and Markos S., Phylogenetic utility of the external transcribed spacer (*ETS*) of 18S-26S rDNA: congruence of ETS and ITS trees of Calycadenia (Compositae). *Mol Phylogenet Evol.,* 10(3):449-63, 1998.

Barkalov, V.Y. and Kozyrenko, M.M., Phylogenetic Analysis of the Far Eastern *Salix* (Salicaceae) Based on Sequence Data from Chloroplast DNA Regions and *ITS* of Nuclear Ribosomal DNA, Botanica Pacifica, *A journal of plant science and conservation*, 3(1): 3-19, 2014b.

Barkalov, V.Y. and Kozyrenko, M.M., Phylogenetic relationships of *Salix* L. subg. *Salix* species (Salicaceae) according to sequencing data of intergenic spacers of the chloroplast genome and ITS rDNA*, Russion Journal of Genetic*, 50(8):828-837, 2014a.

Beismann, H., Barker, J., H., A., Karp, A. and Speck, T., AFPL Analysis sheds light on distribution of two *Salix* species an their hybrid along a natural gradient, *Moleculer Ecology*, 6: 989-993, 1997.

Bilgin, R., Back to the Suture: the distribution of intraspecific genetic diversity in and around Anatolia. *Int. J. Mol. Sci*. 12: 4080-4103, 2011.

Bousquet, J., Strauss, S.H., Doerksen, and Price, A.H., Evolution Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants, *Proc. Natl. Acad. Sci.*. 89, 7844-7848, 1992.

Brunsfeld, S.J., Soltis, D.E. and Soltis, P.S., Patterns of genetic variation in *Salix* section Longifoliae (Salicaceae), *American Journal of Botany*, 78, 855-869, 1991.

Camin, J.H. and Robert, R.S., A Method for Deducing Branching Sequences in Phylogeny, *Evolution,* 19(3), 311-326, 1965.

CBOL*, Plant Working Group A DNA barcode for land plants*. Proc Natl Acad Sci USA., 2009.

Chen, J., Sun, H., Wen, J. and Yang, Y. Molecular Phylogeny of *Salix* L. (Salicaceae) Inferred from Three Chloroplast Datasets and Its Systematic

Implications. International Association for Plant Taxonomy (IAPT). *Taxon*, 59, (1):29-37, 2010.

Clegg, M.T. and Zurawski, G., *Chloroplast DNA and the study of plant phylogeny. In: Soltis PS, Soltis DE, Doyle JJ (eds),* Molecular systematics of plants. Chapman and Hall, London, 1-13, 1993.

Clegg, M.T., Ritland, K. and Zurawski, G., *Processes of chloroplast DNA evolution,* In ''Evolutionary Processes and Theory'' (S. Karlin and E. Nevo, Eds.), Academic Press, New York, 275-294, 1986.

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

Cronquist, A., *Salix*. In: Hitchcock, C. L., A. Cronquist, M. Ownbey, and J. Thompson. Vascular Plants of the Pacific Northwest. Univ. Washington Publ. Biol. 17(2) : 37–70, 1964.

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

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

Dorn, R.D., Cytological and taxonomic notes on North American Salix. *Madroño*, 23: 99, 1975.

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.

Ehleringer, J.R. and Björkman, O., A Comparison of Photosynthetic Characteristics of *Encelia* Species Possessing Glabrous and Pubescent Leaves, *Plant Physiol.* 62, 185-190, 1978.

Erinç, S., *Changes in the physical environment in Turkey since the end of the last glacial,* p. 87-110. In W.C. Brice (ed.), The environmental history of the Near and Middle East since the last ice age. Academic Press, London, 1978.

Evlard A., Druart, P. and Collinet, G., *Using Salix spp. in phytostabilization of metal pollution in soils : an example of phytoremediation appropriate to the brownfields of Wallonia.* Poster in: 19th National Symposium on Applied Biological Sciences. Gembloux, Belgium, 2014.

Excoffier, L., Laval, G. and Schneider, S., Arlequin ver. 3.0: An integrated software package for population genetics data analysis, *Evolutionary Bioinformatics Online*, 1:47-50, 2005.

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

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.

Fritz, R., S., Orians, M.N. and Brunsfeld, S.J., Interspecific hybridization of plants and resistance to herbivores: hypotheses, genetics, and variable responses in a diverse herbivore community, *Oecologia*, 97:106-117, 1994.

Futuyma, D.J., *Evolution,* Sinauer Associates, Sunderland, second edition, MA. 44, 2011.

Gielly, L. and Taberlet, P., The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbc*L sequences*, Mol. Biol. Evol.* 11: 769-777, 1994.

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.

Hamza-Babiker, N., Berthold, H., Gloss, J. and Arnold, C., Chloroplast DNA identification of eight closely related European *Salix* species, *Austrian Journal of Forest Science,*126(3), 175-193, 2009.

Hamzeh, M. and Dayanandan, S., Phylogeny of *Populus* (Salicaceae) based on nucleotide sequences of chloroplast *trn*T– *trn*F region and nuclear rDNA. *American Journal of Botan*y, 91(9): 1398-1408, 2004.

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.

Hasegawa, M., Kishino, H. and Yano, T., Dating of human-ape splitting by a molecular clock of mitochondrial DNA, *Journal of Molecular Evolution,* 22(2): 160–174, 1985.

Herrera, S., Bonkers about biofuels, *Nature Biotechnology,* 24(7), 755-660, 2006.

Hilu, K.W. and Liang, H., The *mat*K gene: sequence variation and application in plant systematics*. Am. J. Bot.* 84:830-839, 1997.

Hilu, K.W., Borsch, T., Muller, K., Soltis, D E., Soltis, P.S., Savolainen, V., Chase, M.W., Powell, M.P., Alice, L.A., Evans, R., Sauquet, H., Neinhuis, C., Slotta, T.A.B., Jens, G.R., Campbell, C.S. and Chatrou, L.W., Angiosperm phylogeny based on matK sequence information, *Am. J. Bot.* 90: 1758-1776, 2003.

Hollingsworth, P.M. and Graham, S.W., Little D.P., Choosing and Using a Plant DNA Barcode. *PLoS ONE* 6(5): e19254, 2011.

Hsiao, C., Chatterton, N.J., Asay, K.H., and Jensen, K.B, Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae(Poaceae), inferred

from nuclear rDNA (internal transcribed spacer) sequences, *Genome,* 38: 221-223, 1995.

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

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

Kimura, A., Über Toisusu, eine neue Salicaceen-Gattung und die systematische Stellung derselben. *Botanical Magazine (Tokyo)*, 42: 287–290, 1928.

Koch, W., *De Salicibus Europaeis commentatio*. Erlangae, 1828.

Kress, W.J., Wurdack K.J., Zimmer, E.A., Weigt, L.A. and Janzen, D.H., Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci*., 102: 8369–8374, 2015.

Lauren-Moreau, A., Pitre, E.F. Argus, G.W., Labrecque, M., and Brouillet, L., Phylogenetic Relationships of American Willows (*Salix* L., Salicaceae), *PLOSONE*, 2015.

Ledig, F.T., *Genetic Diversity in tree species, With special reference to conservation in Turkey and the eastern Mediterranean, The proceedings of International Symposium on In situ Conservation of Plant Genetic Diversity*, 231-247, CRIFC, Turkey, 1998.

Leskinen, E. and Alström-Rapaport, C., Molecular phylogeny of Salicaceae and closely related Flacourtiaceae: evidence from *5.8* S, *ITS* 1 and *ITS* 2 of the rDNA, *Plant Systematic and Evolution,* 215, 209-227, 1999.

Levsen, N., Tiffin, P., and Olson, M., Pleistocene speciation in the genus Populus (Salicaceae), *Systematic Biology*, 2012.

Li, J.H., Bogle, A.L., and Klein, A.S., Interspecific relationships and genetic divergence of the disjunct genus *Liquidambar* (*Hamamelidaceae*). *Rhodora,* 99: 229241, 1997.

Librado, P. and Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25:1451-1452, 2009.

Linnaeus, C., *Species plantarum, exhibentes plantas rite cognitas, ad genera relatas, cum differentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas.* Holmiae, Impensis Laurentii Salvii. L. Salvius, Stockholm., 1753.

Liu, X., Wang, Z. and Dongsheng, W., Phylongeny of *Populus-Salix* (Salicaceae) and their relative genera using molecular data set, *Biochemical and Systematic Ecology,* 68, 210-215, 2016.

Mahdi, J. G., Mahdi, A.J. and Bowen, I.D., The historical analysis of aspirin discovery, its relation to the willow tree and antiproliferative and anticancer potential, *Cell Proliferation Journal (Cell Prolif.),* 39, 147-155, 2006.

Mallet, J., Hybrid Speciation, *Nature,* 446/ 15. doi:10.1038/nature05706, 2007.

Masood M.S., The complete nucleotide sequence of wild rice (*Oryza nivara*) chloroplast genome: first genome wide comparative sequence analysis of wild and cultivated rice, *Gene,* 340:133–139, 2004.

MATLAB and Statistics Toolbox Release, The MathWorks, Inc., Natick, Massachusetts, United States, 2012.

Ministary of Enviroment and Forestry, '*Convention of Biological Diversity Fourth National Repor*t, Republic of Turkey, 2009.

Mishra, A., Agarwal, S., Jain, C.K., Rani, V., High GC content: Critical parameter for predicting stress regulated miRNAs in "*Arabidopsis thaliana*, *Bioinformation,* 4(4): 151-154, 2009.

Mort, M.E., Soltis, D.E., and Soltis, P.S., Francisco-Ortega, J. and Santos-Guerra, A., Phylogenetic Relationships And Evolution of Crassulaceae Inferred From *mat*K Sequence Data, *American Journal of Botany* 88(1): 76–91, 2001.

Mosseler, A., Hybrid performance and species crossability relationships in willows (*Salix*). *Can. J . Bot*. 68: 2329-2338, 1989.

Nakai, T., Chosenia, a new genus of Salicaceae*. Botanical Magazine (Tokyo)* 34: 66–69, 1920.

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

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

Neuhaus, H. and Link, G. The chloroplast Trna Lys (UUU) gene from mustard. *Curr. Genet*. 11: 251–257, 1987.

Nissen, L.R. and Lepp N.W., Baseline concentrations of copper and zinc in shoot tissues of a range of Salix species. *Biomass Bioenergy,* 12:115–20, 1997.

Orians, C.M., Daniel, I.B., Roche, B.M., Fritz, R.S. and Floyd, T., Water availability alters the relative performance of *Salix sericea*, *Salix eriocephala*, and their F1 hybrids. *Can. J. Bot,*. 77: 514–522, 1998.

Palmer, J.D., *Plastid chromosomes: structure and evolution*. In: Vasil IK, Bogorad L, editors. Cell culture and somatic cell genetics in plants. Vol. 7A: The molecular biology of plastids. San Diego (CA): *Academic Press;* 5–53, 1991.

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

Percy, D.M., Argus, G.W., Cronk, Q.C., Fazekas, A.J. and Kesanakurti, P.R., Understanding the spectacular failure of DNA barcoding in willows (*Salix*): Does this results from a trans specific selective sweep? *Mol Ecol.,* 23:4737–4756, 2014.

Pirie, M.D., Balca, M.P., Vargas, Z., Botermans, M., Bakker, F.T. and Chatrou, L.W., Ancient Paralogy In The cpDNA *trn*L-F Regıon in Annonaceae: Implications for Plant Molecular Systematics, *American Journal of Botany* 94(6): 1003–1016, 2007.

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

Przyborowski, J.A. and Sulima, P., The analysis of genetic diversity of *Salix viminalis* genotypes as a potential source of biomass by RAPD markers, *Industrial Crops and Products*, 31, 395–400, 2010.

Pulford, I.D., Watson, C., Phytoremediation of heavy metal-contaminated land by trees—a review, *Elsevier*, *Environment International,* 529–540, 2003.

Rohlf, F.J., NTSYS-pc, Version 2.1d. Exeter Software, Setauket, NY, 2000.

Qureshi, R.A., Gilani, S.A., Gilani, S.J., Sultana, K.N. and Ghufra, M.A., Palynological study of the genus *Salix* L. (Salicaceae) from Pakistan. *Pak. J. Bot.,* 39: 2257–2263, 2007.

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

Sander, M.L. and Ericsson, T., Vertical distributions of plant nutrients and heavy metals in *Salix viminalis* stems and their implications for sampling, *Biomass Bioenergy,* 14:57– 66, 1998.

Savolainen, V., Fay, M.F., Albach, D.C., Backlund, A., Bank, M., Cameron, K.M., and Johnson, S.A., Phylogeny of the Eudicots: A nearly complete familial analysis based on *rbc*L gene sequences. *Kew Bulletin,* 55: 257–309, 2000.

Schleicher, J.C., *Catalogus hucusque absolutus omnium plantarum in Helvetia sponte nascentium.* Bex. Idem, ed. 4. Camberii (1821), 1807.

Schneider, C.K., Salix. In: Ch. S. Sargent. *Plantae Wilsonianae*, 3(1):40–179. Cambridge, 1916.

Serres-Giardi, L., Belkhir, K., David, J. and Glémin, S., Patterns and evolution of nucleotide landscapes in seed plants. *Plant Cell*., 24(4):1379–1397, 2012.

Shaw, J., Lickey, E.B., Schilling, E.E. and Small, R.L., Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III, *American Journal of Botany,* 2007.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T, Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J.K., Yamaguchi-Shinozaki, Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura,

M. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression, *EMBO J.*, 5(9): 2043–2049, 1986.

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.

Sneath, P.H.A. and Sokal, R.R., Numerical taxonomy: the principles and practice of numerical classification. *San Francisco: Freeman,* 573,1973.

Sohma, K., Pollen Diversity in *Salix* (Salicaceae). Science Reports of the Tohoku University, Fourth Series, *Biology*, 40: 77–178, 1993.

Soltis, D.E. and Soltis, P.S., *Choosing an approach and an appropriate gene for phylogenetic analysis,* 2-31, Molecular Systematics of Plants II:DNA Sequencing. Kluwer Academic Publishers, Boston, 1998.

Stebbins, G., L., *Polyploidy, Hybridization, and Invasion of New Habitat.* California, 1950.

Stegemann, S., Keuthe, M., Greiner, S. and Bock, R., *Horizontal transfer of chloroplast genomes between plant species*, 109(7), 2434–2438, 2012.

Suda, Y. and Argus, G.W. Chromosome numbers in some North American *Salix. Brittonia* 20: 191-197, 1968.

Taberlet, P., Gielly, L., Pautou, G. and Bouvet, J., Universal primers for amplification of three non-coding regions of chloroplast DNA, *Plant Mol. Biol.* 17: 1105-1109, 1991.

Tajima, F., Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics,* 123(3), 585-595, 1989.

Tamura, K., Dudley, J., Nei, M. and Kumar, S., MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0, *Mol. Biol. Evol.,* 24:1596-1599, 2007.

Tamura, K., Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C content biases, *Molecular Biology and Evolution*, 9(4): 678–687, 1992.

Terzioğlu, S., Coşkunçelebi, K. and Serdar, B., Contribution to the description of an endemic Turkish Salix species, *Plant Biosystems,* 141(1), 82–85, 2007.

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

Tunçtaner, K., *Çeşitli Söğüt Klonlarının Genetik Varyasyonları ve Türkiye'nin Değişik Yörelerine Adaptasyonları Üzerine Araştırmalar*, Kavak ve Hızlı gelişen Yabancı Tür Orman Ağaçları Araştırma Enstitüsü, 1990.

Vermerris, W., Genetic Improvement of Bioenergy Crops. *Springer,* 347-362, 2008.

Watson, C., Pulford, I.D., and Riddell-Black, D., Screening of Willow Species for Tolerance to Heavy Metals: Comparison of Performance in a Hydroponics System and Field Trials. *International Journal of Phytoremediation,* 5, 351-365, 2003.

Wayne, P. and Knowles, L. L., Inferring Phylogeny Despite Incomplete Lineage Sorting, *Syst. Biol*. 55(1):21–30, 2006.

Weir, B. S., Cockerham, and C. Clark, Estimating F-Statistics for the Analysis of Population Structure. *Evolution* 38(6): 1358, 1984.

Wimmer F., *Salices Europaeae*. Vratislaviae, 1866.

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

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.

Yaltırık, F. and Efe, A., *Otsu Bitkiler Sistematiği,* İ.Ü. Publish No:3568, No: 3, İstanbul, 1989.

Yi, T., Miller, A.J. and Wen, J., Phylogenetic and biogeographic diversificationof *Rhus* (Anacardiaceae) in the Northern Hemisphere, *Molecular Phylogenetic and Evolution,* 33, 861-879, 2004.

Youngbae, S., Thien, L.B., Reeve, H.E. and Zimmer, E.A., Molecular Evolution and Phylogenetic Implications of Internal Transcribed Spacer Sequences of Ribosomal DNA in Winteraceae, *American Journal Botany,* 80(9), 1993.

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

Zimmer, E.A., Hamby, R.K., Arnold, M.L., Leblanc, D.A. and Theriot, E.C., Ribosomal RNA phylogenies and flowering plant evolution, In B. Fernholm, K. Bremer, and H. Jornvall [eds.], The hierarchy of life, 205-214. *Elsevier Science,* Amsterdam, 1989.

Zohary, M., *Geobotanical Foundations of the Middle East*. Gustav Fisher Verlag, Stuttgart/Swets & Zeitlinger, Amsterdam. XXIV , 739 pp., 279, 1973.

Zuckerkandl, E. and Pauling, L., "*Evolutionary Divergence and Convergence in Proteins," in Vernon Bryson and Henry Vogel, eds., Evolving Genes and Proteins* (*The molecular evolutionary clock is not a "metronomic"clock; rather, it "ticks" are stochastic events),* New York: Academic Press, 97-166, 1965.

APPENDICES

Appendix A: Turkish *Salix* L. Species Sampled From Different Locations of Turkey

Subgenus	Section	Code	Species name	Location (Town/Province)
SALIX	**Humboltianae Pax**	Ac-1	***S. acmophylla***	Asma KöprüSuçeken/Batman
	Humboltianae Pax	Ac-2	***S. acmophylla***	Asma KöprüSuçeken/Batman
	Humboltianae Pax	Ac-Fır-7	***S. acmophylla***	Birecik Halfeti/Şanlıurfa
	Humboltianae Pax	Ac-Fır-8	***S. acmophylla***	Birecik Halfeti/Şanlıurfa
	Humboltianae Pax	Ac-Mez-6	***S. acmophylla***	Birecik-Ziyaret Köyü/Şanlıurfa
	Amygdalinae W.Koch	Beh67	***S. triandra* subsp. *triandra***	ÇerkeşOrman Fidanlığı/Çankırı
	Amygdalinae W.Koch	Beh104	***S. triandra* subsp. *triandra***	Üçköy/Çorum
	Amygdalinae W.Koch	Beh136	***S. triandra* subsp. *triandra***	Ardıçlı-Niksar/Tokat
	Amygdalinae W.Koch	Beh129	***S. triandra* subsp. *triandra***	Afyon
	Amygdalinae W.Koch	CorL10/20	***S. triandra* subsp. *triandra***	Uzundere-Tortum/Erzurum
	Amygdalinae W.Koch	Tri1Uz1	***S. triandra* subsp. *triandra***	Uzungöl/Rize
	Amygdalinae W.Koch	Tri2Uz1	***S. triandra* subsp. *triandra***	Uzungöl/Rize
	Amygdalinae W.Koch	Tri2Tos1	***S. triandra* subsp. *triandra***	Tosya-Beşçam/Kastamonu
	Amygdalinae W.Koch	Tri1Tos1	***S. triandra* subsp. *triandra***	Tosya-Beşçam/Kastamonu
	Amygdalinae W.Koch	IHRTri1	***S. triandra* subsp. *triandra***	Ihlara Vadisi
	Amygdalinae W.Koch	IHRTri2	***S. triandra* subsp. *triandra***	Ihlara Vadisi
	Amygdalinae W.Koch	Tri5Cel1	***S. triandra* subsp. *bornmuelleri***	Çeltek-Tersakan/Amasya
	Amygdalinae W.Koch	Tri3Cel1	***S. triandra* subsp. *bornmuelleri***	Çeltek-Tersakan/Amasya
	Pentandrae (Borrer) Schneider	Beh 111	***S. pentandroides***	Topulyurdu/Sivas

	Pentandrae (Borrer) Schneider	Beh102	***S. pentandroides***	Behiçbey Nursery/Ankara
	Pentandrae (Borrer) Schneider	Beh 48a	***S. pentandroides***	Çarşamba/Samsun
	Pentandrae (Borrer) Schneider	Pen2Bag1	***S. pentandroides***	Çoruh-Bağbaşı/Erzurum
	Pentandrae (Borrer) Schneider	Pen1Bag1	***S. pentandroides***	Çoruh-Bağbaşı/Erzurum
	Pentandrae (Borrer) Schneider	HER-Pen*	***S. pentandroides***	Güleman-Ayıpınar/Elazığ
	Pentandrae (Borrer) Schneider	Pen1Ak1	***S. pentandroides***	Ladik/Amasya
	Salix	Beh3	***S. alba***	Akyazı-Vakıf/Sakarya
	Salix	Beh6	***S. alba***	Kızılcahamam-Güvem/Ankara
	Salix	Beh153a	***S. alba***	Konya
	Salix	Kız2-1	***S. alba***	Kızılcahamam/Ankara
	Salix	CorL8-20	***S. alba***	ÇoruhYavuzmah. Köprüsü/Artvin
	Salix	Cor L5-20	***S. alba***	Çoruh/Artvin
	Salix	Cor L3-20	***S. alba***	İspir/ Erzurum
	Salix	Alp8 Bor	***S. alba***	Bor/Niğde
	Salix	Alp6	***S. alba***	Ürgüp/Nevşehir
	Salix	Alp4Ulu	***S. alba***	Uluırmak Köprüsü/ Aksaray
	Salix	Kız1/2	***S. alba***	Kurtboğazıbarajıaltı/Ankara
	Salix	Kny2Y	***S. alba***	Yabani-Ilgın-Çatak/Konya
	Salix	KnyİK	***S. alba***	Yabani-Ilgın-Çatak/Konya
	Salix	Beh57b	***S. excelsa***	Çarşamba-Çelikli/Samsun
	Salix	Beh76	***S. excelsa***	Şenyurt/Artvin
	Salix	C3KIS3	***S. excelsa***	Kışlacık Köyü/Kırklareli
	Salix	C3KIS1	***S. excelsa***	Kışlacık/Kırklareli
	Salix	C1OV1	***S. excelsa***	Ovacık Köyü/Sivas
	Salix	CORL1-24	***S. excelsa***	Derekapı-Yusufeli/Artvin
	Salix	Beh128	***S. fragilis***	Çay/Afyon
	Salix	Beh146	***S. fragilis***	Behiçbey Nursery/Ankara
	Salix	BeyF2	***S. fragilis***	BeynamOrmanı /Ankara
	Salix	Beh160	***S. fragilis***	Akşehir/Konya
	Salix	Beh11	***S. albaxfragilis***	Behiçbey Nursery
	Subalbae	Beh86	***S. babylonica***	Kalecik/Ankara
	Subalbae	Beh 95	***S. babylonica***	YaylacıkKöyü/Amasya
	Subalbae	Beh 108	***S. babylonica***	Tokat
	Subalbae	C4Yay2	***S. babylonica***	ÇoruhYaylacıkÇıkışı/ Artvin
	Subalbae	Cor L4/22	***S. babylonica***	Yusufeli-Çoruh/Artvin
	Subalbae	Cor L5/22	***S. babylonica***	Behiçbey Nursery/Ankara
	Subalbae	IHRBAB	***S. babylonica***	IhlaraValley/Aksaray
VETRIX	**Hastatae Kerner**	Ap1Lad2	***S. apoda***	Ladik/Amasya
	Hastatae Kerner	Apo-H*	***S. apoda***	Çamlıhemşin Şenköy çevresi/Artvin
	Nigricantes	MY1*	***S. myrsinifolia***	Ilgaz/Kastamonu

Nigricantes	MY2*	***S. myrsinifolia***	Ilgaz/Kastamonu
Vetrix Dumort.	ÇorL1/21	***S. caucasica***	Yokuşlu köyü/Artvin
Vetrix Dumort.	Riz2Sira1	***S. caucasica***	Çoruh /Artvin
Vetrix Dumort.	RizCinSira 1	***S. caucasica***	Çoruh-Sırakonaklar/Artvin
Vetrix Dumort.	Riz1Sira1	***S. caucasica***	Çoruh-Sırakonaklar/Artvin
Vetrix Dumort.	Riz2Ay1	***S. caucasica***	Ayder/Rize
Vetrix Dumort.	Riz1Ay1	***S. caucasica***	Ayder/ Rize
Vetrix Dumort.	GKS pedi	***S. pedicellata* subsp. *pedicellata***	Göksu-Ermenek/Karaman
Vetrix Dumort.	Pedi-H*	***S. pedicellata* subsp. *pedicellata***	Maraş
Vetrix Dumort.	Kız 4/2	***S. caprea***	Kızılcahamam Soğuksu Milli Parkı/Ankara
Vetrix Dumort.	CA1KAS1	***S. caprea***	Kastamonu-Çankırı il sınırı
Vetrix Dumort.	CA2KAS1	***S. caprea***	Kastamonu-Çankırı il sınırı
Vetrix Dumort.	CA1BOS1	***S. caprea***	Bostan/Kastamonu
Vetrix Dumort.	Art Kaf 1	***S. caprea***	Kafkasör Yaylası/Artvin
Vetrix Dumort.	Art Kaf 2	***S. caprea***	Kafkasör Yaylası/ Artvin
Vetrix Dumort.	Art Kaf 3	***S. caprea***	Kafkasör Yaylası/Artvin
Vetrix Dumort.	AE-H*	***S. aegyptiaca***	Bahçesaray-yol kenarı/Van
Vetrix Dumort.	AE-ANK*	***S. aegyptiaca***	Tatvan/Bitlis
Vetrix Dumort.	E.A.1	***S. aegyptiaca***	Kars-Erzurum Yolu
Vetrix Dumort.	Beh43A	***S. cinerea***	Akyazı Gebeş/Sakarya
Vetrix Dumort.	C5Kar6	***S. cinerea***	Çubuk-Karagöl/ Ankara
Vetrix Dumort.	C5Kar5	***S. cinerea***	Çubuk-Karagöl/ Ankara
Vetrix Dumort.	Cin1Bag1	***S. cinerea***	Çoruh-Bağbaşı/Erzurum
Vetrix Dumort.	Cin1Bag3	***S. cinerea***	Çoruh-Bağbaşı/Erzurum
Vetrix Dumort.	Cin2Bag2	***S. cinerea***	Çoruh-Bağbaşı/Erzurum
Vetrix Dumort.	Cin1Uz1	***S. cinerea***	Çoruh-Bağbaşı/Erzurum
Vetrix Dumort.	Beh 126	***S. pseudomedemii***	Zile/Tokat
Vetrix Dumort.	Beh 130b	***S. pseudomedemii***	Behiçbey Nursery/Ankara
Vetrix Dumort.	Bey P4	***S. pseudomedemii***	Beynam Forest/Ankara
Vetrix Dumort.	Bey P2	***S. pseudomedemii***	Beynam Forest/Ankara
Vetrix Dumort.	Bey P1	***S. pseudomedemii***	Beynam Forest/Ankara
Vetrix Dumort.	PSE-ANK*	***S. pseudodepressa***	Sarıkamış/ Kars
Vetrix Dumort.	PSE-H*	***S. pseudodepressa***	GümüşDamla Köyü/Bayburt
Vimen Dumort.	E.V.1	***S. viminalis***	Nehir Başı/Erzurum

Vimen Dumort.	E.V.2	***S. viminalis***	Erzurum
Vimen Dumort.	Arm7Olt3	***S. armenorossica***	Oltu/ Erzurum
Vimen Dumort.	Arm2Olt3	***S. armenorossica***	Oltu/ Erzurum
Vimen Dumort.	Arm1Bag2	***S. armenorossica***	Bağbaşı-Çoruh/Erzurum
Vimen Dumort.	Arm2Bag3	***S. armenorossica***	Bağbaşı-Çoruh/Erzurum
Vimen Dumort.	Arm1Bag3	***S. armenorossica***	Bağbaşı-Çoruh/Erzurum
Canae Kerner	Myr0Ilg1	***S. elaeagnos***	Ilgaz/Kastamonu
Canae Kerner	El2Tos1	***S. elaeagnos***	Tosya-Beşçam/Kastamonu
Canae Kerner	El2Bos1	***S. elaeagnos***	Bostan/ Kastamonu
Canae Kerner	KsEla1	***S. elaeagnos***	Cide/Kastamonu
Canae Kerner	KsEla2	***S .elaeagnos***	Ağlı/Kastamnou
Helix Dumort.	Elb4Alb3	***S. elbursensis***	Çoruh-Alanbaşı/Artvin
Helix Dumort.	Elb1Alb1	***S. elbursensis***	Çoruh-Alanbaşı/Artvin
Helix Dumort.	Her-El*	***S. elbursensis***	Çoruh-Alanbaşı/Artvin
Helix Dumort.	Elb5Alb1	***S. elbursensis***	Çoruh-Alanbaşı/Artvin
Helix Dumort.	Kız3-1	***S. amplexicaulis***	Kızılcahamam-SoğuksuMilliParkı/ Ankara
Helix Dumort.	C5Kar11	***S. amplexicaulis***	Çubuk-Karagöl/ Ankara
Helix Dumort.	C5Kar10	***S. amplexicaulis***	Çubuk-Karagöl/ Ankara
Helix Dumort.	Kız3/2	***S. amplexicaulis***	Kızılcahamam-Soğuksu Milli Parkı/ Ankara
Helix Dumort.	Amp2Kuz 1	***S. amplexicaulis***	Ilgaz/Kastamonu
Helix Dumort.	Amp1Dag 1	***S. amplexicaulis***	Ilgaz/Kastamonu
Helix Dumort.	Beh10a	***S. amplexicaulis***	Ilgaz/Kastamonu
Helix Dumort.	Her-Riz*	***S. rizeensis***	İkizdere/Rize
Helix Dumort.	Riz-H*	***S. rizeensis***	İkizdere/Rize
Helix Dumort.	Pur1Dan2	***S. purpurea* subsp. *leucodermis***	Köyceğiz-Dana deresi /Muğla
Helix Dumort.	Pur1Gok3	***S. purpurea* subsp. *leucodermis***	Gökçeova Gölü/ Muğla
Helix Dumort.	Pur2Ag1	***S. purpurea* subsp. *leucodermis***	Köyceğiz-Ağla/Muğla
Helix Dumort.	Pur8Dan2	***S. purpurea* subsp. *leucodermis***	Köyceğiz-Dana deresi /Muğla
Helix Dumort.	Pur5 Dan2	***S. purpurea* subsp. *leucodermis***	Köyceğiz-Dana deresi /Muğla
Helix Dumort.	Pur2 Gök 3	***S. purpurea* subsp. *leucodermis***	Köyceğiz-Dana deresi /Muğla
Cheilophilae Hao	Her-Wil*	***S. wilhelmsiana***	İkizdere/Rize
Cheilophilae Hao	Wil-H*	***S. wilhelmsiana***	Tortum/Erzurum
Cheilophilae Hao	E.W.1	***S. wilhelmsiana***	Kars Erzurum Road

* Species samples from herbarium

Appendix B: Agarose Gel Images For Every Studied Region

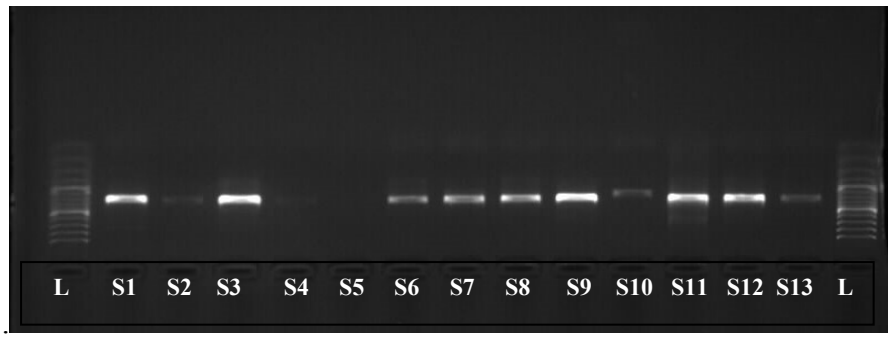

Figure B.1 PCR products of *trn*T-L IGS region run in 1.5% agarose gel (L: Gene ruler, 100-3000bp)

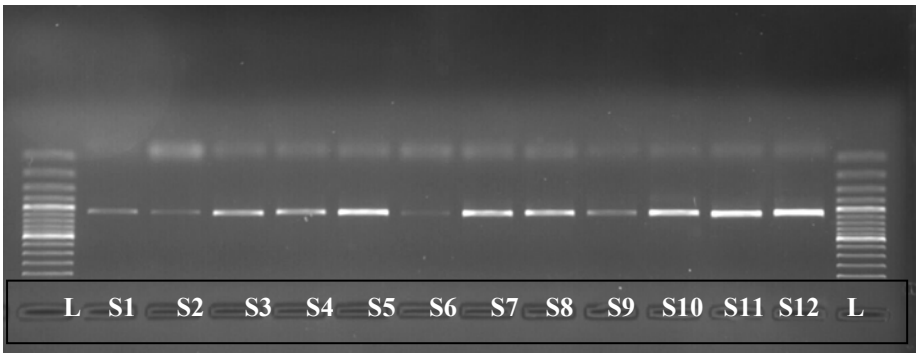

Figure B.2 PCR products of *trn*L intron region run in 1.5% agarose gel (L: Gene ruler, 100-3000bp)

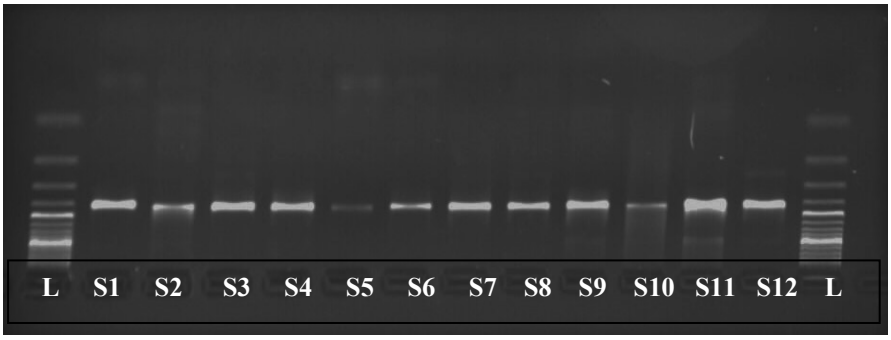

Figure B.3 PCR products of *trn*L-F IGS region run in 1.5% agarose gel (L: Gene ruler, 100-3000bp)

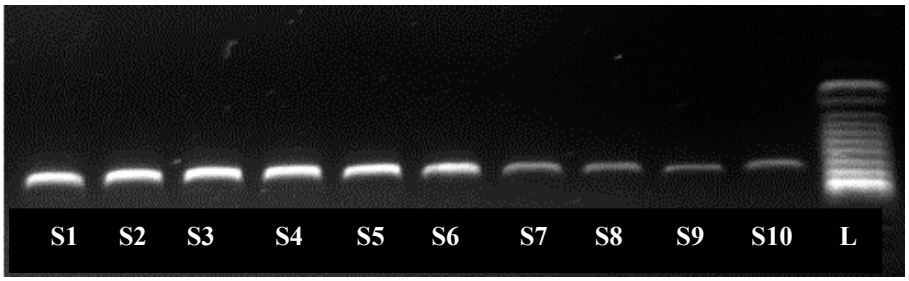

Figure B.4 PCR products of *mat*K-1 region run in 1.5% agarose gel(L: Gene ruler, 100-3000bp)

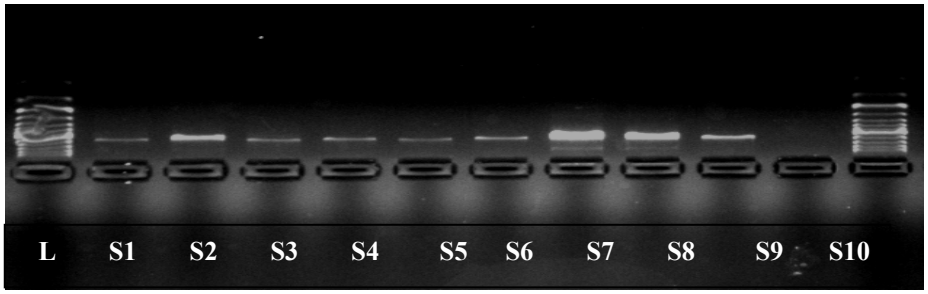

Figure B.5 PCR products of *mat*K-2 region run in 1.5% agarose gel (L: Gene ruler, 100-3000bp)

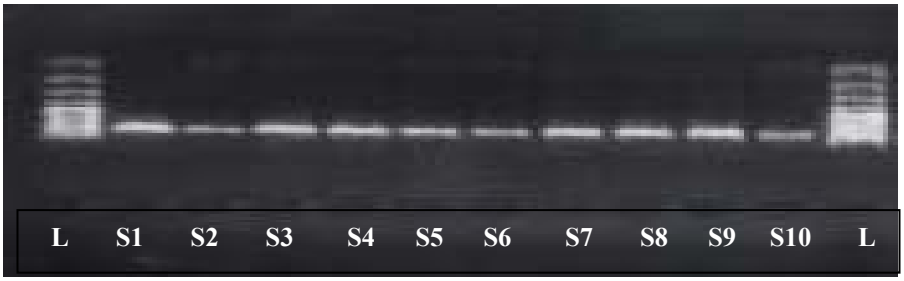

Figure B.6 PCR products of *rbc*L-1 region run in 1.5% agarose gel (L: Gene ruler, 100-3000bp)

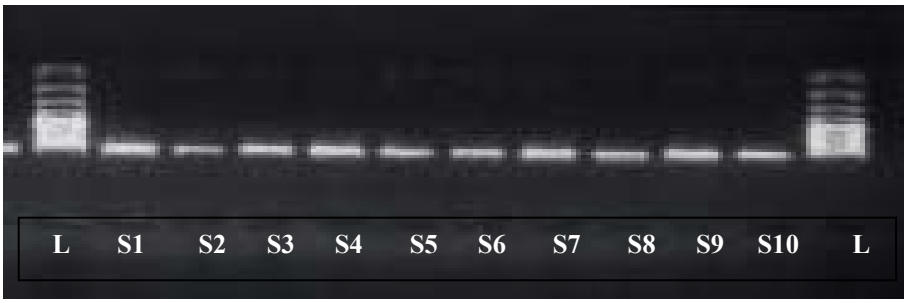

Figure B.7 PCR products of *rbc*L-2 region run in 1.5% agarose gel (L: Gene ruler, 100-3000bp)

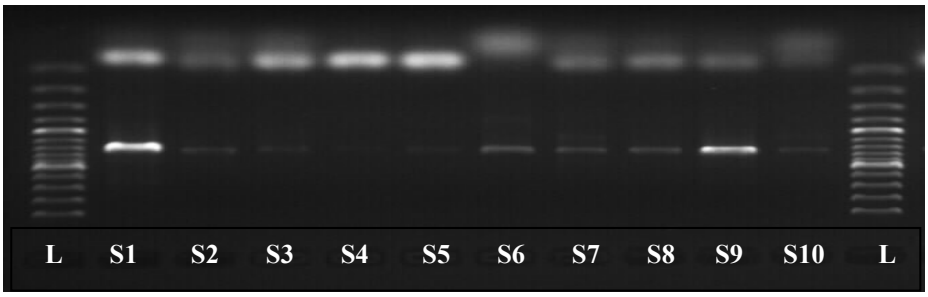

Figure B.8 PCR products of *ITS* L-4 region run in 1.5% agarose gel (L: Gene ruler, 100-3000bp)

Appendix C: Chromatogram Structures for Each Studied Regions

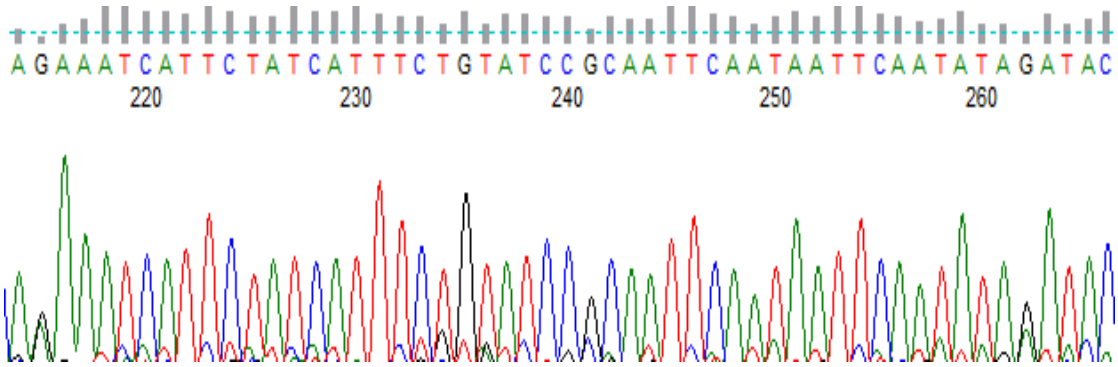

Figure C.9 A chromatogram example for *trn*T-F region

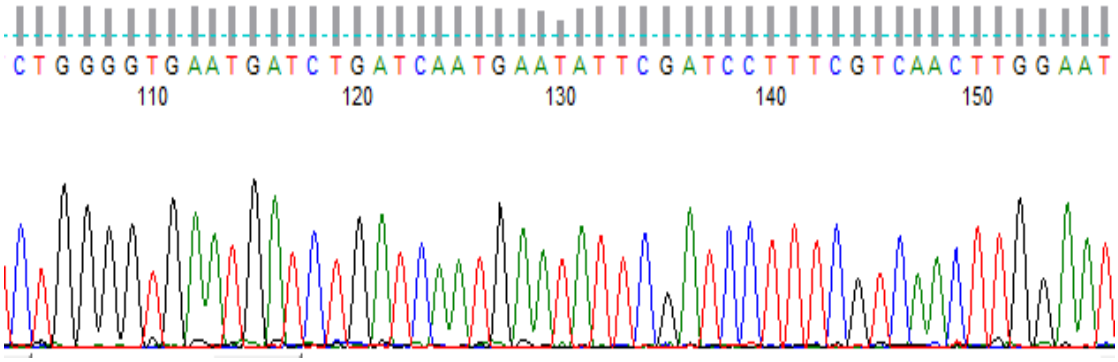

Figure C.10 A chromatogram example for *trn*L region

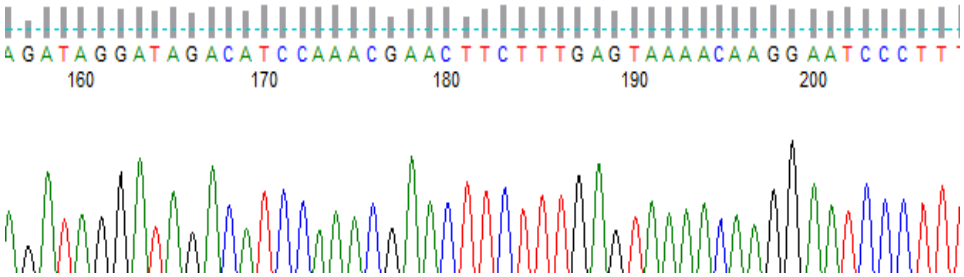

Figure C.11 A chromatogram example for *trn*L-F region

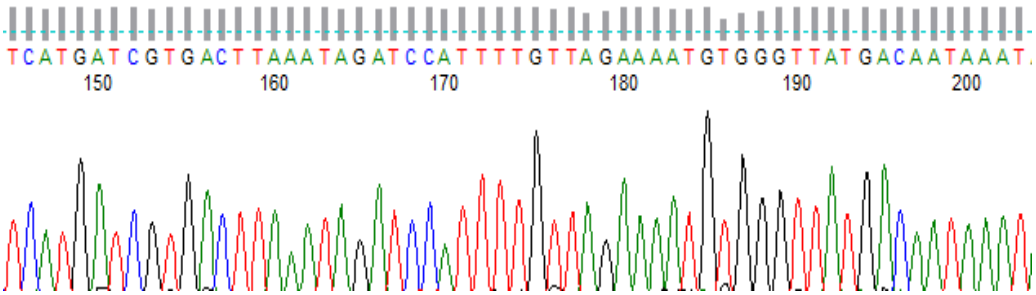

Figure C.12 A chromatogram example for *mat*K-1 region

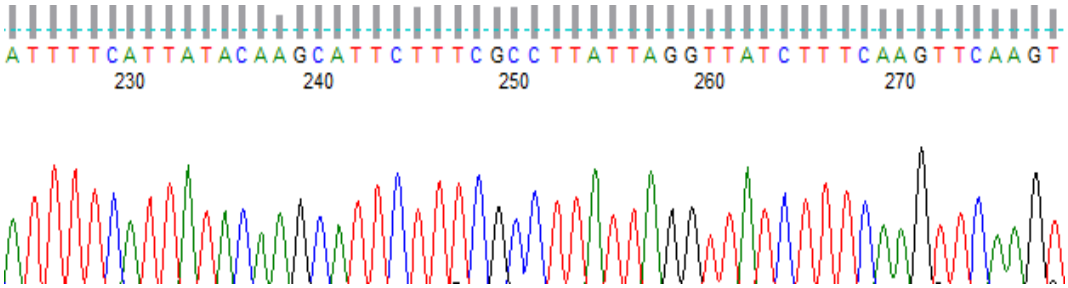

Figure C.13 A chromatogram example for *mat*K-2 region

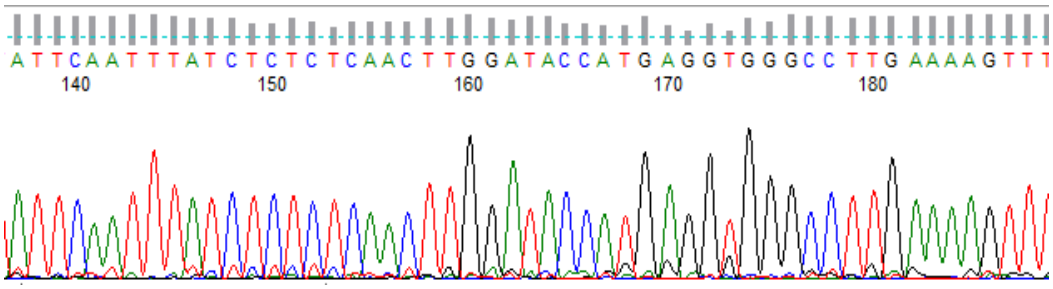

Figure C.14 A chromatogram example for *rbc*L-1 region

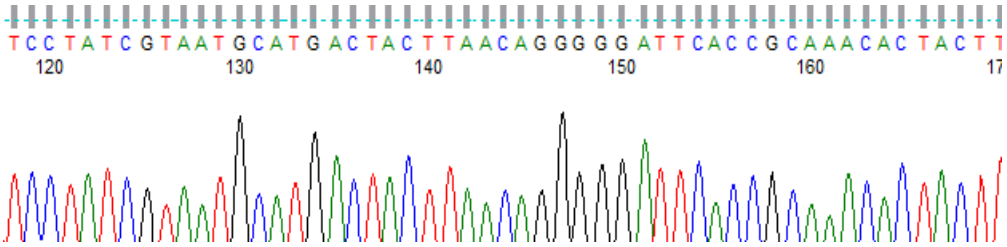

Figure C.15 A chromatogram example for *rbc*L-2 region

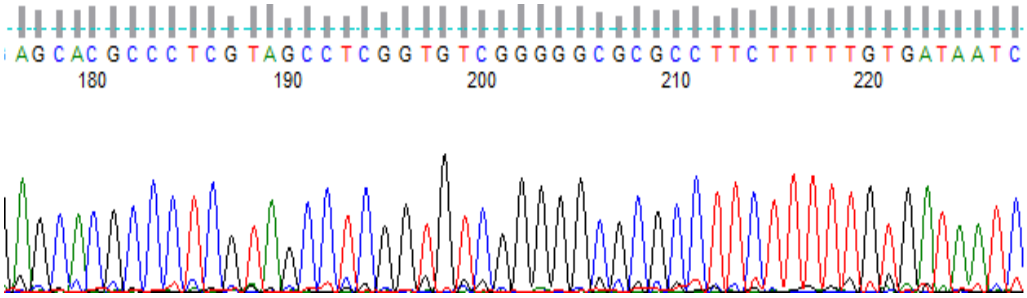

Figure C.16 A chromatogram example for *ITS* L-4 region

CURRICULUM VITAE

General Information

Date of Issue	:
ID Number	: 17707455286
Name Surname	: Pelin Acar
Mail Address	: METU, Department of Biology, Lab 251, Ankara
Birth Date and Place	: 19.02.1984 Antalya
Phone	: 0 312 210 51 60 GSM: 0 533 721 60 15
e-mail	: pkeske@metu.edu.tr

Education

Period of Study	Degree	University	Department
2010/-	Doctorate	METU	Molecular Biology
2006/2009	Master	Ankara University	Ecology and Plant Sciences
2002/2006	Bachelor	Ankara University	Biology

Academic and Professional Experience

Period	Position	Location	Department
2008 (3 months)	Visiting Scholar	Blanes Botanical Garden/Spain	-
2010/-2017	Research Assistant	METU	Biology