The Assessment of Genetic Diversity and Population Structure of Endemic Scutellaria yildirimlii by ISSR Marker for Conservation Purposes

Merve YILDIRIM (merveornek@hotmail.com)
Ankara University
Gül Nilhan TUĞ
Ankara University
Ahmet Emre YAPRAK
Ankara University

Research Article

Keywords: Scutellaria yildirimlii, Endemic, Conservation, Population genetics, ISSR

Posted Date: September 12th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2038976/v1

Abstract

Scutellaria yildirimlii is an endemic species from Turkey and belongs to the family Lamiaceae which has a high number of medicinal and aromatic plants. It is crucial to determine the genetic potential of these rare and endemic species for conservation; therefore, 111 samples from 5 populations of *S. yildirimlii* were investigated with 15 Inter Simple Sequence Repeats (ISSR) primers. The percentage of polymorphic loci (PPL) of *S. yildirimlii* was determined as 93.9% at the species level and 56.5% at the average population level. The Shannon's information index (I) and Nei's gene diversity (H) were calculated as 0.292 and 0.183, respectively. Gene flow among populations (Nm) was obtained as 2.984. Also, the variation between and within the population was 16% and 84%, respectively. The distribution of populations concerning gene flow shows a correlation between Unweighted Pair-Group Method with Arithmetic Averages (UPGMA) dendrogram and Principal Coordinate Analysis (PCoA) analysis. Structure analysis (ΔK=3) revealed that K test yielded maximum peaks for 3 groups. In addition, the population distribution resembled both the UPGMA and PCoA dendrograms.

1. Introduction

Lamiaceae is one of the most prominent angiosperm families, with 236 genera and more than 7000 species (Harley et al. 2004). *Scutellaria* L., also known as skullcap, is one of the largest genera of the family, with 470 species (WCSP 2022). The principal diversity centers of *Scutellaria* are mostly Central Asia, Afghanistan, and the Irano-Turanian phytogeographical region, while the East Mediterranean region and the Andes are considered secondary diversity hotspots (Paton 1989, 1990). Turkey is rich in *Scutellaria* species diversity since the members of this species are distributed all over the country. However, some of these species are at high risk of extinction because of their partial and small populations (Minareci and Pekönür 2017). Turkey is home to 39 *Scutellaria* species, of which 17 are endemic (endemism ratio of 43.58%) (Güner et al. 2012). Although the extinction of species is considered a part of evolutionary processes related to ecological factors, today, the major reasons for the loss of biodiversity are human activities, climate changes and habitat loss due to stochastic incidents (Primack 2006). Unfortunately, extinction processes occur more rapidly than new species emerge (Frankham et al. 2002). Genetic diversity plays a crucial role in species' long-term survival and adaptability. Obtaining data about the genetic structures of the species with multiple and separate populations is vital in case the species consist of separate populations because knowledge about genetic diversity within and between the populations has become a prerequisite for the establishment of a strategic plan for conservation program (Crema et al. 2009). Molecular markers are commonly used to determine plant populations' genetic structure (Weber and May 1989; Liu 1998). ISSR (Inter Simple Sequence Repeat), which is a PCR (Polymerase Chain Reaction) based molecular marker, primers with double, triple, quadruple and quintuple repetitive nucleotides are used, and the region between two microsatellites can be amplified with these primers (Zietkiewicz et al. 1994). Because of their highly polymorphic nature, this technique aids in assessing variation among species and detecting population structure (Boydak et al. 2021). ISSR markers have some advantages like primer design without knowledge of genomic sequence, the

requirement of a small amount of DNA and low costs compared with RFLP (Restriction Fragment Length Polymorphism), SSR (Simple Sequence Repeat), and AFLP (Amplified Fragment Length Polymorphism) techniques. Also, it produces more reliable and reproducible bands due to its higher binding temperature and extended primers compared to RAPD (*Random Amplification of Polymorphic DNA) (Qian et al. 2001)*. On the other hand, it has disadvantages such as the inability to distinguish heterozygosity in loci because it is a dominant marker, the determination of binding temperatures of the primers separately, and the possibility that particles of similar sizes are not homologous (Kesawat and Das Kumar 2009). Given that it was more appropriate for the research topic and the setup of our research facility, ISSR was chosen for this investigation. Under the scope of biodiversity conservation, it is crucial to draw attention to the future of endangered species for us and the next generations. *Scutellaria yildirimlii* is an edaphic endemic species, first introduced in 2013 from the marly-gypsaceous soils of Central Anatolia. It has a narrow distribution area with fragmented populations (Cicek and Yaprak 2013). According to IUCN (2012) Red List Criteria, it is included in EN (Endangered) category (Yıldırım et al. 2019). In this study, it was aimed to (i) determine genetic diversity level of natural populations of *S. yildirimlii* using ISSR fingerprinting technique, (ii) evaluate the genetic differentiation level between populations, and (iii) develop strategies for conserving the species in light of the study's findings.

2. Materials And Methods

2.1. Plant material

Literature and herbaria data (ANK, HUB, GAZI and ESSE) were utilised to accurately identify the reported distribution areas of *S. yildirimlii* and revealed that the species inhabits only the provinces of Ankara and Eskişehir at six different localities that were all studied in 2016 (Figure 1). Due to the negligible distance between them, the subpopulations from Asagikepen (~28 ha) and Yeşilkoy (~0,5 ha) were considered to be a single population (Yıldırım et al. 2019). Based on the population sizes of these localities, a total of 22 samples, 17 from Asagikepen and five from Yesilkoy, were analysed. In total, fresh leaf samples of 111 different individuals were collected from all indicated populations of the species (Table 1). Throughout the sampling process, the distances between sampled individuals were defined considering the size of each population and the individuals' distribution patterns. The fresh leaf samples were labelled, dried in silica gel-filled bags, and stored at –80 °C until DNA extraction.

Table 1

The geographical locations of *S. yildirimlii*

Population	Location	Latitude (N) Longitude (E) Altitude (m)	Sample size (individual code)
Ayas	A4 16th km of Ayas-Ankara route, Aysanti pass, gypseous soils	40° 05' 37.8" 032° 26' 17.1" 1205 m	23 (1-23)
Kizlarkayasi	B4 Ankara, Polatli-Kizlarkayasi	39° 40' 35.9" 032° 00' 12" 755 m	21 (24-44)
Oglakci	B3 24th km of Sivrihisar-Polatli route, near Oglakci, Eskisehir, clayish hillsides	39° 33' 11.9" 031° 43' 35.9" 849 m	22 (45-66)
Asagikepen	B3 At the intersection of Yesilkoy road and Afyon-Sivrihisar route, near Asagıkepen village Eskisehir, 872 m, gypseoushillsides	39° 19' 44.8" 031° 27' 27.7" 877 m	5 (67-71)
	B3 Near Asagikepen village in Sivrihisar, Eskisehir, gypseous soils.	39° 22' 15.2" 031° 29' 17.4" 960 m	17 (72-88)
Kavuncu	A3 North-East of Kavuncu village in Gunyuzu, Eskisehir, gypseous step	39° 24' 39.9" 031° 54' 15.1" 742 m	23 (89-111)
Total			**111**

2.2. DNA extraction and ISSR-PCR amplification

Genomic DNA extraction was realized using the “Macherey-Nagel NucleoSpin® Plant II” kit with a denoted protocol. With the help of a Nanodrop spectrophotometer (Thermo Scientific Nanodrop One-W), the purities and concentrations of isolated DNA samples were determined. Afterwards, the samples were diluted to 10 ng/µl and stored at –20 ^{o}C for PCR applications. 59 universal UBC-ISSR (University of British Columbia, Canada) primers were screened at PCR amplification using “Applied Biosystems Veriti 96 Well Thermal Cycler”. 15 primers with the highest polymorphism, repeatability and clearest bands were used among the scanned primers. PCR-ISSR amplification was performed for each primer in two parts due to the exceeding number of samples. In the first part, PCR-ISSR amplification of 88 samples from Ayas, Kizlarkayasi, Oglakci and Asagikepen populations; in the second part, 23 samples from Kavuncu population with the repetition of 8 samples from Asagikepen population were performed. ISSR amplifications were realized with a reaction volume of 20 µl containing 1XTaq DNA polymerase buffer (with $MgCl_2$), 200 µM dNTPs, 0.4 µM primer, 1 unit Taq DNA polymerase enzyme and 10 ng genomic DNA applying Touchdown PCR (TD) program (Don et al. 1991). Amplification reaction includes the following steps; initial denaturation step for 5 min at 95 ^{o}C followed by 1 cycle; denaturation 30 s at 95 ^{o}C, annealing 45 s at 65-55 ^{o}C, extension 1 min 30 s at 72 ^{o}C, followed by 15 cycles; denaturation 30 s at 95 ^{o}C, annealing 30 s at 45 ^{o}C, extension 1 min 30 s at 72 ^{o}C, followed by 20 cycles; final extension 7 min at 72 ^{o}C followed by 1 cycle. After combining with a 6X loading dye, the amplified DNA fragments were separated by 2% (w/v) agarose gel electrophoresis containing 0.5X TBE (Tris-Borate-EDTA) buffer at a specific voltage of 90 V for 5 hours, and ethidium bromide (EB) was used to stain the gel. Band sizes were determined using "abm DNA ladder" molecular weight marker ranging from 100 pb – 3 kb. The gels were visualized and photographed under UV via BioRad Molecular Imager DocXR+.

2.3. Data analysis

To create a binary data matrix, clear and reproducibly amplified ISSR bands were scored as either present (1) or absent (0). All gathered data was combined and utilised to analyse genetic diversity parameters such as observed number of alleles (*Na*), effective number of alleles (*Ne*) (Kimura and Crow 1964), Nei’s gene diversity (*H*) (Nei 1973), Shannon’s information index (*I*) (Lewontin 1972), the number and percentage of polymorphic loci (*NPL* and *PPL*), total genetic diversity (H_T), genetic diversity within a population (H_S), and coefficient of genetic differentiation among populations ($G_{ST}=1-H_S/H_T$) with POPGENE 1.32 (Yeh et al. 1997). This software was also performed to calculate gene flow between populations (*Nm*) based on G_{ST} value using the equation $Nm=0.5(1-G_{ST})/G_{ST}$ (McDermott and McDonald 1993). The UPGMA (Unweighted Pair-Group Method with Arithmetic Averages) dendrogram, which illustrates genetic relationships between populations, was generated using the MEGA (Molecular Evolutionary Genetics Analysis) 6.06 software based on the matrix of Nei's genetic distance (Nei 1978; Tamura et al. 2013). Additional UPGMA dendrogram depicting genetic distances between individuals was

constructed using SYN-TAX 2000 software with Nei's genetic distance data on the basis of Jaccard similarity coefficient (Jaccard 1908; Podani 2001). Using a hierarchical analysis of molecular variance (AMOVA) software (Excoffier et al. 1992) generated by the GenAlEx 6.5 programme, pairwise population comparisons were utilised to explain the partition of genetic diversity within and among populations. AMOVA also computed ΦPT statistics, the analogue for binary data to F_{ST} statistics (Fixation index). The Principal Coordinate Analysis (PCoA) was performed using GenAlEx 6.5 software to provide the spatial presentation of the relative genetic distances between individuals and populations (Peakall and Smouse 2012). The correlation between geographic and genetic distance matrices was calculated at a population level using the Mantel test (Mantel 1967). The genetic structure of *S. yildirimlii* populations was performed using an admixture model with a Bayesian clustering algorithm implemented in the STRUCTURE software v 2.3.4 (Pritchard et al. 2000). For each K (subpopulation) value ranging from 2 to 14, twenty iterations were realized with a burning of $5x10^4$ interactions, followed by $3x10^5$ Monte Carlo Markov Chain (MCMC) (Wu et al. 2015). The Evanno method was used to calculate ΔK, the most likely number of subpopulations, with STRUCTURE HARVESTER 2.3.4 software (Evanno et al. 2005; Earl and VonHoldt 2012). The model with the highest K value was deemed to describe the data best. Populations were assigned to clusters using a threshold value of ≥ 0.80. Individuals that did not match this criterion were deemed admixed (Celik et al. 2016).

3. Results

Analysis was conducted on 111 samples of *S. yildirimlii* from 5 populations. Among the 59 ISSR markers analyzed, the 15 primers with the highest polymorphism results were chosen (Table 2). 311 of 331 bands generated by these primers were polymorphic (93.9%). The primers with the highest polymorphism rates, UBC 807 and 812, had a value of 100%, while the primer with the lowest polymorphism rate, UBC 818, had a value of 81.2%. The primer band range was 16 to 26. The average number of bands per primer was 22.1, with a total of 20.8 polymorphic bands. The length of the amplified bands ranged from 200 to 2000 bp.

Table 2

Polymorphism analysis of *S. yildirimlii* from ISSR primers

	Marker	Sequence (5'-3')	T_m(°C)	TNB	NPB	PPB%	Band size (bp)
1	UBC 807	$(AG)_8T$	TD-56	26	26	100	200-2000
2	UBC 808	$(AG)_8C$	TD-56	26	25	96.1	500-2000
3	UBC 810	$(GA)_8T$	TD-56	17	15	88.2	500-2000
4	UBC 812	$(GA)_8A$	TD-56	20	20	100	300-2000
5	UBC 818	$(CA)_8G$	TD-56	16	13	81.2	200-2000
6	UBC 826	$(AC)_8C$	TD-56	25	24	96.0	400-2000
7	UBC 835	$(AG)_8YC$	TD-56	25	24	96.0	300-2000
8	UBC 836	$(AG)_8YA$	TD-56	26	25	96.1	200-2000
9	UBC 840	$(GA)_8YT$	TD-56	26	25	96.1	200-2000
10	UBC 842	$(GA)_8YG$	TD-56	19	18	94.7	300-2000
11	UBC 856	$(AC)_8YA$	TD-56	24	23	95.8	300-2000
12	UBC 888	$BDB(CA)_7$	TD-65	25	23	92.0	300-2000
13	UBC 889	DBD(AC)7	TD-65	18	15	83.3	600-2000
14	UBC 890	VHV(GT)7	TD-56	17	15	88.2	400-2000
15	UBC 891	HVH(TG)7	TD-56	21	20	95.2	500-2000
			Total	**331**	**311**	**93.9**	**200-2000**

T_m: annealing temperature, **TNB:** total number of bands, **NPB:** number of polymorphic bands, **PPB:** percentage of polymorphic bands, **Y**=(C,T), **B**=(C,G,T), **D**=(A,G,T), **V**=(A,C,G), **H**=(A,C,T)

Both population and species-level genetic diversity parameters for *S. yildirimlii* were analysed. At the population level, 187 polymorphic loci (NPL) were detected with a percentage of 56.5% (PPL), the mean number of effective alleles (Ne) was 1.263, while the mean number of observed alleles (Na) was 1.565. For Nei's gene diversity (H) and Shannon's information index (I), mean values of overall populations were computed as 0.158 and 0.242, respectively. In addition, the following species-level values were obtained; polymorphic loci number (NPL) was found as 311 with a percentage of 93.9% (PPL), and the numbers of observed alleles (Na) and effective alleles (Ne) were revealed to be 1.934 and 1.296, respectively. Species-level value of Nei's gene diversity (H) was 0.183, while Shannon's information index (I) was 0.292 (Table 3).

Table 3

Genetic analysis of *S. yildirimlii* at population and species levels

Pop. Name	N	N_a±S	N_e±S	H±S	I±S	NPL	PPL(%)
Ayas	23	1.547 ± 0.499	1.262 ± 0.354	0.155 ± 0.190	0.237 ± 0.271	181	54.7
Kizlarkayasi	21	1.538 ± 0.499	1.252 ± 0.339	0.152 ± 0.185	0.234 ± 0.266	178	53.8
Oglakci	22	1.613 ± 0.488	1.280 ± 0.344	0.169 ± 0.187	0.262 ± 0.266	203	61.3
Asagikepen	22	1.601 ± 0.490	1.275 ± 0.353	0.165 ± 0.189	0.254 ± 0.268	199	60.1
Kavuncu	23	1.526 ± 0.501	1.244 ± 0.346	0.145 ± 0.186	0.223 ± 0.266	174	52.6
Average		1.565 ± 0.496	1.263 ± 0.348	0.158 ± 0.188	0.242 ± 0.268	187	56.5
Species level	111	1.934 ± 0.250	1.296 ± 0.336	0.183 ± 0.178	0.292 ± 0.246	311	93.9

N: number of individuals, **N_a:** number of observed alleles, **N_e:** number of effective alleles, **H:** Nei's (1973) genediversity, **I:** Shannon's information index (Lewontin, 1972), **NPL:** number of polymorphic loci, **PPL:** percentage of polymorphic loci, **S:** Standarddeviation

Based on all loci data, the values of Nei's gene diversity parameters (Nei 1987) were obtained as follows; total genetic diversity (H_T) was 0.183, while genetic diversity within the population (H_S) was 0.157. Using the $G_{ST}=1-H_S/H_T$ formula, the value of the coefficient of genetic differentiation among populations (G_{ST}) was calculated to be 0.144, and gene flow (Nm) between populations was estimated as 2.984 (Table 4). Nm values can be grouped into three categories as follows: Nm ≥1 is considered a high, 0.250<Nm <0.990 is considered an intermediate and 0 <Nm<0.249 is considered a low gene flow (Slatkin 1981, 1985; Caccone 1985; Waples 1987). Relatedly, Wright suggested that generally, if Nm<1, local differentiation of populations will result, whereas if Nm>1, population differentiation will be minimal among them (Wright 1949).

Table 4

Nei's analysis of gene diversity for all loci (Nei 1987)

All loci	N	H_T ±S	H_S±S	G_{ST}	N_m
Average	111	0.183 ± 0.031	0.157 ± 0.024	0.144	2.984

N: number of individuals, **H_T:** total genetic diversity, **H_S:**genetic diversity within population**, G_{ST}:** coefficient of genetic differentiation; $G_{ST}=1-H_S/H_T$, **N_m:** estimate of gene flow; $N_m=0.5(1-G_{ST})/G_{ST}$, **S:** Standard deviation

The AMOVA test revealed that 84% of the total genetic variation occurred within populations, and only 16% was noted among populations which were statistically significant ($p<0.001$) with the permutations based on 999 iterations. Nm result (2.984) also indicated a high level of gene flow and low genetic differentiation among populations and, therefore, corroborates a low value for ΦPT (PhiPT)=0.159 (Table5)

Table 5

Statistics of molecular variance analysis (AMOVA) of *S. yildirimlii*

Source	df	SS	MS	VC	V%	Φ-statistics	P-value
Among populations	4	563.967	140.992	5.128	16%	Φ_{PT} 0.159	0.001
Within populations	106	2881.655	27.185	27.185	84%		0.001
Total	110	3445.622	-	32.313	100%		

df: degrees of freedom, **SS**: sum of square, **MS**: mean square, **VC**: variance components, **V%**: percentage of variance, **Φ_{PT}**(PhiPT): genetic differentiation among populations, **P:** statistical significance

Nei's unbiased genetic distance (Nei 1978) between populations ranged from 0.0219 to 0.0512. Samples from Ayas were the most distantly related to samples of Asagikepen, with a correlation between geographic distance (113 km) and genetic distance (0.0512) among populations along with their identity as 0.9501, the lowest of all the groups studied. In addition to the geographic distance (29 km), the genetic distance between the other pairwise combinations of populations was relatively low, with the least genetic difference between samples of Asagikepen and Oglakci (0.0219). These two groups have the highest genetic identity with a rating of 0.9784 (Table 6). Statistically significant p-value ($p<0.001$) supported that there was a moderately positive correlation between genetic and geographic distance as a result of Mantel test (Rxy=0.493) concerning Nei's identity data (Nei 1978).

Table 6

Nei's unbiased measures of genetic identity (above diagonal) and genetic distance (below diagonal) (Nei 1978)

Population	Ayas	Kizlarkayasi	Oglakci	Asagikepen	Kavuncu
Ayas	****	0.9748	0.9588	0.9501	0.9554
Kizlarkayasi	0.0255 (60 km)	****	0.9715	0.9644	0.9733
Oglakci	0.0421 (84 km)	0.0290 (27 km)	****	0.9784	0.9630
Asagikepen	0.0512 (113 km)	0.0362 (55 km)	0.0219 (29 km)	****	0.9621
Kavuncu	0.0457 (88 km)	0.0270 (31 km)	0.0377 (22 km)	0.0387 (36 km)	****

Geographic distances (km) between population pairs were also indicated below diagonal in parentheses

Using Nei's binary genetic distance matrix (Nei 1978), UPGMA and PCoA cluster analyses were used to study the relationship between *S. yildirimlii* populations (Figures 2 & 3). The dendrogram generated by the UPGMA algorithm was separated into two main clusters, the first cluster of which includes Ayas and Kizlarkayasi populations in Ankara province, as well as Kavuncu population in Eskişehir. The second major cluster consisted of Oglakci and Asagikepen with Yesilkoy populations from Eskişehir province (Figure 2).

According to the PCoA analysis results, the first two primary components described 46.99% and 28.03% of the total variation, respectively, while the third components explained 90.44%. Along the first axis of the PCoA graph, the populations clustered into two major groups: Ayas, Kizlarkayasi, Kavuncu populations, and Oglakci, Asagikepen (with Yesilkoy) populations (Figure 3). The PCoA graph and UPGMA dendrogram complement one another.

The genetic relationship among 111 individuals of the species was established based on the Jaccard similarity coefficient by UPGMA clustering analysis using SYN-TAX 2000 program (Podani 2001). The dendrogram generated by the UPGMA algorithm was separated into two main clusters, the first cluster of which included Ayas and Kizlarkayasi populations in Ankara province, as well as Kavuncu population with 2 individuals from Asagikepen (87 and 88) and 3 individuals from Oglakci population (46, 47 and 54) in Eskişehir province. The second major cluster consisted of Oglakci and Asagikepen with Yesilkoy populations from Eskişehir province (Figure 4). Another genetic relationship among individuals was created based on a distance matrix by PCoA analysis using GenAlEx 6.5 program (Peakall and Smouse 2012).

The first two primary components described 7.73% and 5.46% of the total variation, respectively, while the third component explained 16.7%. Along the first axis of the PCoA graph, the individuals were clustered into three major groups. The first cluster consisted of the individuals of Ayas and Kizlarkayasi populations, the second one consisted of the individuals from Oğlakçi and Asagikepen populations, and the third one consisted of the individuals from Kavuncu population (Figure 5). The PCoA and UPGMA cluster analysis were compatible with each other.

Bayesian (Structure) analysis was conducted to determine the genetic structure among 111 *S. yildirimlii* genotypes. This clustering approach assigns individuals to clusters based on genotype. The Structure program estimates the most likely number of clusters (K) by calculating the log probability of data for each K value and using ΔK statistics described by Evanno et al. (Evanno et al. 2005). The best K value for representing *S. yildirimlii* genotypes was K=3, with the highest peak, and the second peak was observed at K=5 (Figure 6).

4. Discussion

From 15 ISSR primers, we were able to analyse 111 individuals from 5 populations of the endemic and fragmentally distributed *S. yildirimlii* and extract 331 bands, of which 311 are polymorphic. It has been determined that 93.9% of the species' bands exhibit polymorphism. Previous studies showed that Lamiaceae family members like *Scutellaria baicalensis (Bai et al. 2013), Cunila spicata (Echeverrigaray et al. 2016)*, *Thymus daenensis* subsp. *daenensis* (Rahimmalek et al. 2009), *Salvia miltiorrhiza (Zhang et al. 2013),* the percentage of polymorphic bands based on ISSR markers varies from 85.32% (*Cunila spicata*) to 96.04% (*Scutellaria baicalensis*). We can conclude that *S. yildirimlii* has a comparatively high proportion of polymorphic bands at the species level compared to these findings. Nei's gene diversity (H), Shannon's information index (I), and percentage of polymorphic loci (PPL) were calculated for *S. yildirimlii* at species level as H_{sp}=0.183±0.178, I_{sp}=0.292±0.246, PPL_{sp}=93.9%, and at population level as H_{pop}=0.158±0.188, I_{pop}=0.242±0.268, PPL_{pop}=56.5%. When compared with the genetic diversity results of *Scutellaria baicalensis* (H_{sp}=0.246, PPL_{sp}=96.04% and H_{pop}=0.206 PPL_{pop}=73.44%) (Bai et al. 2013; Echeverrigaray et al. 2016), *Lamiophlomis rotate* (H_{sp}=0.291, PPL_{sp}=96.73% and H_{pop}=0.166 PPL_{pop}=51.81%) (Liu et al. 2006), *Satureja khuzistanica* (H_{sp}=0.306, PPL_{sp}=98.33% and H_{pop}=0.262, PPL_{pop}=74.88%) (Hadian et al. 2017) based on ISSR markers and allozyme analysis of *Scutellaria montana* (H_{sp}=0.374 and H_{pop}=0.287, PPL_{pop}=75.42%) (Cruzan 2001) from *Lamiaceae* family genetic diversity of *S. yildirimlii* presents comparable results at the species level but less so at the population level. The genetic diversity of *S. yildirimlii* reveals similar results to that of *Primula apennina* (H_{sp}=0.242, I=0.318, PPL_{sp}=96.95%) and *Uechtritzia armena* (H_{sp}=0.192, I=0.333, PPL_{sp}=96.21%) both endemic and threatened (Crema et al. 2009; Yıldırım Doğan et al. 2016). Consequently, *S. yildirimlii* can be classified as a perennial endemic herbaceous plant with considerable genetic diversity.

According to the number of individuals from the most to the least, the populations are in the following order: Kizlarkayasi, Asagikepen (Yesilkoy), Kavuncu, Oglakci and Ayas (Yıldırım et al. 2019) whereas the order of genetic diversity values (H, I, and PPL) is Oglakci, Asagikepen (Yesilkoy), Ayas, Kizlarkayasi and Kavuncu. These findings demonstrate that the size of a population is not necessarily proportional to its genetic diversity. PCoA and UPGMA analysis, both, revealed that the two populations with the highest genetic diversity, Asagikepen and Oglakci, are clustered together (Figure 2 & 3). Thus, it may be concluded that these two populations are not genetically isolated from one another and that gene flow between them is on going. Ayas population is the third most genetically diverse among the other populations despite having the smallest population size, with only 587 individuals (Yıldırım et al. 2019). It is estimated

that this population was substantial recently. However, it is assumed that there has been a sharp decline since 2010 due to the expansion of agricultural areas, reforestation, habitat degradation and loss (Ayyıldız 2010). Although having a considerable number of individuals (48000), Kizlarkayasi population ranks only fourth in terms of genetic diversity (Yıldırım et al. 2019). This outcome can be attributed primarily to genetic drift. Given that the habitat is steppe and the species is perennial, it is likely that a natural catastrophe such as a fire in the past induced genetic drift. Kavuncu population exhibits the lowest genetic diversity among the other populations of *S. yildirimlii*, and this was supported by PCoA and UPGMA analysis showing that it was apart from the others (Figure 2 & 3). This finding is consistent with earlier isolation from the other populations.

S. yildirimlii exhibited higher genetic diversity when compared to *Scutellaria indica*, a perennial plant, on both allozyme (H_{sp}=0.101, H_{pop}=0.008 PPL_{pop}=2.36%) and RAPD (H_{sp}=0.139, H_{pop}=0.03 PPL_{pop}=8.94%) bases. As a result of its low genetic diversity, *S.indica* displays a high level of genetic differentiation between its populations (allozyme G_{ST}=0.92 and RAPD G_{ST}=0.81). This *Scutellaria* species with floral dimorphism produces seeds through cross-pollination of its casmogamic flowers and self-pollination of its cleistogamic flowers. Therefore, these values were compatible with self-fertilized reproductive systems (Hamrick and Godt 1996). These values suggest that, in comparison to perennial self-fertilized species, *S. yildirimlii* has a higher level of heterozygosity and a higher value of gene flow (Hs=0.183 and G_{ST}=0.144). A high level of heterozygozity and polymorphism, along with a low level of inter-population differentiation (G_{ST}), are generally associated with outbreeding plant species (Loveless and Hamrick 1984). To comprehend the genetic differentiation of a species, it is necessary to have in-depth knowledge of pollination biology, seed dispersal systems, and the reproductive system. However, cross-pollination and pollination by insects (generally Hymenoptera, Lepidoptera, and Diptera) or birds are common in Lamiaceae family as the members of this family have floral morphology that allows bees to suck nectarine and provide a good source of nectarine and pollen (Watson and Dalwitz 1992; Judd et al. 2015). *S. yildirimlii* populations were found to have a G_{ST} of 0.144, indicating low levels of genetic differentiation between them. A value of 0.159 for Φ_{PT} was obtained with G_{ST} to assess the level of genetic differentiation between populations. These values corroborate one another, and we can conclude that *S. yildirimlii* populations exhibit a modest degree of genetic differentiation.

Variations revealed between and within populations as an outcome of AMOVA were 16% and 84%, respectively, demonstrating that genetic variation within the population is highly prevalent. The value of gene flow between *S. yildirimlii* populations, computed over the G_{ST} value, is significant (N_m=2.984). Gene flow $N_m \geq 1$ is generally sufficient to prevent major differenciation between populations, and it is attributed to the plant's ability to disperse pollen or seed over long distances (Sözen et al. 2017).

The Mantel test indicates a moderately positive and statistically significant relationship between genetic and geographic distances (R_{xy}=0.493, $p<0.001$). This correlation suggests that these two parameters are not directly related, which could be due to topographical differences. UPGMA and PCoA cluster analyses also confirm the results of the Mantel test. For instance, Ayas population is clustered with Kizlarkayasi,

which is geographically the nearest (60 km) as well as in terms of genetic distance (0.9748). Although Oglakci population is geographically closer to Kavuncu population by 22 km, it is genetically more similar to Asagikepen population (0.9784) by 29 km and clustered with it.

STRUCTURE analysis shows that the highest peak and second peak for representing *S. yildirimlii* genotypes were found at K=3 and K=5, respectively. These results for K=3 are consistent with those of UPGMA and PCoA. Concerning K=3, the first cluster A included 21 genotypes (18.9%), of which all from Kavuncu population had a probability of membership (threshold value) qi > 80%, while two genotypes with qi < 80% were admixed. The second cluster B also had 21 genotypes (18.9%), of which all from Ayas population had a probability of membership qi > 80%, while two genotypes with qi < 80% were admixed. There were 27 genotypes (24.3%) in group C, of which 11 were from Oglakci population, and 16 were from Asagikepen (with Yesilköy). The highest number of genotypes was admixed with the number of 42 (37.8%), which had the probability of membership qi < 80%, including all genotypes (21) from Kizlarkayasi population (genotype 24-44), with 2 from Ayas (genotype 2 and 23), 11 from Oglakci (genotype 45-49, 52-56 and 60), 6 from Asagikepen (genotype 82 and 84-88) and 2 from Kavuncu (genotype 102 and 110) populations as mentioned above (See Supplementary).

UPGMA dendrogram (Figure 4), except two individuals from Asagikepen population (87 and 88) and three individuals from Oglakci population (46, 47, and 54) that are clustered together with Kizlarkayasi population. The distance between the populations of Kizlarkayasi and Asagikepen is 55 km, while the distance between the populations of Kizlarkayasi and Oglakci is 27 km. This suggests the possibility of secondary seed dispersal between populations via birds.

5. Conclusion

In-situ conservation must be planned for all natural populations of *S. yildirimlii* to protect the species' existing genetic diversity. This is especially important for Ayas and Oglakci populations, which have relatively small populations that are found in restricted areas with a high level of genetic diversity. Ex-situ conservation techniques are also critical for future environmental challenges. Attempts can be undertaken to create new populations in similar habitats. Furthermore, preventative steps can be implemented to address issues such as habitat fragmentation and loss caused by agricultural expansion and road building that preventing gene flow between populations. Through leaflets and informative signage, locals should also be informed about conservation initiatives to reduce the pressure caused by unregulated and overgrazing in the species' distribution areas.

Declarations

Acknowledgements

This research was supported by Ankara University Research fund (Project no: 16H0430005). We are grateful for the support of Dr. Gül AYYILDIZ, Dr. Ali Murat KESER, Dr. Banu AVCIOĞLU and Asst. Prof. Dr.

İbrahim ÇELİK.

Conflict of interest

The authors have declared that no conficts of interest exist.

References

1. Ayyıldız G (2010) Aysantı Beli (Ayas-Ankara) Florasının Tehdit Altındaki Türleri. Dissertation, Gazi University
2. Bai C, Wen M, Zhang L, Li G (2013) Genetic diversity and sampling strategy of Scutellaria baicalensis germplasm resources based on ISSR. Genet Resour Crop Evol 60:1673–1685
3. Boydak M, Teker T, Gazdağli A, et al (2021) ISSR genotyping of Phoenix theophrasti natural populations in Turkey and Crete (Greece) and P. dactylifera. Nord J Bot 39:1–11
4. Caccone A (1985) Gene flow in cave arthropods: a qualitative and quantitative approach. Evolution (N Y) 39:1223–1235
5. Celik I, Camci H, Kose A, et al (2016) Molecular genetic diversity and association mapping of morphine content and agronomic traits in Turkish opium poppy (Papaver somniferum) germplasm. Mol Breed 36:1–13
6. Cicek M, Yaprak AE (2013) Scutellaria yildirimlii (Lamiaceae), a new species from Turkey. Phytotaxa 132:53–58
7. Crema S, Cristofolini G, Rossi M, Conte L (2009) High genetic diversity detected in the endemic Primula apennina Widmer (Primulaceae) using ISSR fingerprinting. Plant Syst Evol 280:29–36
8. Cruzan MB (2001) Population size and fragmentation thresholds for the maintenance of genetic diversity in the herbaceous endemic Scutellaria montana (Lamiaceae). Evolution (N Y) 55:1569–1580
9. Don RH, Cox PT, Wainwright BJ, et al (1991) 'Touchdown'PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 19:4008
10. Earl DA, VonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361
11. Echeverrigaray S, Albuquerque M, Zacaria J, Longaray AP (2016) Population Structure and Genetic Diversity of the Medicinal and Aromatic Plant Cunila spicata Benth.(Lamiaceae) Based on Inter-Simple Sequence Repeat (ISSR) Markers. J Adv Biotechnol 5:652–658
12. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620
13. Excoffier L, Smouse PE, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

14. Frankham R, Ballou SEJD, Briscoe DA, Ballou JD (2002) Introduction to conservation genetics. Cambridge University Press
15. Güner A, Aslan S, Ekim T, et al (2012) Türkiye bitkileri listesi:(damarlı bitkiler). Nezahat Gökyiğit Botanik Bahçesi Yayınları
16. Hadian J, Raeisi S, Azizi A, et al (2017) Genetic diversity of natural populations of medicinally valuable plant Satureja khuzistanica Jamzad based on ISSR markers. Brazilian J Bot 40:771–781
17. Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. Philos Trans R Soc London Ser B Biol Sci 351:1291–1298
18. Harley RM, Atkins S, Budantsev AL, et al (2004) Labiatae—In: Kaderit, JW (Ed.), Kubitzki, K.(Ed. in chief). The families and genera of vascular plants, vol 7. Flowering plants. Dycotiledons: Lamiales (except Acanthaceae including Avicenniaceae)
19. Jaccard P (1908) Nouvelles recherches sur la distribution florale. Bull Soc Vaud Sci Nat 44:223–270
20. Judd W, Campbell C, Kellogg E, et al (2015) Plant Systematics: A Phylogenetic Approach. Sinauer Associates Sunderland, Massachusetts
21. Kesawat MS, Das Kumar B (2009) Molecular markers: it's application in crop improvement. J Crop Sci Biotechnol 12:169–181
22. Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. Genetics 49:725–738
23. Lewontin RC (1972) The apportionment of human diversity. In: Evolutionary biology. Springer, pp 381–398
24. Liu BH (1998) Statistical genomics: linkage, mapping, and QTL analysis. CRC press
25. Liu J, Wang L, Geng Y, et al (2006) Genetic diversity and population structure of Lamiophlomis rotata (Lamiaceae), an endemic species of Qinghai–Tibet Plateau. Genetica 128:385–394
26. Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. Annu Rev Ecol Syst 15:65–95
27. Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220
28. McDermott JM, McDonald BA (1993) Gene flow in plant pathosystems. Annu Rev Phytopathol 31:353–373
29. Minareci E, Pekönür S (2017) An Important Euroasian Genus: Scutellaria L. Int J Second Metab 4:35–46
30. Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci 70:3321–3323
31. Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590
32. Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York
33. Paton A (1989) The phytogeography of Scutellaria L. Notes from R Bot Gard (United Kingdom)
34. Paton A (1990) A global taxonomic investigation of Scutellaria (Labiatae). Kew Bull 399–450

35. Peakall ROD, Smouse PE (2012) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research - an update. Bioinformatics 28:2537–2539
36. Podani J (2001) Computer Programs for Data Analysis in Ecology and Systematics. User's Manual, Budapest
37. Primack RB (2006) Essentials of conservation biology. Sinauer Associates Sunderland
38. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
39. Qian W, Ge S, Hong D-Y (2001) Genetic variation within and among populations of a wild rice Oryza granulata from China detected by RAPD and ISSR markers. Theor Appl Genet 102:440–449
40. Rahimmalek M, Bahreininejad B, Khorrami M, Sayed Tabatabaei BE (2009) Genetic variability and geographic differentiation in Thymus daenensis subsp. daenensis, an endangered medicinal plant, as revealed by inter simple sequence repeat (ISSR) markers. Biochem Genet 47:831–842
41. Slatkin M (1981) Estimating levels of gene flow in natural populations. Genetics 99:323–335
42. Slatkin M (1985) Rare alleles as indicators of gene flow. Evolution (N Y) 39:53–65
43. Sözen E, Hilooglu M, Kandemir A (2017) Genetic diversity of local endemic Teucrium leucophyllum Montbret & Aucher ex Bentham.(Lamiaceae) in Turkey. Indian J Pharm Educ Res 51:s195–s199
44. Tamura K, Stecher G, Peterson D, et al (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729
45. Waples RS (1987) A multispecies approach to the analysis of gene flow in marine shore fishes. Evolution (N Y) 41:385–400
46. Watson L, Dalwitz MJ (1992) The Families of Flowering Plants: Descriptions, Illustrations, Identification, and Information Retrieval. Version: 26th October 2019. delta-intkey.com
47. WCSP (2022) World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. http://wcsp.science.kew.org/. Accessed 3 Jul 2022
48. Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396
49. Wright S (1949) The genetical structure of populations. Ann Eugen 15:323–354
50. Wu FQ, Shen SK, Zhang XJ, et al (2015) Genetic diversity and population structure of an extremely endangered species: the world's largest Rhododendron. AoB Plants 7:plu082
51. Yeh FC, Yang RC, Boyle TBJ, et al (1997) POPGENE, the user-friendly shareware for population genetic analysis. Mol Biol Biotechnol centre, Univ Alberta, Canada 10:295–301
52. Yıldırım Doğan N, Kandemir A, Osma E (2016) Genetic diversity and variability among populations and ecological characteristics of the Uechtritzia armena Freyn (Asteraceae) endemic to Turkey. 4:20–27
53. Yıldırım M, Ayyıldız G, Keser AM, et al (2019) Current Population Sizes, Distribution Areas and Re-Evaluated IUCN Categories of Rare and Endemic Species from Central Anatolia, Turkey: Salsola grandis, Scutellaria yildirimlii and Sideritis gulendamii. Biyolojik Çeşitlilik ve Koruma 12:151–160

54. Zhang Y, Li X, Wang Z (2013) Diversity evaluation of Salvia miltiorrhiza using ISSR markers. Biochem Genet 51:707–721
55. Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20:176–183

Figures

Figure 1

The distribution areas of *S. yildirimlii*

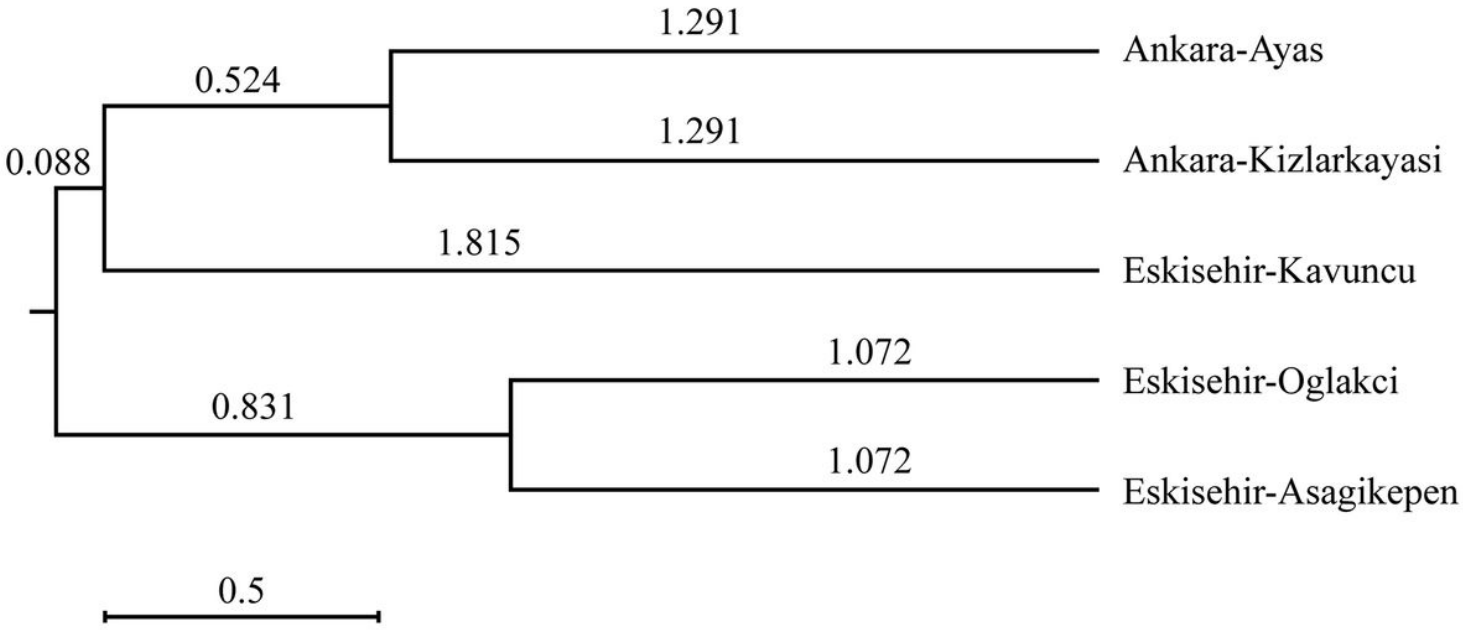

Figure 2

UPGMA dendrogram based on distances among *S. yildirimlii* populations

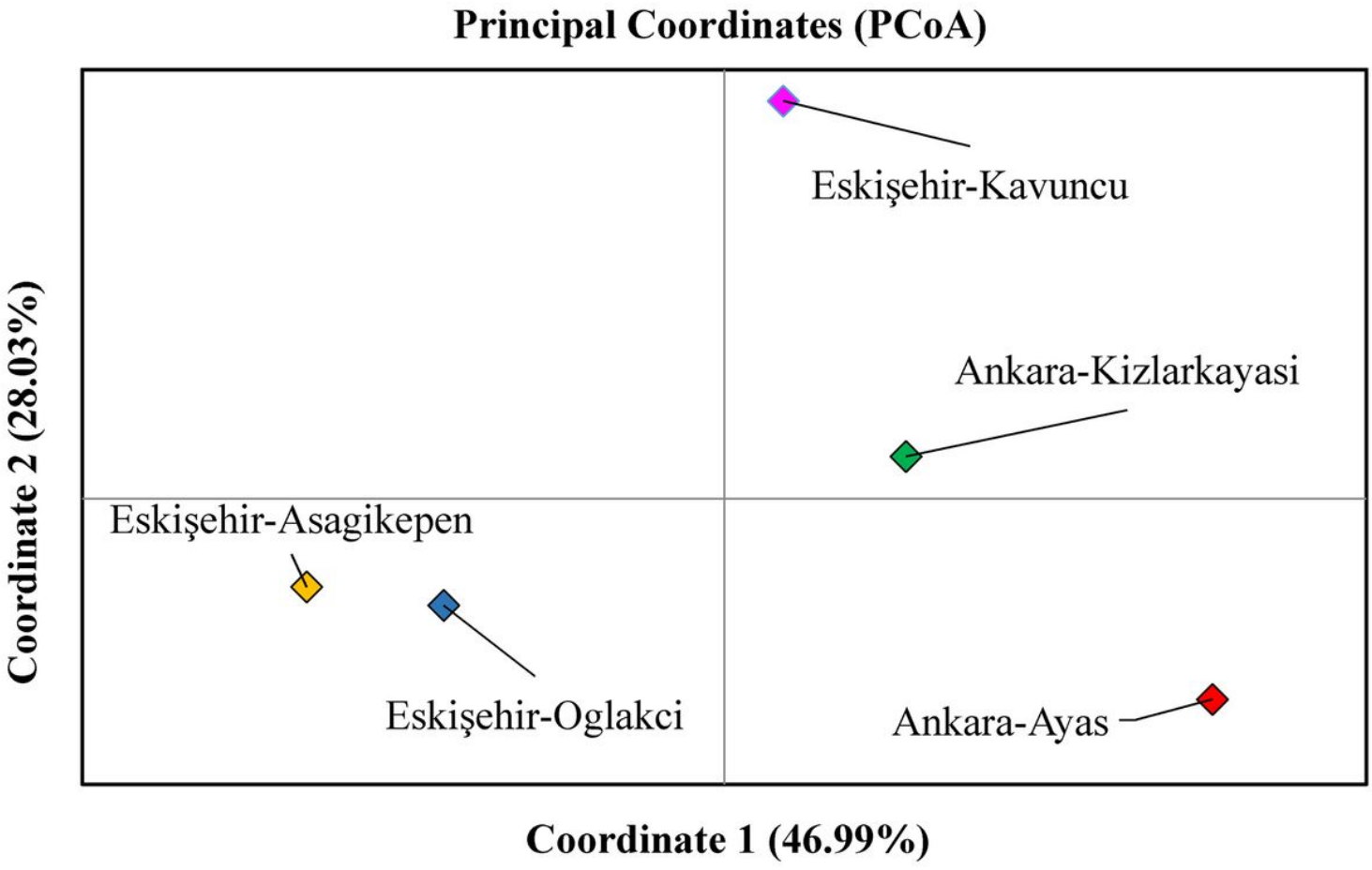

Figure 3

Principal Coordinate Analysis (PCoA) graph of *S. yildirimlii* populations

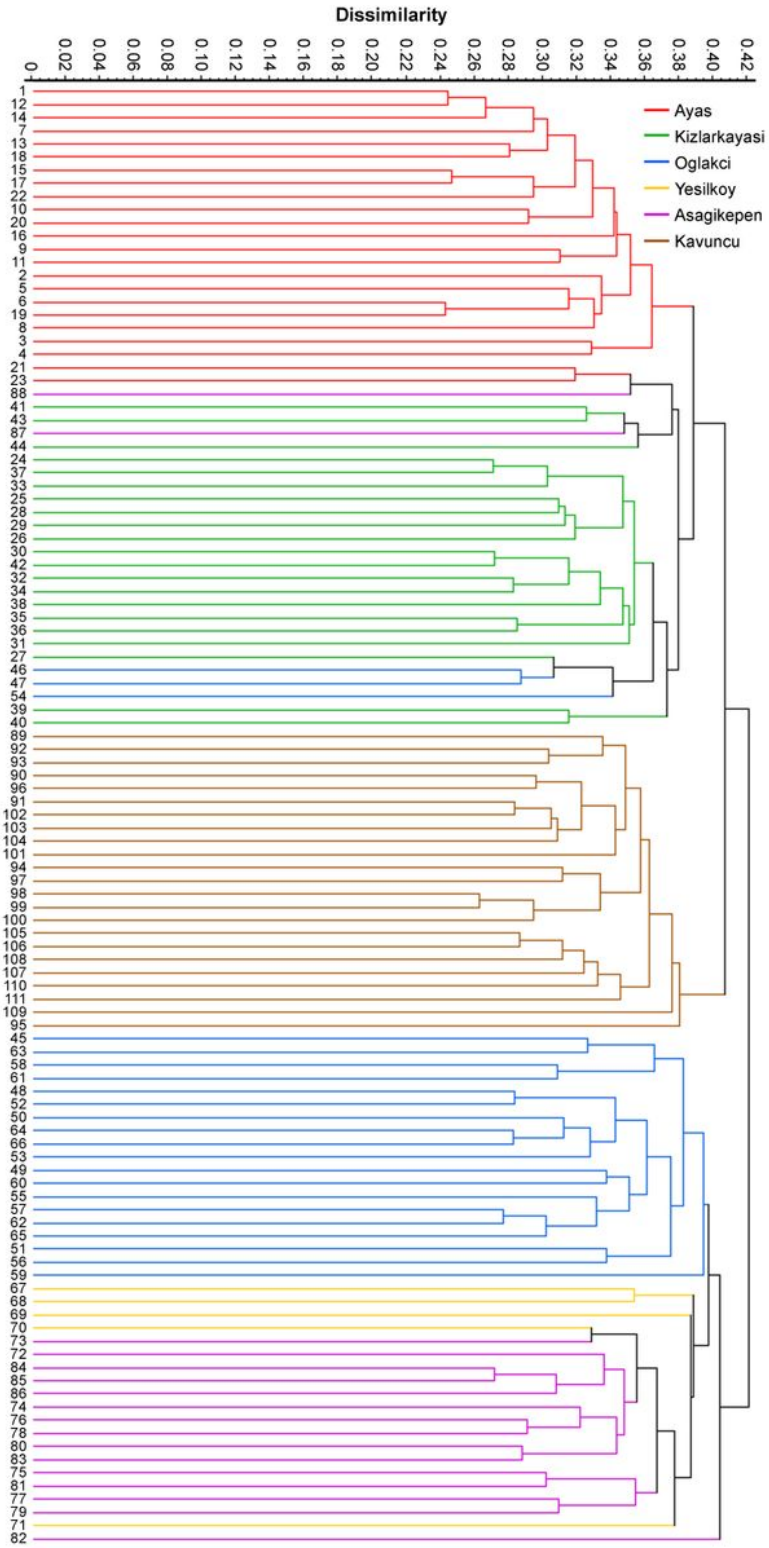

Figure 4

UPGMA dendrogram showing the genetic distance between 111 genotypes of *S. yildirimlii*

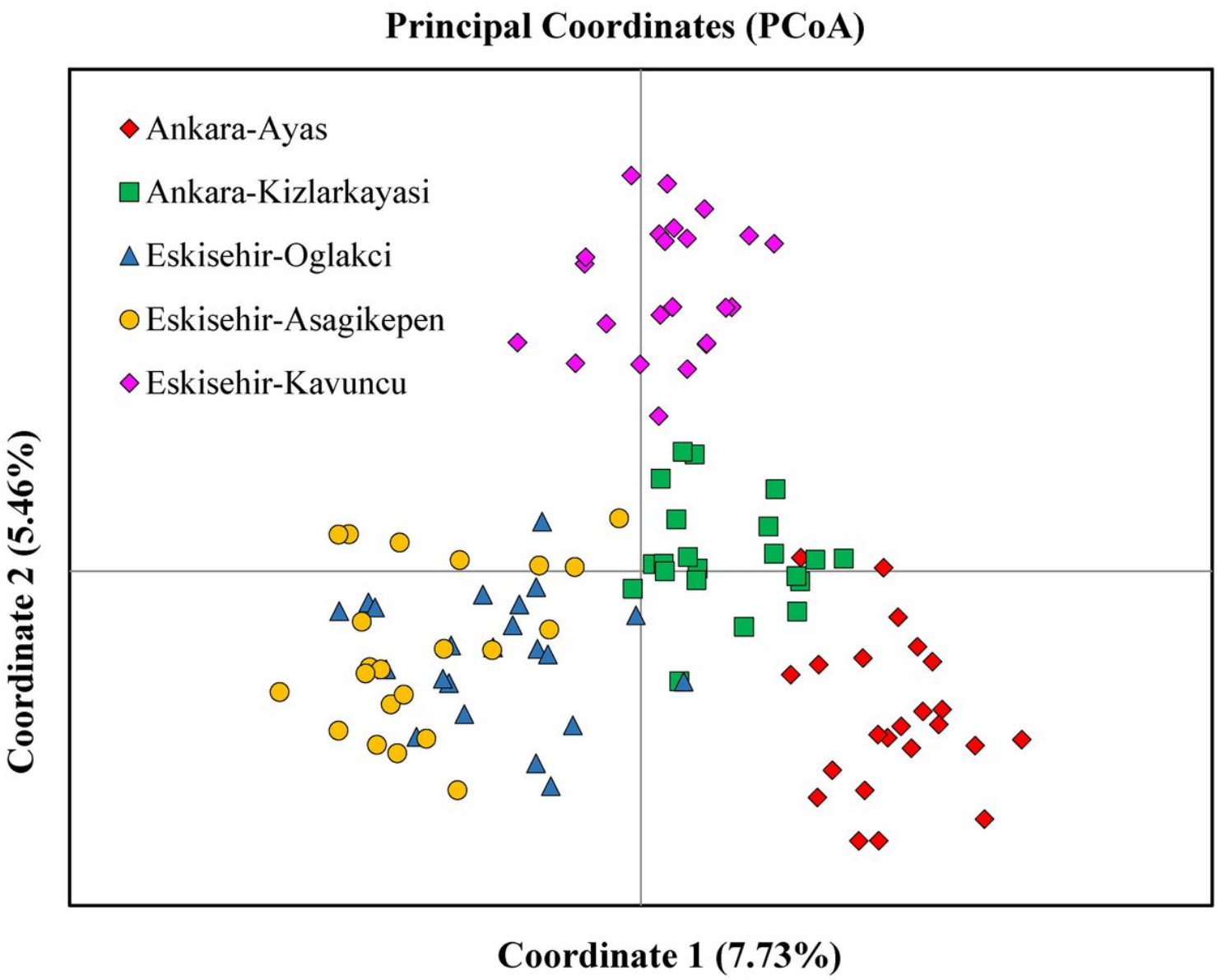

Figure 5

Principal Coordinate Analysis (PCoA) graph showing the spatial distribution of 111 *S. yildirimlii* genotypes

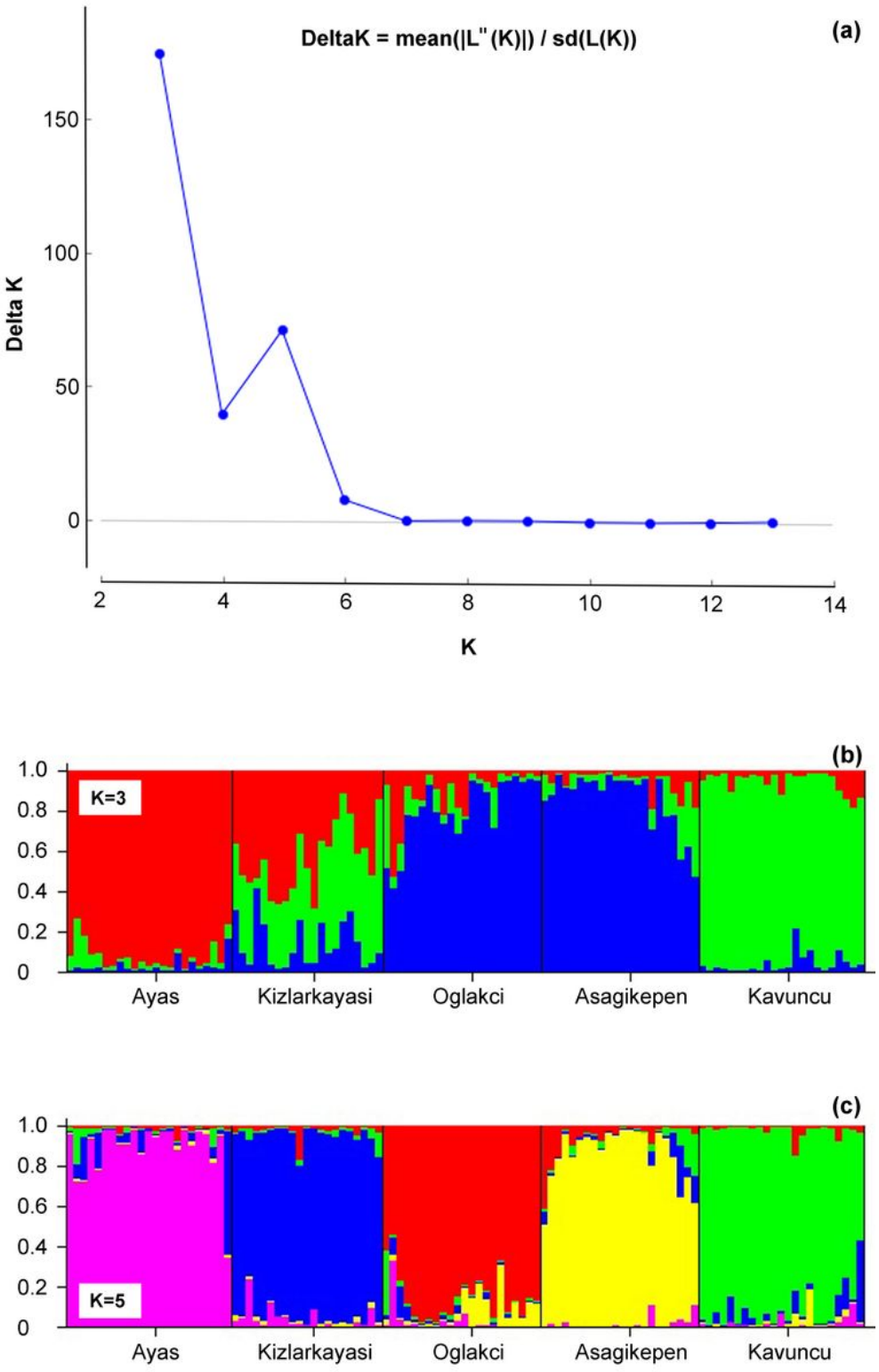

Figure 6

(a) STRUCTURE plot of population structure analysis and Delta-K values,
(b) population structure analysis K = 3 and (c) population structure analysis K = 5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.docx