Phylogenetic Analysis of Brown Bear *(Ursus arctos)* Populations Across Türkiye

by

Elif Çeltik

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Molecular Biology and Genetics

KOÇ ÜNİVERSİTESİ

July, 2023

Phylogenetic Analysis of Brown Bear *(Ursus arctos)* Populations Across Türkiye

Koç University
Graduate School of Sciences and Engineering
This is to certify that I have examined this copy of a master's thesis by

Elif Çeltik

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

____
Prof. Dr Çağan H. Şekercioğlu

____
Asst. Prof. İsmail Kudret Sağlam

____
Asst. Prof. Emrah Çoraman

Date: ____________________________

[To my beloved family]

ABSTRACT

Phylogenetic Analysis of Brown Bear *(Ursus arctos)* Populations Across Türkiye
Elif Çeltik
Master of Science in Molecular Biology and Genetics
July, 2023

Brown bears (*Ursus arctos*) are omnivores distributed across Eurasia and North America. Previous studies based on mitochondrial DNA (mtDNA) revealed significant differences in genetic variation between European and Middle Eastern brown bears, with each region characterized by distinct haplogroups. Türkiye, as a crossroads between three continents, is an important region for brown bear diversity, as it harbors haplotypes from both regions. However, previous information on brown bear genetic diversity in Türkiye was based on only mtDNA sequences and limited sample sizes. Therefore, the role Türkiye has played in the evolutionary history of brown bears is still largely unknown. Here, using genome-wide sequencing data together with the mtDNA D-loop region sequences, we aim to fill this gap by investigating the patterns of brown bear genetic diversity and differentiation across Türkiye and nearby regions. Our results demonstrate the geographical distribution of haplogroups across Türkiye and show that Turkish populations have three distinct mtDNA haplogroups clustered into western and eastern geographic groups. Genome-wide SNP analysis supported this conclusion, with high patterns of genetic diversity and distinct geographic structuring. Collectively, our results indicate that the evolutionary dynamics influencing brown bear populations across Türkiye and the Middle East are distinct and Türkiye's brown bears have higher genetic diversity than other world populations.

ÖZETÇE

Türkiye'deki Bozayı Populasyonlarının Filogenetik Analizi
Elif Çeltik
Moleküler Biyoloji ve Genetik, Yüksek Lisans
24 Temmuz 2023

Bozayılar, Avrasya hattından Amerika'ya kadar çok geniş bir bölgede yaşayan omnivorlardır ve Türkiye'de de geniş bir coğrafi alana dağılmış şekilde yaşamaktadırlar. Mitokondriyal DNA (mtDNA) kullanılarak yapılan önceki çalışmalar, Avrupa ve Orta Doğu bozayıları arasındaki genetik varyasyonda önemli farklılıklar ortaya çıkardı, ve her bölge farklı haplogruplarla karakterize edildi. Üç kıta arasında bir kavşak konumunda bulunan Türkiye, her bölgeden haplotipler barındırması nedeniyle genetik çeşitlilik açısından önemli bir bölgedir. Türkiye bozayılarının filogenetiği üzerine yapılan daha önceki çalışmalar sadece mtDNA sekanslarını içeriyordu, ve sınırlı örneklemeye ve dağılıma sahipti. Bu nedenle Türkiye ve Orta Doğu bozayılarının evrim ağacındaki yeri tam olarak anlaşılamamıştır. Bu çalışmada genom dizileme verilerini ve mtDNA'nın D-loop dizilerini kullanarak Orta Doğu populasyonlarının evrimsel geçmişine ışık tutmayı, Türkiye ve yakın bölgelerdeki genetik çeşitlilik ve farklılaşma modellerini araştırmayı amaçladık. Bulgularımız, Türkiye popülasyonlarının batı ve doğu coğrafi gruplar halinde kümelenmiş üç farklı mtDNA haplogrubuna sahip olduğunu gösterdi ve her bir haplogrupun Türkiye genelindeki coğrafi dağılımını ortaya koydu. Genom verimizle yaptığımız SNP analizi de yüksek genetik çeşitlilik ve net olarak ayrılmış populasyon yapısıyla bu sonucu destekler niteliktedir. Sonuçlarımız, Türkiye ve Orta Doğu'daki bozayı populasyonlarını etkileyen evrimsel dinamiklerin farklı olduğunu ve Türkiye bozayılarının diğer dünya popülasyonlarına göre daha yüksek genetik çeşitliliğe sahip olduğunu göstermektedir.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Prof. Dr. Çağan Şekercioğlu, for his continuous support and guidance throughout my MSc studies. His passion for science will always inspire me, it was a great chance to study with one of my role models. I am grateful that Prof. Dr. Çağan Şekercioğlu and Dr. Morteza Naderi gave me the opportunity to work on the Tubitak project and conduct my MSc research at Koç University. I would especially like to thank Dr. Morteza Naderi for funding me as a master's student and for his contributions to my studies.

I want to thank Asst. Prof. İsmail Kudret Sağlam for his advice, valuable time, and endless support. I am grateful for his teaching me critical bioinformatics skills and how to analyze genomic data. I would specially thank him for making time for me every week and providing helpful feedback and suggestions for my study. With his contributions, I believe that I gained scientific knowledge and perspective. Most importantly, I improved my computational skills, which was a top priority during my MSc studies.

I would especially like to thank Dr. Mübeccel Çisel Kemahlı Aytekin. She was my supporter and mentor during computational and wet lab studies. I thank her for teaching me bioinformatics. She improved my scientific thinking ability with the solutions she found to scientific problems.

I thank my committee member Asst. Prof. Emrah Çoraman, and to my colleagues in my lab, Ercan Sıkdokur, Elif Başarır, Seray Altıntaş, who provided feedback and encouragement. Special thanks to Ercan Sıkdokur, Elif Başarır, and Seray Altıntaş for helping me with the laboratory processes. We completed the project together with great effort and hard work, and they always supported me. In addition, I want to thank Emrullah Yılmaz for his assistance with computational analyses.

Most notably, I want to thank my family for their unwavering love and support, which keeps me inspired and self-assured. They always believed in me and supported me.

This thesis was funded by a Tübitak 1001 project (219Z066) titled ”Landscape genetics and niche modeling of brown bear populations in Türkiye”.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ANGSD	Analysis of Next Generation Sequencing
Cq	Quantification cycle
DNA	Deoxyribonucleic Acid
dsDNA	Double-stranded deoxyribonucleic acid
Hd	Haplotype diversity
kya	Thousands years ago
LGM	Last Glacial Maximum
ML	Maximum Likelihood
NCBI	National Center for Biotechnology Information
PCA	Principal component analysis
PCR	Polymerase Chain Reaction
PopART	Population Analysis with Reticulate Trees
RAD	Restriction-site associated
SNP	Single Nucleotide Polymorphism
TMRCA	Time to the Most Recent Common Ancestor

Chapter 1

INTRODUCTION

1.1 The biology of brown bear

Ursus arctos, composed of the Latin words *Ursus* (bear) and *arctos* (also bear), is included in the genus *Ursus* because they have bear characteristics such as large bodies, long hair, short legs and tails. They belong to the Ursidae family and Caniforms, with long snouts and non-retractible claws. Ursidae are most closely related to Canidae, which includes dogs, foxes, jackals, wolves, and coyotes. The order of brown bears is Carnivora, characterized by sharp teeth and claws, in the class Mammalia. Mammals, including bears, have hair, three middle ear bones, sweat glands, a four-chambered heart, and a neocortex. If the mammal is female, it also has mammary glands. They also have a well-develeoped neocortex, a brain region located in the cerebral cortex responsible for reasoning, sensory perception, cognition, and language.

Brown bear (*Ursus arctos*) is one of the largest carnivore species in Europe. Brown bears typically have a dark brown color, but they can range in color from a light tone to black, and the size of brown bears may change depending on the amount of available food. They are omnivores, and their diet varies in content according to local food availability, season, location (latitude, longitude, and altitude), as well as environmental conditions such as temperature and snow cover. They usually eat vegetation such as grasses, nuts, berries, sedges, bulbs, and roots in Türkiye. Brown bears are easily identified by their large heads, long noses, and small round ears. Males are much larger and heavier than females in all populations. According to habitat, a male's body mass can change between 140 and 360 kilograms, while females may weigh between 100 and 200 kilograms. Brown bears have a stumpy tail that blends in with the rest of the bear's thick hair. Bears' legs are short and muscular, with huge paws and long claws. Young bears, in particular, are good climbers because of their claws (Euronatur, n.d.).

The annual cycle of bears occurs in four primary physiological stages: winter dormancy, hypophagia, normal activity, and hyperphagia. The length of winter dormancy may change depending on where they live. For instance, dormancy can last up to seven months in the north, whereas some individuals do not undergo dormancy and stay active all year in warmer regions. After they enter the den for dormancy, their metabolism begins to slow down. Their heart rate, breathing, and oxygen intake all decrease. Bears stop eating, urinating, and defecating. The next stage is hypophagia. Bears gradually resume their normal activity level during the first 2-3 weeks after leaving the den. They continue to eat and drink minimally. Their energy requirement and their appetite rise gradually. Adult bears spend all their energy on mating, while mothers devote all their energy to caring for their newborn cubs. Hyperphagia is the next step. Bears begin to gain weight around the end of summer, and their daily caloric requirements might reach 20.000 kcal. Bears drink a lot throughout these months to digest all the food they ingest. A bear's urine output can reach 16 liters per day. Feeding takes up most of a bear's time, sometimes up to 14 hours per day, to gain body fat and lean muscle (Nelson et al., 1983).

1.2 The distribution of brown bears

To understand the distribution of brown bears, there are several molecular and traditional methods. Genetic information in DNA may be used to answer questions regarding bear population distribution. Through molecular analyses, the characteristics of natural populations, such as gene flow, reproductive success, genetic diversity, and evolutionary lineages, can be defined. Furthermore, traditional methods such as paleontological and morphological data can be used to understand the evolutionary history of species.

Our study focuses on the distribution of brown bear populations in Türkiye. Habitat type and quality are the main factors that determine the distribution of species. Brown bears can live in a variety of habitats (Ambarlı, 2006, 2012). Most brown bears in Türkiye live in natural habitats around the Black Sea and in Eastern Anatolia. In addition, they can be found in other parts of Türkiye, including the Aegean, Mediterranean, Marmara, and Central Anatolia regions (Figure 1.1). Eastern Türkiye's forested lands are not as extensive as those around the Black Sea. Thus, suitable habitat is fragmented, and this may lead to sub-populations. Furthermore, isolated populations may occur in western Türkiye since human activity has rendered many forest corridors impassable (Can & Togan, 2004).

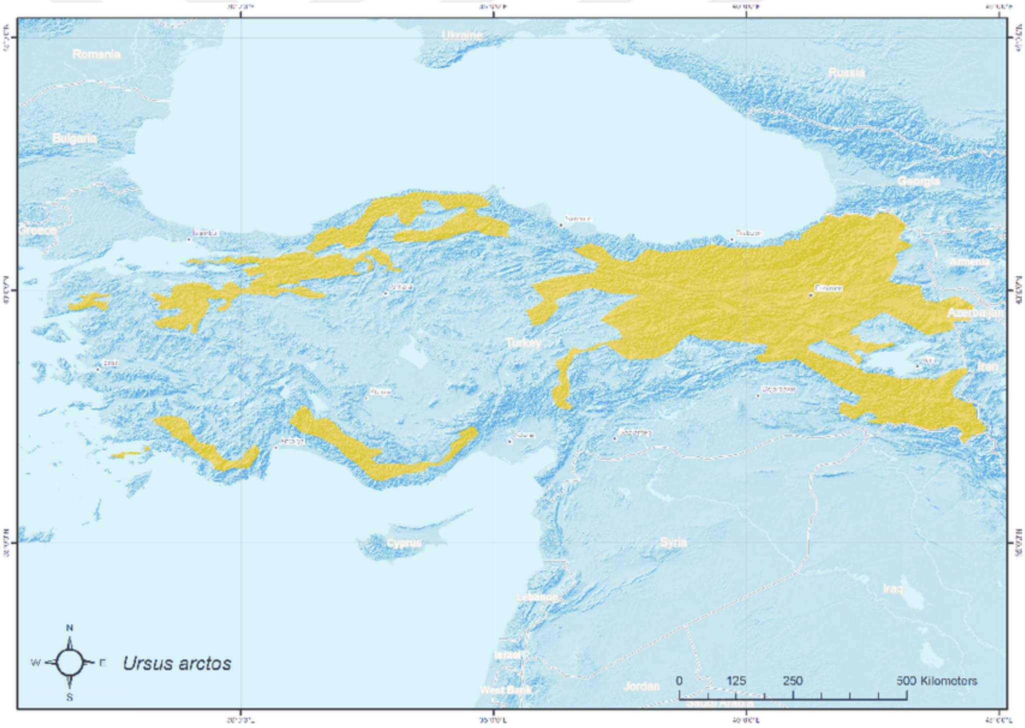


Figure 1.1: The distribution of the brown bears in Türkiye (Ambarli et al., 2016)

1.3 Phylogenetic trees

Phylogenetic trees are used to analyze the evolutionary connections of species or populations and are one of the main tools of this study. A phylogenetic tree can be constructed to study evolutionary relationships, diversification, speciation, and subclade determination (Mclennan, 2010). Jean-Baptiste Lamarck is attributed with creating the first branching diagrams to illustrate evolution (Lamarck, 1809, 1815). Then, Charles Darwin drew a branching diagram in his notebook (1837) before publishing a phylogenetic tree in his book on the Origin of Species. Today, phylogenetic trees are a crucial tool for understanding the relationships between species over time (Dees, Momsen, Niemi, & Montplaisir, 2014; Nyffeler et al., 2005).

A phylogenetic tree calculates the distance between sequences and their common ancestors (B. G. Hall, 2013; S. Kumar et al., 2001). It consists of external/internal nodes and branches. External nodes demonstrate the actual sequences of existing species, while internal nodes indicate hypothetical ancestors. Nodes are connected to each other through branches (B. G. Hall, 2013). Phylogenetic trees can be positioned in a variety of ways, including sideways, inverted, curved branches, and radial trees. The branching patterns provide the same information regardless of how the tree is drawn: evolutionary lineage and divergence patterns. To build phylogenetic trees, different types of data containing morphological data such as organ types, structural characteristics, skeletal configurations, and molecular data such as mtDNA, ribosomal RNA, or any genes of interest may all be used (MacDonald & Thanukos, n.d.). Molecular data such as DNA or protein sequences are most used to create phylogenetic trees. A phylogenetic tree is constructed in three steps. The first goal is to identify and obtain homologous DNA or protein sequences. The sequences are then aligned. Finally, the algorithm estimates a phylogenetic tree from the aligned sequences. There are several algorithms to construct the phylogenetic tree. Typically, the algorithm begins with a good tree, and branches are randomly swapped, with each swap determining whether the tree improved by the criterion. The most common criteria are parsimony and model-based strategies such as bayesian and maximum likelihood trees. Parsimony assumes that evolutionary events are uncommon and therefore chooses the reconstruction with the fewest postulated events, which are the most accurate. Model-based techniques begin by assuming a specific model for the probability of evolutionary events occurring and then comparing trees using those models.

Maximum likelihood trees calculate the probability of seeing the data we observe given the tree we assume and choose the tree with the highest probability while Bayesian trees compute the probability of getting the tree we assume given the data we have and choose the tree with the highest probability. After building the tree, we should evaluate its reliability. The standard statistical way of assessing a tree's validity is to resample data to see if we get the same tree every time. The bootstrapping approach is typically used to resample and generate new datasets. The tree with the highest criterion score is displayed, as are the bootstrapping values for each node (B. G. Hall, 2013).

1.4 Molecular markers

A phylogenetic tree can be built using a variety of molecular markers. Nuclear DNA, Y chromosome, and mitochondrial DNA markers are three different types of molecular markers. Nuclear and mitochondrial DNA are two types of DNA found in mammalian cells. Nuclear DNA is transmitted from both parents and is present in the nucleus of cells. Each nuclear chromosome has two copies of cells, one from the mother and the other from the father. The Y chromosome is also located in the cell nucleus, but it differs from other nuclear DNA chromosomes in that it has a unique characteristic. It is passed down from the father to his offspring. mtDNA, located in the mitochondrion, is a cytoplasmic cellular organelle and is inherited from female to offspring (Lansman et al., 1983). These various inheritance patterns influence how phylogenetic trees are interpreted. mtDNA may give information about maternal evolutionary history, whereas Y-chromosome markers give information about paternal evolutionary history (Herrero et al., 1999).

Mitochondrial DNA (mtDNA) has its genome located in the matrix of the mitochondria. Each organelle in mammalian cells usually contains many identical copies of mtDNA, and mtDNA is only passed through the female germline (Kivisild, 2015).The mammal mitochondrial genome is highly conserved and composed of double-stranded circular DNA molecules. Circular mitochondrial DNA contains a non-coding segment known as a displacement (D) loop (Kasamatsu et al., 1971) also known as the control region, shown in Figure 1.2. Large parts of the D loop region display high levels of variation compared to the coding regions and accumulate mutations over time. Thus, the D-loop region is critical for studies about the evolutionary history of vertebrates (Larizza et al., 2002). The length of D-loop region varies between 880 and 1400 bp. The length can be extended due to repetitive sequences observed in different species (Sbisà et al., 1997). Since D-loop region is the most rapidly evolving region of the mitochondrial genome (Lopez et al., 1997), we decided to compare the D-loop sequences of brown bear populations.

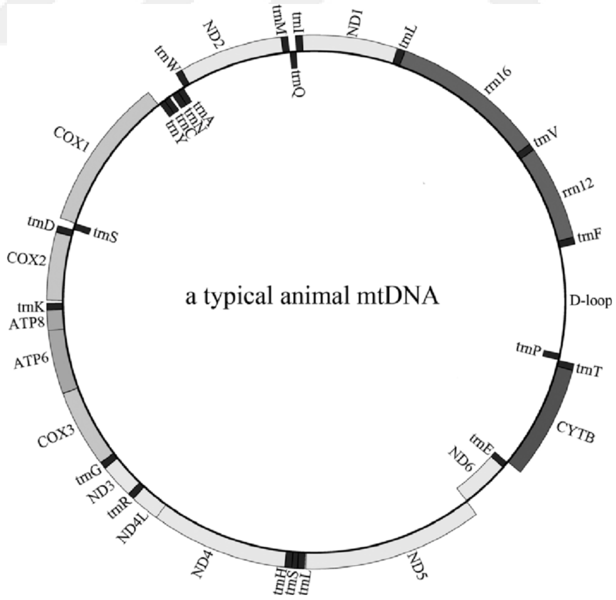


Figure 1.2: Circular representation of animal mitochondrial DNA (Wehausen et al., 2004)

We also used restriction site-associated DNA (RAD) as a nuclear DNA marker, which is a fractionated genome examining 0.1 to 10 of a genome. RAD sequencing mainly targets orthologous regions distributed throughout the genome by applying restriction sites. As a result, providing comparative genomic data for molecular systematics may be a more cost-effective technique than sequencing and assembling whole genomes. RAD markers can be utilized for multiple purposes, including genetic variation detection, phylogenetics, population structure, and genome-wide association studies (Rubin et al., 2012). Studies by Cariou et al. (2013) and Hohenlohe et al. (2010) claimed that RAD sequences are suitable for interspecific and intraspecific genetic comparisons.

1.5 Brown bear evolutionary history

Ursine bears are a mammalian subfamily with six extant species that are morphologically and ecologically distinct. The subfamily includes the American black bear *(Ursus americanus)*, Asian black bear *(Ursus thibetanus)*, sun bear *(Helarctos mayalanus)*, sloth bear *(Melursus ursinus)*, brown bear *(Ursus arctos)*, polar bear *(Ursus maritimus)*, and many extinct taxa (Kutschera et al., 2014). The evolutionary relationships between bears are demonstrated in a Bayesian tree that used complete mitochondrial genomes (Figure 1.3). According to this study, the sister group of polar and brown bears is the extinct cave bear (*Ursus spelaeus*) with all tree species clustering together. Interestingly, brown bears in the ABC islands diverged from grizzly bears and are shown to be closer to polar bears. Sun bear, Asiatic black bear, and American black bear were clustered separately from the other bear species. Spectacled bear comprises the sister group to all ursine bears.

Polar and brown bears are members of the Ursinae subfamily, which has six extant, morphologically, and ecologically different species, but hybridization of some ursine species has been reported. It is known that natural breeding between grizzlies and polar bears produces hybrid offspring in bears. Furthermore, Asiatic black bear and sun bear can hybridize. Molecular phylogenetic analyses could not precisely reconstruct the relationship between the six ursine bear species, because of small population sizes for most bears and possibly different habitats, ecology and morphology (V. Kumar et al., 2017).

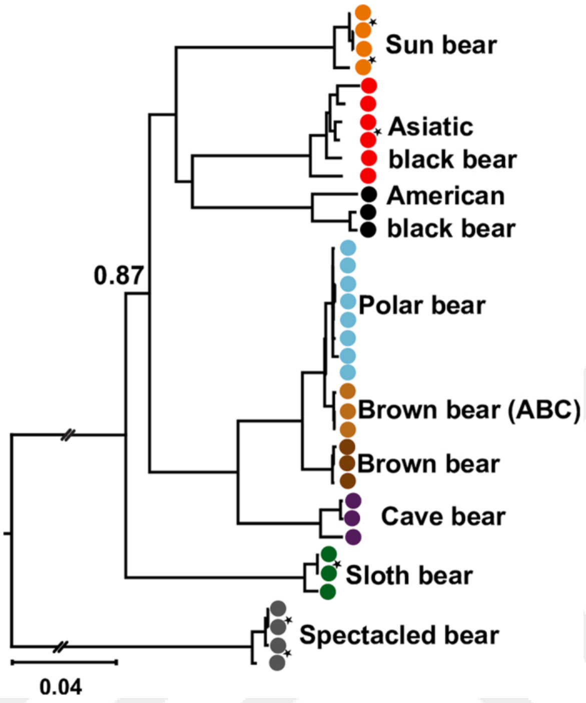


Figure 1.3: The phylogenetic tree of bear species (Tao et al., 2014)

Brown bears is the only bear species in Türkiye. Brown bears are classified into seven haplogroups: Clade 1 (West of Europe), Clade 2a (ABC islands), Clade 3 (Eastern Europe, Western Alaska, Eastern Russia, Hokkaido, and Anatolia) (Hassanin, 2015; Korsten et al., 2009; Murtskhvaladze et al., 2010; Saarma et al., 2007; Tammeleht et al., 2010), Clade 4 (Grizzly bears), Clade 5 (Tibet), Clade 6 (Gobi bear and Himalaya) (Tumendemberel et al., 2019) and Clade 7 (Middle east) (Çilingir et al., 2016). The distribution of these clades are illustrated in Figure 1.4, based on data gathered from the literature (Hirata et al., 2013; Mizumachi et al., 2020; Waits et al., 1998). Clades 3, 1 and 7 are divided into subclades. Clade 3 is the most widespread, with three subclades (3a1, 3a2, and 3b). 3a1 is a widely distributed clade in Eurasia, Russia, and Alaska. 3a2 is found in central Hokkaido, while 3b is found in eastern Hokkaido (Hirata et al., 2013; Tumendemberel et al., 2019). Clade 1 also has three subclades with Eastern Europe (1b), Scandinavia (1a), and Taurus–

Levant (1d) (Çilingir et al., 2016; Hassanin, 2015; Kohn et al., 1995; Taberlet & Bouvet, 1994). Clade 7 is divided into two subclades Middle East-Iran 7a, Middle East-Türkiye 7a and Middle East-divergent 7b (Çilingir et al., 2016).

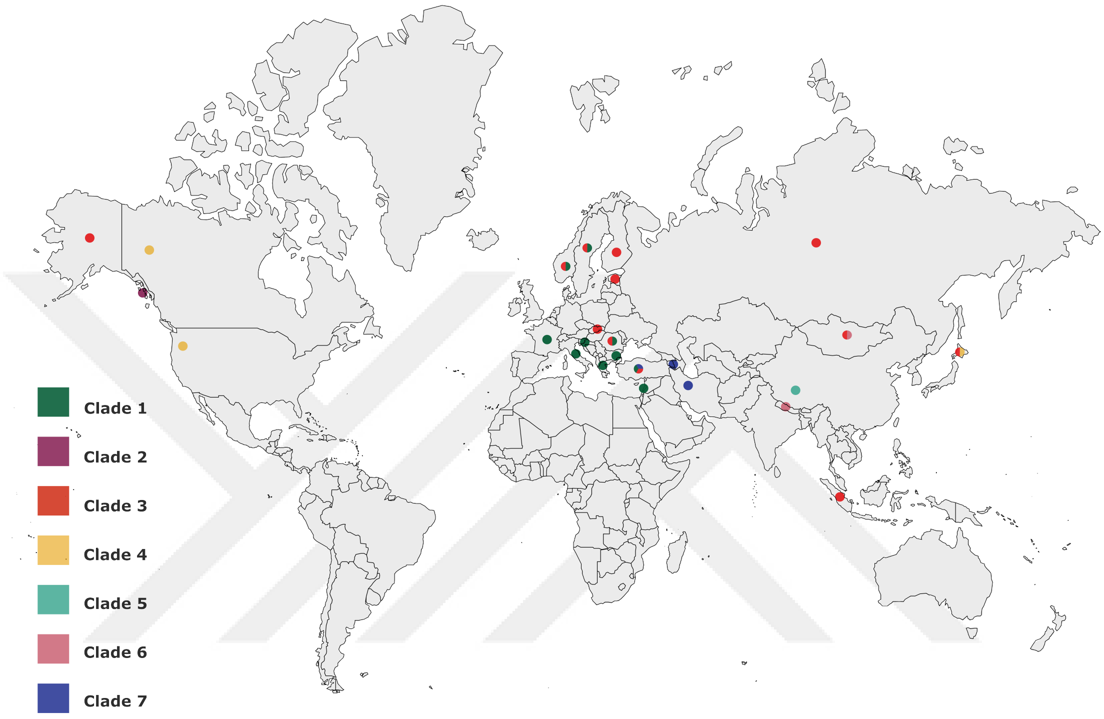


Figure 1.4: The haplogroup distribution of brown bear populations around the world

The studies of Taberlet and Bouvet (1994) and Kohn et al.(1995) revealed significant genetic differences within European groupings and a clear distinction between western and eastern populations (Matsuhashi et al., 2001). The studies of Taberlet and Bouvet on mtDNA control region indicated that there are two main lineages in European brown bears that are eastern (clade 3a) and western (clade 1). The eastern lineage spreads over the Eurasian Continent, from the northeastern region of Europe to the Far Eastern region of Russia (Korsten et al., 2009; Murtskhvaladze et al., 2010; Saarma et al., 2007; Tammeleht et al., 2010). The Western lineage is divided into two groups: Iberian (clade 1a) and Balkan/Italian (clade 1b) (Kohn et al., 1995; Taberlet & Bouvet, 1994).

The phylogenetic tree of brown bears in Türkiye was also studied, and the mitochondrial DNA control region was used in a recent study by Çilingir et al. (2016). Both analyses using the the median-joining network and Bayesian phylogenetic tree demonstrate that Anatolian brown bear haplotypes are subdivided into three separate maternal lineages: Clade 1, Clade 3, and Clade 7. These are further subdivided into five different subclades: Middle East-Iran 7a, Middle East-Türkiye 7a, Middle East-divergent 7b, Taurus-Lebanon 1d, Italy-Balkan, Western Anatolia 1b. Subclade 7a was constituted of all Iranian and Turkish specimens, including those from the Artvin region. Two specimens were highly distinct from the rest of the Turkish subclade 7a members, with a posterior probability value of 0.95. They defined these samples as a new subclade named clade 7b. One originated from Özgüven, Artvin, in Northeastern Türkiye, while the other was not linked to any major lineage in the previous publication. Clade 1b consists of the Middle East, Eastern Europe and Lebanon regions. Two samples from western and southwestern Türkiye revealed different haplotypes from subclade 1b in Çilingir's study. These samples are linked with clade 1d occuring in ancient samples from Lebanon published by Calvignac et al. (2009) and Çilingir et al. (2016) has demonstrated that subclade 1d is not extinct.

On the other hand, the studies indicated that Iranian populations are highly different from bears seen in nearby regions containing Turkish and Lebanese populations (Calvignac et al., 2009; Talbot & Shields, 1996). According to Ashrafzadeh et al. (2016)'s study on the Iranian brown bears, the mtDNA of Iranian populations formed a lineage that was situated between the western and eastern lineages. Çilingir et al. (2016) demonstrated that clade 7, known as the Iran clade, is also found in Türkiye. We will reassess the situation of Iran using our five mtDNA and RAD-seq samples taken from tissues.

1.6 The divergence times of brown bear

A variety of factors are assumed to be involved in the emergence of genetic variations that result in the formation of haplogroups. Past climatic changes are expected to have a major impact on brown bear population size and the distribution of intraspecific variation in brown bears (Taberlet et al., 1998). The majority of populations had to live in refugia during the Pleistocene glaciations. Many species started to extend their ranges northward after the end of the last glacial maximum (LGM), approximately 18.000 years ago (bp). It is estimated that migration to refugia during glacial eras and outward migration during interglacial periods led to changes in animal genomes (Saarma et al., 2007). Phylogenetic studies of brown bears using mitochondrial markers show signs of migration associated with glacial periods (Molodtseva et al., 2022).

The geographic distribution of brown bears in Europe and their mtDNA haplotypes were found to be strikingly consistent. The Eastern and Western lineages of the brown bear populations in Europe are separated by the mtDNA control region sequences into two distinct clades. While the Western lineage is made up of small, fragmented, and threatened populations, the Eastern lineage consists mainly of large and robust populations. Western and eastern bears, respectively, seemed to have colonized the majority of Europe from Iberian and Caucasian/Carpathian refuges, according to Taberlet et al. (1998) and Hewitt (2000). Following the melting of the Scandinavian ice cap, both expansions collided in central Sweden and created a hybrid zone (Taberlet et al. 1995). However, it is still unclear when the eastern spread began as Hewitt (2000) noted. Although the brown bear is a model species for studying Pleistocene and Holocene displacements, the genetic background of the Eastern lineage is poorly understood (Saarma et al., 2007).

There are other assumptions regarding the distribution of brown bears. It is also assumed to have an impact that factors such as bottlenecks in northern populations led to decreasing genetic diversity (Hewitt 1996; 2000), stochastic processes occurred due to rapidly growing human activity in Holocene and geographical boundaries and female philopatry. Geographical barriers, such as mountains, rivers, cities, and straits, can restrict gene flow between populations, which results in genetic differences. For instance, Mizumachi et al. (2020) suggested that the Bosporus and Dardanelles straits, which are situated between the Balkans and Turkey, were likely the biogeographical border between clades 1b and 7 of brown bears in the post-Pleistocene. On the other hand, de Jong et al. (2023) claimed that divergence times of mtDNA haplogroups occurred prior to LGM glacial refugia. They stated that the genetic variations in mtDNA of brown bears are the result of random differential fixation of ancestral alleles influenced by geographical boundaries and female philopatry.

1.7 Significance of This Study

The brown bear populations in Türkiye are significant in terms of evolutionary diversification due to the geographical location of Anatolia (Ambarli et al., 2016). Türkiye populations have a wide variety of haplotypes, including Eurasian, Middle Eastern, Western European, and a recently discovered Middle Eastern clade, according to previous studies using mitochondrial DNA (Calvignac et al., 2009; Çilingir et al., 2016). However, extensive sampling is needed to better understand the phylogeny of brown bears in Türkiye since past studies had small sample sizes, covered limited distributions, and only used mtDNA that is maternally inherited. The aim of this thesis is to investigate the evolutionary history and the genetic diversity of Türkiye's brown bear populations. The study covers every region of Türkiye where bears live, as well as some regions of Iran. We utilized 91 scat samples from Türkiye and 5 tissue samples from Iran for phylogenetic analysis with mitochondrial D-loop control sequences. We used 10 tissue samples, 5 from Türkiye and 5 from Iran for phylogenetic analysis of restriction site-associated DNA (RAD) sequences.

We analyzed and compared phylogenetic trees using two different markers. Because the mtDNA sequence only offers information on female-mediated gene flow, an assessment of genetic diversity and phylogeny using RAD-seq data provided a different point of view. We assume to gain a better understanding of the phylogenetic patterns of Middle Eastern brown bears with comparative analysis and comprehensive sampling.

Chapter 2

METHODS

2.1 Sampling

We collected scat samples from all regions in Türkiye where brown bear lives by covering the Eastern Anatolia, Southeastern Anatolia, Central Anatolia, Black Sea, Marmara and Mediterranean Regions. The scat samples were collected by recording coordinates, altitude, scat content, and freshness using ArcGIS Survey123. More than 600 samples were collected as a result of extensive fieldwork. Figure 2.1 demonstrates the locations of samples chosen for amplification of the D-loop control region. In addition, tissue samples from Bolu, Ankara, Tunceli and Hakkari were obtained from the Biology Department of Ankara University. In addition, DNA extracted from blood samples collected by KuzeyDoğa Association from Sarıkamış, Kars region were used for this study. Phylogenetic analysis based on RAD sequences were constructed using five tissue samples from Iran, and one blood and four tissue samples from Türkiye. Figure 2.2 shows the locations of these samples.

Five Iranian tissue samples were gathered from accident-related mortalities such as vehicle collisions, and all laboratory procedures were conducted at the Faculty of Agriculture and Environment's biodiversity lab, Arak University.

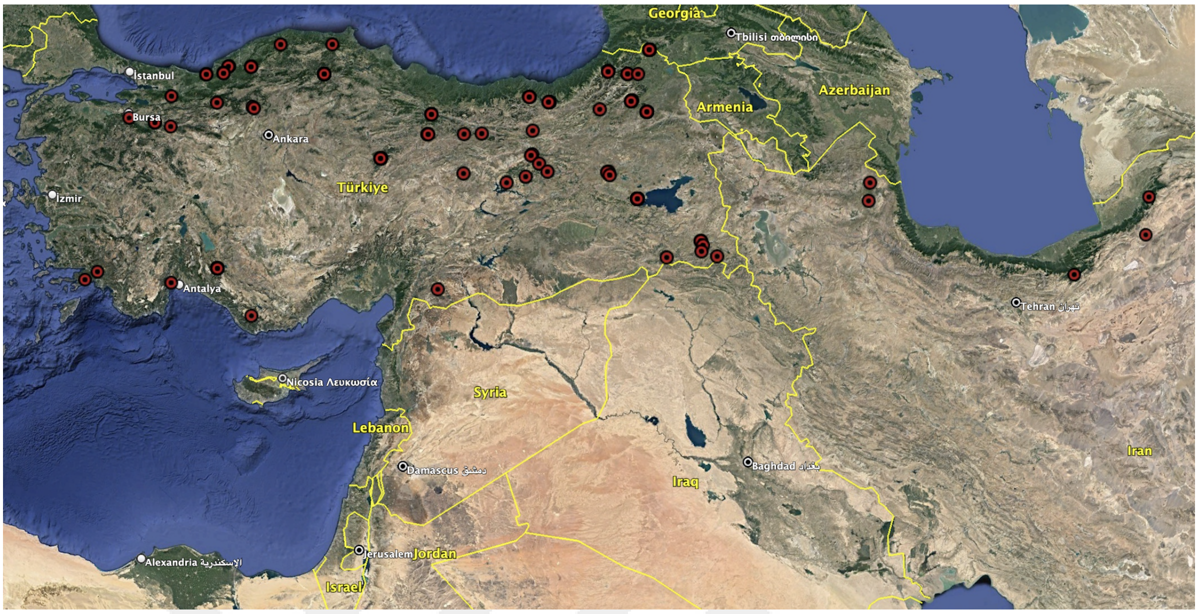


Figure 2.1: The sampling localities of brown bear scat samples for mtDNA D-loop control region

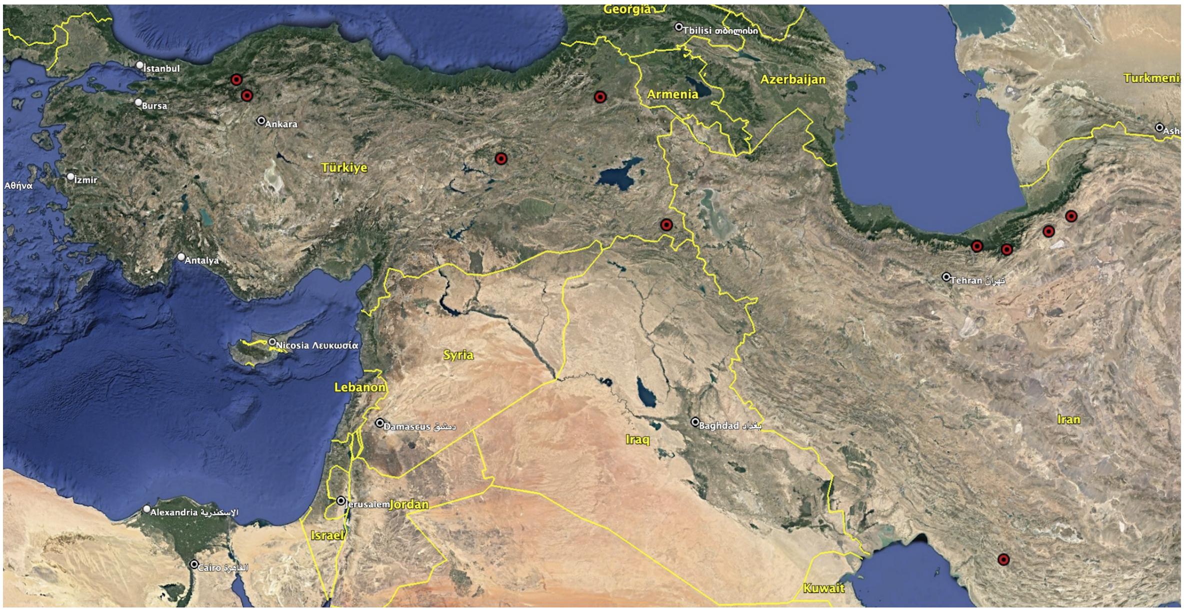


Figure 2.2: The sampling localities of brown bear scat samples for RAD-sequences

2.2 DNA extraction

DNA extraction was performed at the Biodiversity and Evolutionary Genetics Laboratory (KUBEGL) using Qiagen QIAmp DNA Stool Mini Kit (Qiagen, Germany). All of the processing was conducted with sterile equipment. We took a sample of the outer surface of the scat with a scalpel blade because the outer layer of scat includes intestinal epithelial cells (Stenglein et al., 2010; Wehausen et al., 2004). The maximum amount of stool was added into 15 ml of the tube, and then the tube was filled with distilled water (dH20) and vortexed. The tubes were left overnight on a shaker in a 4°C room. On the second day, samples were centrifuged at 3000 rpm for 2 minutes, and the supernatant was taken into new 15 ml tubes. These tubes were centrifuged at 4750 rpm for 30 minutes, and the supernatant was discarded. Then, inhibitEx buffer was added to each stool sample of a 15 ml tube and vortexed until the samples were thoroughly homogenized. The samples are again centrifuged at 4750 rpm for 3 minutes to pellet the samples. 25 ul of proteinase K was added into new 2 ml of tubes. 750 ul of supernatant was added into the same tube. Then, 600 ul of AL buffer was added, and the tubes were vortexed for 15 seconds. The samples were incubated at 60 °C overnight. Last day, 600 ul of absolute ethanol was added into the lysate and mixed. Then, 700 ul of the samples were added into the spin column and centrifuged at 14000 rpm for 1 min. Until all lysate had been loaded, this step was repeated. 500 ul of AW1 buffer was added to the spin column. 500 ul of AW2 buffer was added at 14000 rpm for 3 min. Then, the empty tube was again centrifuged for 3 min. In the end, the spin column was transferred to a new 1.5 ml tube, and 55 ul of ATE buffer was added directly onto the middle of the membrane. The tube was incubated at 5 min at room temperature, then centrifuged at 14000 rpm for 1 minute to elute DNA, and this step was repeated. Eluted DNA was preserved at -20 °C.

2.3 Quantitative polymerase chain reaction (qPCR)

DNA samples were amplified with G10P primer, which is a species-specific primer, to clarify that the samples contain bear DNA. This primer may effectively distinguish brown bear samples from other animal species (Paetkau & Strobeck, 1994). PCR master mix was prepared with 8 μl of Syber, 10 μl of H_2O, and 1 μl G10P primer for each sample, and PCR reactions were carried out in 20 μl volumes containing 16 μl of master mix and 4 μl of DNA. Thermal cycling conditions are indicated in Table 2.1. The quantitative polymerase chain reaction (qPCR) utilizes fluorescence produced by a fluorochrome attached to double-stranded DNA. A fluorophore is released when the probe is digested during elongation, or fluorochrome attached to a probe may fluoresce when it binds to the target throughout DNA synthesis. The Cq (quantification cycle) value is important while evaluating qPCR results, and it means that the number of PCR cycles required for the fluorescence to reach a quantification threshold (A. Ruiz-Villalba, 2021). We considered the samples with a good signal and a Cq value between 20 and 25 to be successful and preserved them for next experiments.

Table 2.1: Quantitative real-time PCR (qPCR) conditions

Step	Temperature	Time	Number of cycles
Initial denaturation	93 °C	3 min	1
Denaturation	95 °C	10 min	39
Annealing	58 °C	10 min	
Extension	72 °C	20 min	
Final extension	65 °C	5 min	1

2.4 DNA quantification

The Qubit 2.0 Fluorometer and Qubit dsDNA BR Assay Kit were used for DNA measurement of extracted scat samples. The standards were prepared, and first introduced to the machine for calibration. The sample was then placed into the sample chamber. The volume of the sample was set to 5 μl and the unit was set to ng/μl. Then, the 5 μl of sample and 195 μl of the working solution were added into each tube. The working solution was prepared by using 199 μl of BR buffer and 1 μl of BR reagent for each sample. Standard was prepared by using 190 μl of working solution and 10 μl of standards.

2.5 qPCR for D-loop amplification and agarose gel electrophoresis

We utilized the forward primer L15995 ('5-CTCCACTATCAGCACCCAAAG-3') and reverse primer L15995 ('5-GGAGCGAGAAGAGGTACACGT-3') for the amplification of the hypervariable segment of the mitochondrial DNA control region. According to the entire genome of *Ursus arctos* mitochondrion, L15995 forward primer binds to mtDNA sequences 16532-16553 and 16858-16878 (AF303110.1). 346 bp is the length of the PCR product. The qPCR conditions in section 2.3 were applied using primer L15995. We carried out qPCR by beginning with samples containing high Qubit values and kept samples with a good signal for running them in an agarose gel.

%2 agarose gel was prepared and run in 1X TAE (Tris, Acetic Acid, EDTA) buffer. 5 µl of the IZW safe stain was added to 100 µl of TAE. The samples were then loaded by mixing with 6X Loading Dye (Biomatik) and run at 100 volts for 30 minutes. Following the end of the electrophoresis, the bands were visualized and photographed with the VWR Gel Imager2. We expected to visualize the band between 300-400 bp with good quality. One of the gel images is shown in Figure 2.2. The successfully amplified samples were stored.

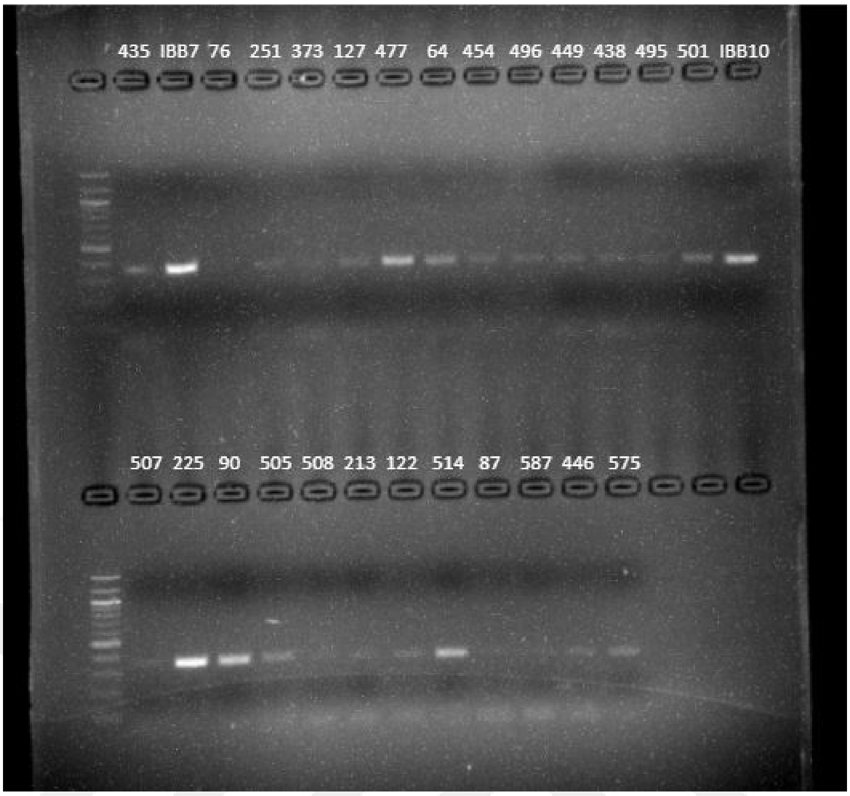


Figure 2.3: Agarose gel electrophoresis of qPCR products amplified with primer L15995 for 346 bp segment of the D-loop control region

2.6 PCR Clean-up

We performed PCR clean-up to remove primer dimers and SYBR dye using Promega Wizard® SV Gel and PCR Clean-Up kit. First, we mixed 30 ul of membrane binding solution with the qPCR product. The mixture was added into the collection tube containing SV minicolumn. It was incubated at room temperature for 1 minute then, it was centrifuged at 13.000 rpm for 1 minute. The liquid at the bottom of the tube was removed, and the minicolumn was reinserted. The column was then filled with a 700 ul membrane wash solution. The flowthrough was discarded after 1 minute of centrifugation at 13.000 rpm. The step was repeated with 500 ul of membrane wash solution and centrifuged for 5 minutes at the same speed. Finally, the collection tube was emptied and centrifuged for 1 minute to eliminate any remaining ethanol. The minicolumn containing cleaned DNA was inserted into the new centrifuge tube. 25 ul of nuclease-free water was added into the minicolumn, then waited for 3 minutes. This step was repeated. Then, it was centrifuged at 13.000 rpm for 1 minute for elution. The minicolumn was discarded, and the DNA was stored at 4 °C.

2.7 Phylogenetic analysis of mtDNA D-loop region sequences

The D-loop control region of scat samples was sequenced by Sanger Sequencing. BioEdit 7.2.5 was used to edit and trim the sequences (T. Hall, 2013). Following the assembly of Sanger sequencing data, all sequences were merged with D-loop control region sequences from NCBI GenBank, which are listed in the Supplementary tables A.4-A.5. Then we aligned all of the mtDNA control region sequences and obtained aligned sequences with a length of 269 bp. We eliminated sequences less than this length. The FASTA formatted sequences were analyzed using the DNASp 6.12 (Rozas et al., 2003) to determine haplotypes. Haplotype diversity was calculated by comparing the SNPs within each sample. Then, the median-joining network and haplotype map were constructed using PopART to understand the haplotypes in Türkiye and the relationships of Türkiye haplotypes with other sequences.

We performed maximum likelihood (ML), maximum parsimony (MP), and Bayesian Inference (BI) trees to investigate the evolutionary relationships of Turkish brown bears with different statistical methods. The outgroup was determined as Ursus Maritimus. The maximum likelihood tree was constructed by RaxML 8.0 (Stamatakis, 2006). ML trees calculate the probability of seeing the data we observe given the tree we assume and choose the tree with the highest probability. The model selection analysis was performed by MEGA and gave K2+G, shown in Table ??. According to this model, GTRGAMMA was chosen from the RaxML manual. The maximum parsimony tree was constructed using PAUP* (Swofford et al., 2002). Parsimony assumes that evolutionary events are uncommon and therefore chooses the reconstruction with the fewest postulated events, which are the most accurate. Last, The Bayesian tree was constructed using BEAST v2.7.4. This model computes the probability of getting the tree we assume given the data we have and

Table 2.2: Substitution model selection by MEGA

Model	#Param	BIC	AICc	lnL	Invariant
K2+G	165	3849.10075	2533.69974	-1100.581051	n/a
T92+G	166	3850.035388	2526.677749	-1096.054616	n/a
T92+G+I	167	3853.390731	2522.076653	-1092.738535	0.5453208959
K2+G+I	166	3855.377605	2532.019967	-1098.725725	0.5381769701
HKY+G	168	3863.344823	2524.074492	-1092.721827	n/a
HKY+G+I	169	3867.301971	2520.075576	-1089.706648	0.5579987473
TN93+G	169	3873.29566	2526.069264	-1092.703492	n/a
TN93+G+I	170	3877.670588	2522.488316	-1089.897202	0.533200921
GTR+G	172	3895.309504	2524.216046	-1088.729154	n/a

chooses the tree with the highest probability. The Hasegawa-Kishino-Yano (HKY) 1985 Substitution Model was used for this analysis, Gamma category count was arranged as 4. Accordingly, we used a coalescent constant population size tree and a strict clock model. We conducted four independent analyses for every 5000 generations from 50,000,000 generations, and 20 percent of each analysis were discarded. MCMC trace files generated from Bayesian analysis were examined with Tracer v1.7.2, demonstrated in Figure A.3. We also performed Bayesian MCMC-based divergence using BEAST v2.7.4 under the HKY substitution model by using the same parameter with Bayesian tree. We calibrated the time according to the divergence time of brown bears and modern polar bears. All analyses were merged with LogCombiner and summarized with TreeAnnotator (Bouckaert et al., 2014). Bootstrapping was performed 1000 times in all three analyses to determine and construct confidence intervals. The tree with the highest criterion score is displayed, as are the bootstrapping values for each node. We visualized summarized trees in FigTree 1.4.4(Rambaut, 2018).

2.8 Phylogenetic analysis of Restriction-site Associated DNA (RAD) sequences

We use RAD-sequencing tissue samples from Iran and Türkiye regions for phylogenetic analysis, shown in Table 2.3. We aligned whole genome sequences obtained from GenBank (Table A.3) and RAD-seq samples from Iran and Türkiye with the reference genome of Ursus arctos (GCF023065955.1) with the BWA-MEM algorithm (Li & Durbin, 2009). We converted aligned reads to bam format, filtered them for proper pairs, and eliminated PCR duplicates using samtools 0.1.19 (Li et al., 2009).

Table 2.3: RAD sequences from Türkiye and Iran regions

Sample ID	Type	City	Region
IBB4	Tissue	Shiraz	Iran
IBB5	Tissue	Shahroud	Iran
IBB9	Tissue	Semnan	Iran
IBB15	Tissue	Shahroud	Iran
IBB225	Tissue	Amol	Iran
BTR010	Blood	Kars	Türkiye
BTR221	Tissue	Hakkari	Türkiye
BTR226	Tissue	Bolu	Türkiye
BTR227	Tissue	Ankara	Türkiye
BTR422	Tissue	Tunceli	Türkiye

We selected common loci found in all populations with a sequence length between 200-500 bp by considering parsimony informative sites. We started by filtering our Türkiye and Iran bam files with minimum mapping quality of 20 (minMapQ 20), ignoring bases with quality less than 20 (minBase 20), minimum depth of 4 (minDepth 4), and maximum depth of 100 (maxDepth 100) using samtools (Danecek et al., 2021). The depth intervals were then computed using mosdepth (Pedersen & Quinlan, 2018). We merged contig intervals and converted bam to bed with bedtools with a maximum 200bp length of gap between loci intervals to merge them into a single locus interval. Then, we convert bam to fasta by filtering with a minimum probability of variable sites of 0.2 and min probability of parsimony informative sites of 0.3 to obtain the top 1000 locus by AMAS (Borowiec, 2016). Last, we aligned sequences from individuals for each locus using clustalo (Sievers et al., 2011). Then, we randomly selected 100 of them. We extracted the same loci from WGS sequences obtained in GenBank and aligned them with our populations with clustalo(Sievers et al., 2011). We filter this data by eliminating the sequences shorter than 300 bp and containing more than 50 N. We continue the analysis with the remaining 91 loci. We also eliminated Slovenia and Maritimus samples because we could not obtain a reasonable tree with a high posterior probability. We loaded our 91 most informative parsimony loci into StarBeast3 template in Beast2 and used the default parameters. We applied the MCMC analysis to estimate the posterior probabilities of the phylogenetic tree with 10,000,000 iterations and visualize it with FigTree v1.4.4 (Rambaut, 2018). These loci were used to create a maximum likelihood tree as well. We created 1000 bootstrap trees for each locus with RAxML and used bootstrap trees to construct a coalescent species tree with ASTRAL 5.7.1 (Zhang et al., 2018).

We carried out principal component analysis for our tissue and blood samples. We did genotype calling for aligned sequences using ANGSD for 15 samples, 5 from Iran and 10 from Türkiye. We computed allele frequencies (-doMaf 1) and genotype likelihood (-GL 1). We ran an analysis to determine the sites found in at least 10 individuals (-minInd 10). Therefore, we acquired a total of 246.033 SNPs. We estimated the between-individual covariance matrix using genotype probability using PCangsd in Python. Then we visualized our analysis in R.

Chapter 3

RESULTS AND DISCUSSION

3.1 Phylogenetic analysis of mitochondrial DNA

3.1.1 Haplotype diversity

To understand the relationships between individuals and populations, we performed a haplotype diversity analysis by using the hypervariable D-loop domain of the mitochondrial DNA control region. The haplotype data analysis revealed 18 haplotypes in Türkiye brown bears, as shown in Figure 3.1; further haplotype information is available in Table A.1 and Table A.2. A total of 33 mutations were calculated in Turkish populations with a haplotype diversity (Hd) of 0.8806 by DnaSP 6 (Rozas et al., 2017).

When the haplotype network in Figure 3.1 was assessed, the variations between haplogroups could be seen. Clade 7 was formed by five haplotypes. In this haplogroup, Haplotype 1 is highly divergent from other haplotypes with 7 mutational differences and is distributed in the eastern part of Türkiye. Other haplotypes of this clade have one SNP difference. Haplotype 11 is found in Artvin and Ardahan regions, whereas Haplotype 15 is found in Yozgat and Sinop regions. Haplotype 14 is found in Sivas and Yozgat. Haplotype 4 is only found in Antalya. In summary, clade 7 is widespread, encompassing the Mediterranean, Black Sea, Eastern, and Central Anatolian regions.

Clade 1 is composed of six haplotypes which different from Clade 7 in eight SNPs. Subclade 1d clade is highly divergent from clade 1b with 6 SNP differences and contains haplotype 5 and haplotype 7 covering Corum, Kastamonu, and Mersin. Haplotype 5 is only found in Mersin, whereas haplotype 7 is found in Çorum and Kastamonu. Subclade 1b clade comprises Muğla, Düzce, Zonguldak, Bilecik, and Bursa.

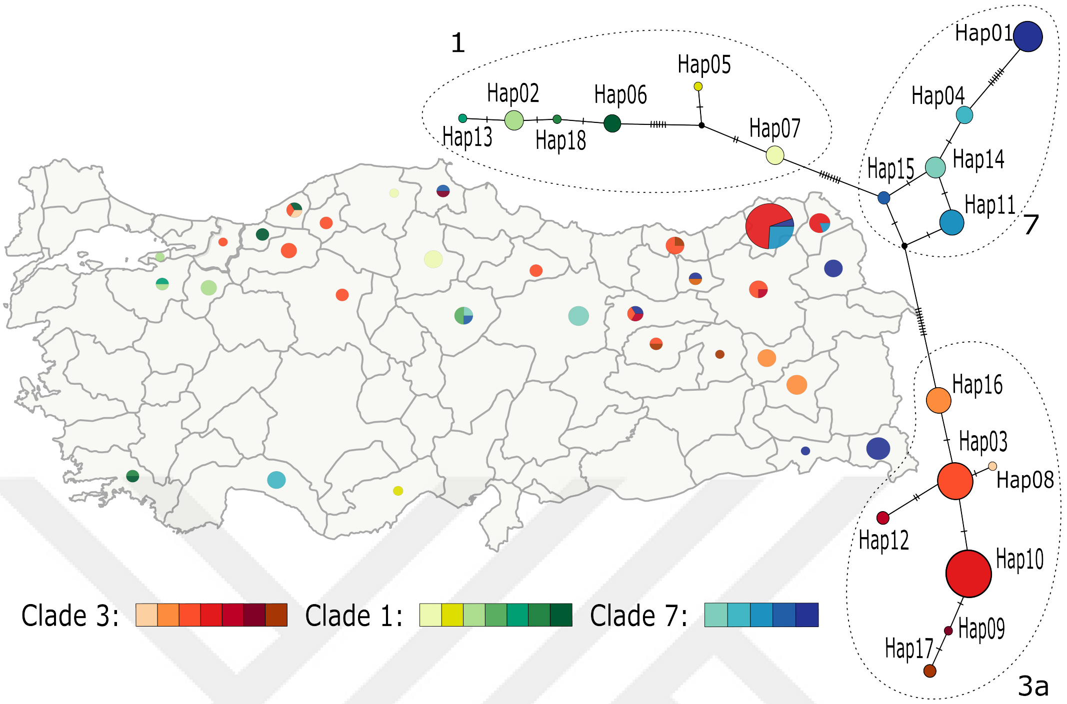


Figure 3.1: Median-joining haplotype network of Türkiye brown bears placed on a map by PopART. The diameter of the circle is proportional to the haplotype frequency in this study.

Clade 1 is distributed in the Western part of Türkiye and is not found in the Eastern part of Türkiye. Another haplogroup called Clade 3a has the largest number of individuals with 7 distinct haplotypes that have 10 SNP differences from Clade 7. Clade 3a is widely distributed across Türkiye, mostly in the eastern region, except for Marmara and the Mediterranean regions. There are also regions where haplotypes overlapped. Two main lineages 7 and 3, overlapped in Ardahan, Artvin, Erzincan, and Bayburt regions. Lineages 1 and 3a overlapped in Düzce, whereas clades 7 and 1 overlapped in Yozgat. In conclusion, the result of haplotype data analysis has shown that Türkiye populations have high haplotype diversity with 18 different haplotypes and are clustered into two major lineages, Eastern and Western clades.

3.1.2 Median joining network

We constructed a haplotype network to analyze and visualize connections between Türkiye and other world populations. We performed the median-joining network in PopART (Population Analysis with Reticulate Trees) (Bandlet et al., 1999) with the haplotypes determined by the mutation differences in mtDNA control region, shown in Figure 3.2.

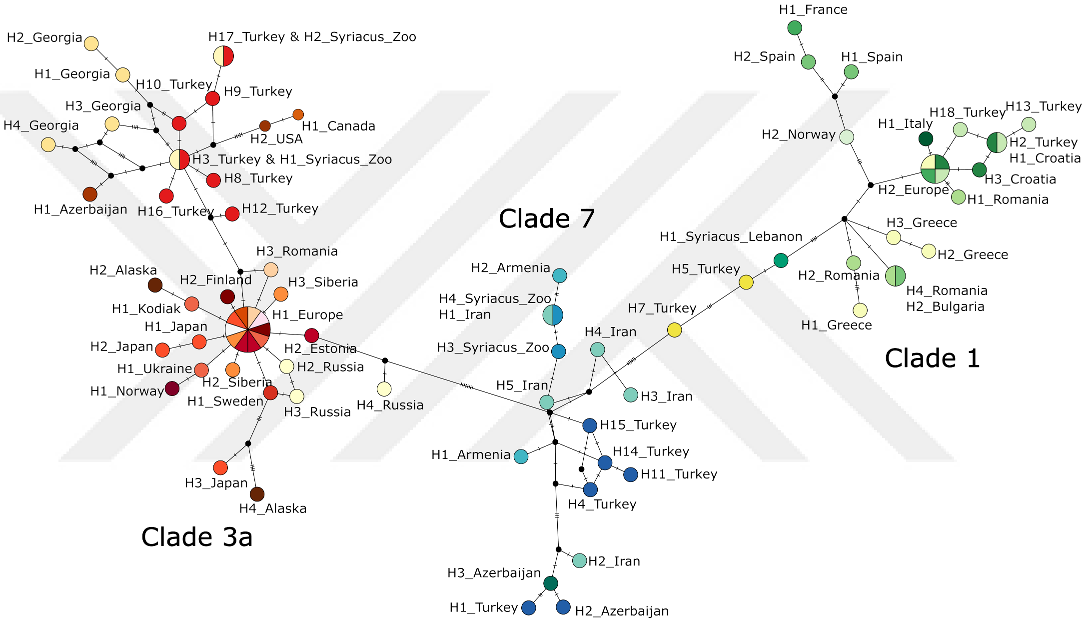


Figure 3.2: Median-joining network analysis of Türkiye haplotypes by PopART. Each line represents a single mutational change.

The results showed that the haplotypes in Türkiye are grouped into three distinct haplogroups containing clade 3, clade 1, and clade 7. Clade 1 comprises three groups: Iberian lineage (1a), Balkan, Italian lineage (1b) (Kohn et al., 1995; Mizumachi et al., 2020; Taberlet & Bouvet, 1994) and Taurus-Levant lineage (1d) (Calvignac et al., 2009; Çilingir et al., 2016). Haplotypes 5 and 7 were clustered with Lebanon and formed 1d and were seperated with clade 1 populations with 4 SNP differences. Haplotypes 5 and 7 in clade 1d were significantly different from each other with 3 SNPs. Norway, Spain, and France formed clade 1a and were differentiated from other clade 1 groups by three SNP differences; this subclade does not exist in Türkiye. Other Türkiye haplotypes in this haplogroup clustered with European populations including Croatia, Greece, Bulgaria, and Romania and formed the clade 1b.

Clade 7, referred to as the Middle East clade, consists of haplotypes from Armenia, Iran, Azerbaijan and Türkiye with two *Syriacus* haplotypes. All Iranian haplotypes match with this clade. Haplotype 1 from Türkiye, Haplotype 2 and 3 from Azerbaijan, and Haplotype 2 from Iran clustered together and split from other clade 7 haplotypes.

Clade 3a, including Eastern European, Western Alaska, and Japan (Matsuhashi et al., 2001), was grouped with Türkiye, Georgia, and Azerbaijan samples. We noticed that Middle Eastern and North American populations including Canada and the United States formed a separate cluster than other clade 3 populations. These American samples matched with Clade 4 (Grizzly bears). It shows that Grizzly bears are closely related to Middle eastern brown bears.

Furthermore, *Syriacus* samples appeared to be distributed into three clades. Clade 3a has a *Syriacus* Zoo sample identical to Haplotype 3 from Türkiye. Haplotype 4 of the *Syriacus* Zoo sample matches with clade 7, and the Lebanon *Syriacus* sample matches with clade 1. Therefore, this figure supports that brown bear populations in Türkiye are divided into three haplogroups, and there was no distinct cluster formed by *Syriacus* specimens.

3.1.3 Phylogenetic trees with three statistical methods

We constructed phylogenetic trees using three different methods to combine the results by increasing confidence to better understand relationships in species (Figure 3.3). We performed three statistical methods for phylogenetic analysis, namely maximum parsimony, maximum likelihood, and Bayesian Inference. The outgroup was determined as *Ursus maritimus*. The results of the tree models were generally consistent with each other and the median-joining network. Maximum parsimony and Bayesian inference trees estimate high support values, and the maximum likelihood has relatively low support values. Three phylogenetic trees all have the same individuals within each clade. While clade 7 and 3a split from the same node in the Bayesian tree with the posterior probability value of 0.59, clade 7 and 1 split from the same node with the posterior probability value of 0.22 in the Maximum likelihood tree and and posterior probability value of 1 in maximum parsimony tree.

Bears of Western Europe, the western region of Türkiye, and Lebanon were classified as Clade 1. Türkiye, which contains two subclades of 1, appears to provide a connection between European (1b) and Levant (1d) populations. Clade 7 consists of Azerbaijan, Armenia, Türkiye, and Iran regions. It includes five Türkiye haplotypes containing Kars, Ardahan, Erzincan, and Hakkari (Eastern region), Bayburt, Artvin, and Sinop (Black sea region), Şırnak (Southeastern region), Sivas and Yozgat (Central Anatolia region) and Antalya (Mediterranean region). Clade 7 was not found in Georgia, although this clade is present near its borders with Türkiye and Armenia. Clade 3a contains Eastern Europe, Western Alaska, Central Hokkaido, and Türkiye. Iran samples containing nine samples were classified into five haplogroups and all haplotypes were matched with clade 7. Therefore, ML, BI, and MP analyses resulted in similar tree topologies, and gave consistent results with median joining network.

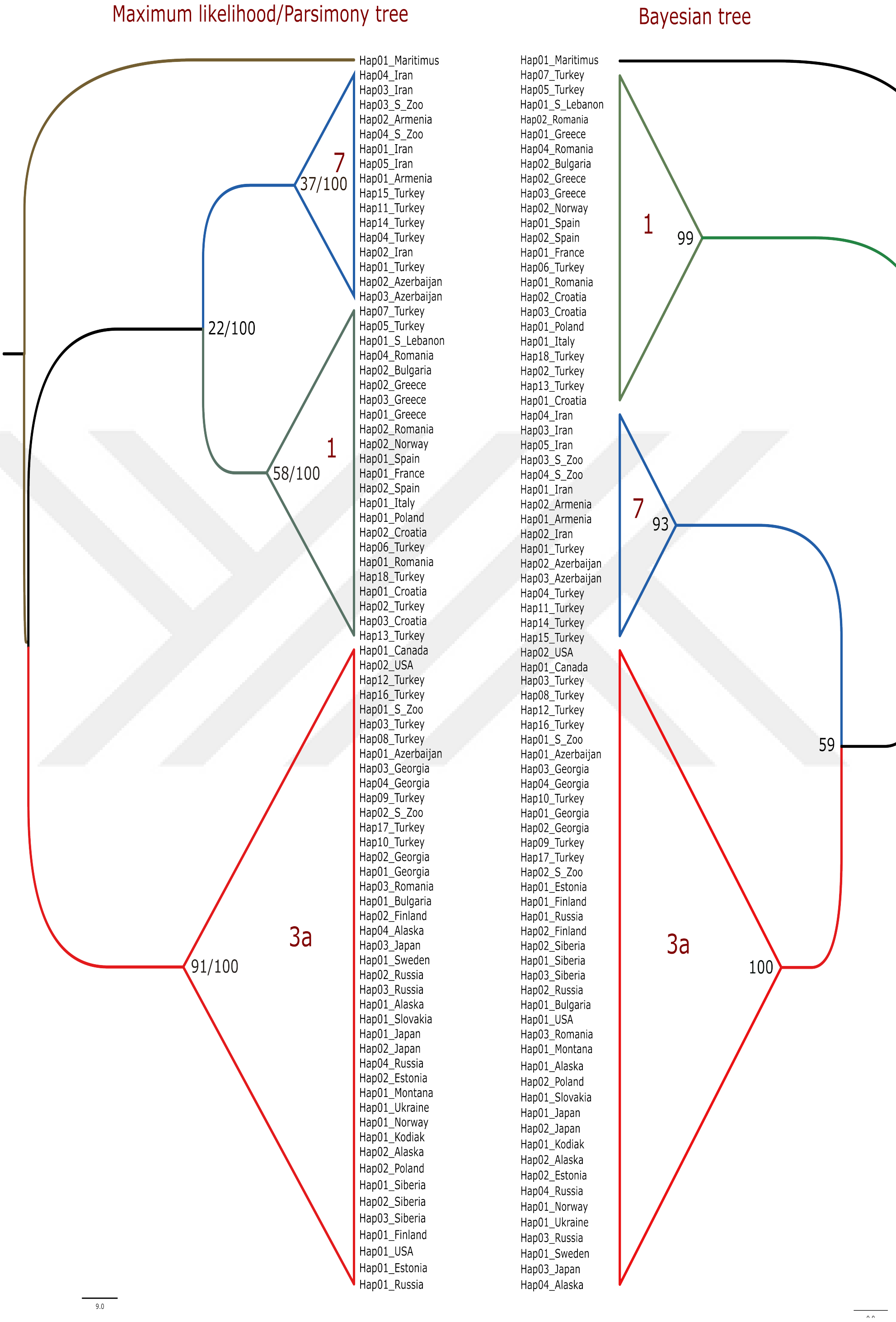


Figure 3.3: Phylogenetic trees (Maximum likelihood, Maximum Parsimony and Bayesian Inference) based on mtDNA D- loop sequences. Support branch values for ML and MP bootstraps and BI posterior probabilities are represented by numbers at the nodes, respectively.

3.1.4 Divergence time tree

We constructed a divergence time tree to better understand how and when clades diverge in time scale. The divergence time between all brown bear populations and modern polar bears was estimated to be 264 kya by Lan et al. (2022) and 263 kya (162-400) by Davison et al. (2011). We built our divergence time tree based on these estimates (3.5).

Our Bayesian estimates demonstrated that clade 3 diverged into a new group around 130.500 years ago. It is slightly closer to the estimate of Hirata et al., 2013 of 165 kya (63–292) and the estimate of Davison et al., 2011 of 92 kya (51–133). Furthermore, the divergence tree demonstrated that the Middle eastern populations comprising Georgia, Türkiye and *Syriacus* samples created a subgroup by clustering separately from other clade 3 populations. Time to most recent common ancestor (TMRCA) of other Eurasian haplogroups, including clades 7 and 1, was about 186 800 years ago according to our Bayesian estimates. When Clade 1 was assessed, TMRCA of Clade 1ab and Clade 1d was calculated to be 105900 years ago. Davison et al. (2011) estimated it to be 100 kya (49–164). This estimate is generally consistent with the previous studies (Calvignac et al., 2009; Hirata et al., 2013). Clade 1d diverged 35.800 years ago which is earlier than previous estimates (Çilingir et al., 2016).

According to Çilingir et al. (2016), clade 7b was formed by an Artvin sample (KT438638) and one unknown sample. Our findings supported the division of Clade 7 into subclades. Artvin sample (KT438638) clustered with Iran, Azerbaijan and haplotype 1 of Türkiye samples, including from the eastern region of Türkiye with Kars, Bayburt, Şırnak, Erzincan and Hakkari. Other subclade 7a was widespread with Armenia, Iran, *Syriacus* and Türkiye samples containing Antalya, Sivas, Yozgat, Artvin, Ardahan and Sinop. The distribution of subclades of clade 7 was illustrated in Figure 3.4. The time of most common ancestor of clade 7 subclades was estimated as 104.600 years ago, whereas the previous estimate was about 50 kya. We found the split time of subclade 7a as 51.800, in contrast to the previous estimate of 21 kya (Çilingir et al., 2016).

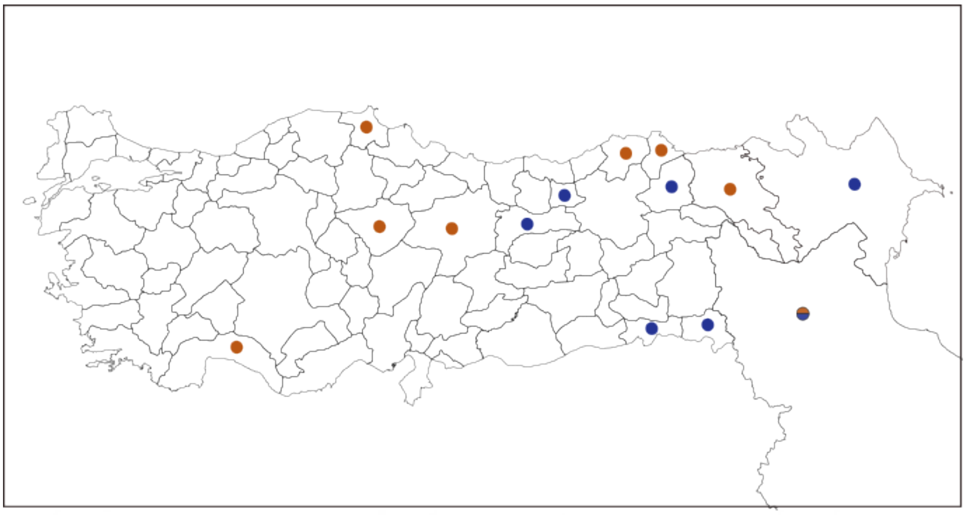


Figure 3.4: Distribution of brown bear subclades belonging to Clade 7. Blue dots represent clade 7b, orange dots represent clade 7a.

The estimations about divergence times of brown bear haplogroups were associated with the glacial periods of the Pleistocene. As LGM period was between 26,000–19,000 years ago, our findings of divergence time did not match with this hypothesis. Divergence occured before LGM period as de Jong et al. (2023) claimed. We agree with the statement of de Jong et al. (2023) that the genetic variations of mtDNA is not the consequence of glacial periods, but rather is the result of random differential fixation of ancestral alleles influenced by geographical boundaries and female philopatry.

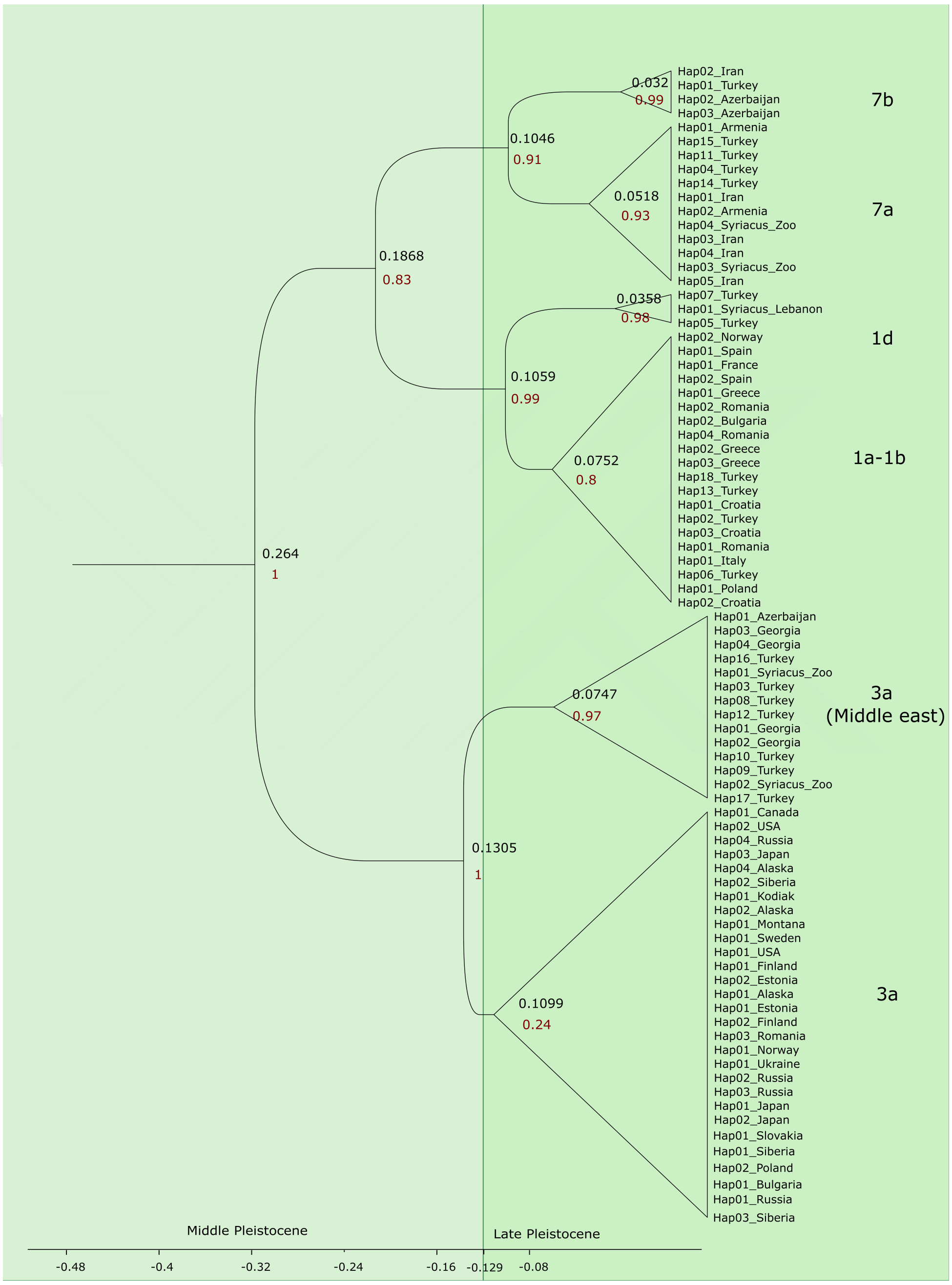


Figure 3.5: Divergence time estimation of brown bears using Beast2. Numbers colored with black represent the divergence times, and numbers colored with red represent the posterior probabilities.

3.2 Phylogenetic analysis for RAD-seq data

3.2.1 Maximum-likelihood species-tree

We constructed a maximum likelihood phylogenetic tree using Parsimony informative sites to assess the evolutionary relationships between populations using RAD-seq data. Figure 3.6 indicates that Iran and Türkiye brown bears may be each other's closest relatives based on evolutionary lineage. They are followed by brown bears in Georgia, then Japan and Russia of the Eastern Palearctic regions. On the other hand, Italy and Slovakia bears are each other's closest relatives followed by the Slovenia and Sweden populations, with European populations clustering together. These findings support our mtDNA phylogeny findings and differentiation between Eastern and Western lineages. Bears in eastern regions containing Japan, Russia, Türkiye, Georgia and Iran are distinct from bears in western regions containing European countries. Furthermore, the ASTRAL tree revealed that the populations of Türkiye and Iran were closer to the Eurasian clade (clade 3).

3.2.2 Bayesian inference species-tree

We constructed a Bayesian inference phylogenetic tree using parsimony informative sites to understand the evolutionary relationships between populations using RAD sequences from Türkiye and Iran, and whole genome sequences from world populations. According to the Bayesian tree, the Middle Eastern populations are associated with the European populations, while Georgian brown bears appear to be evolutionarily distant from Middle Eastern bears (Figure 3.7). This tree is not consistent with the other phylogenetic trees. As we could not obtain reasonable tree and high posterior probability values with the outgroup *Ursus Maritimus* and Slovenia sample, we excluded them from this analysis. In addition, we conducted this Bayesian analysis using *Americanus* as the outgroup; this tree can be found in Supplementary Figure A.2.

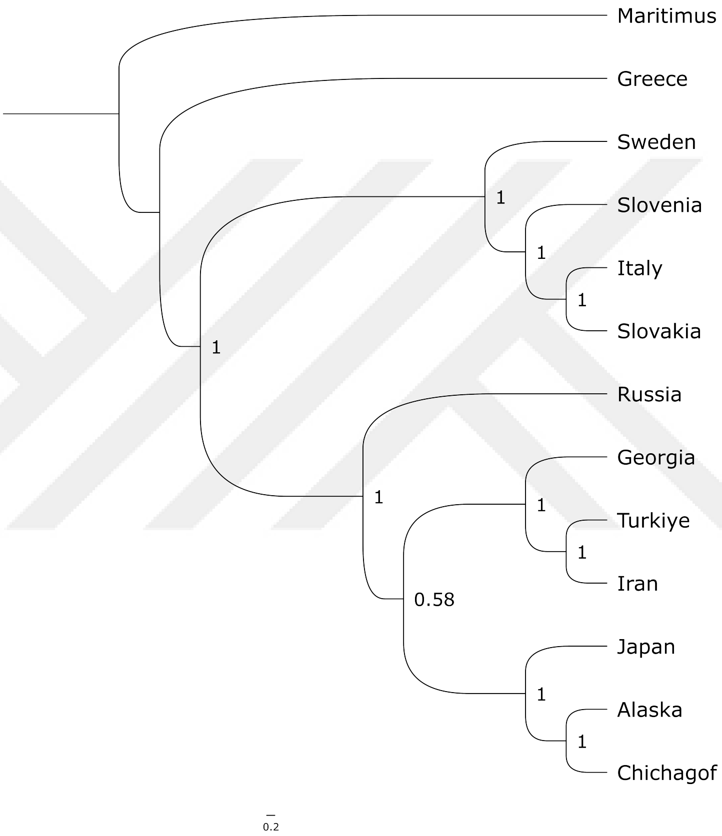


Figure 3.6: ASTRAL species tree of brown bears based on the Maximum likelihood inferred from RAD sequences from Türkiye and Iran, and whole genome sequences from world populations

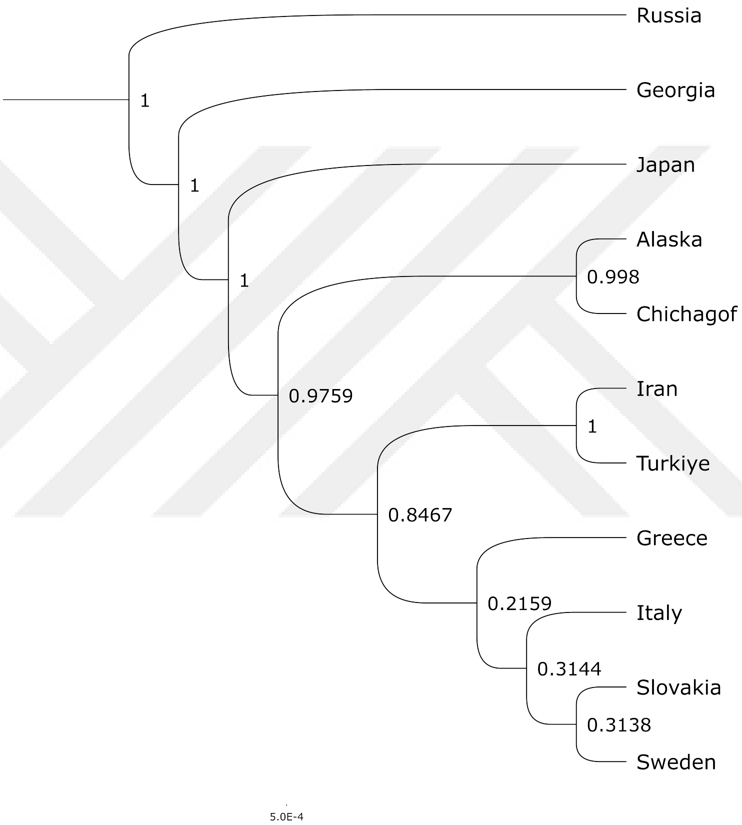


Figure 3.7: Bayesian inference tree of brown bears based on RAD sequences from Iran and Türkiye and whole genome sequences from world populations by Starbeast3 template of Beast2

3.2.3 Principal Component Analysis

We ran an analysis to determine the sites found in at least 10 individuals (-minInd 10) for a total of 15 samples containing 5 Iran samples and 10 Türkiye samples. As a result, we acquired a total of 246.033 SNPs. We performed principal component analysis (PCA) using these SNPs to identify population structure and determine groups with similar genetic profiles. PCA analysis showed that populations in western Türkiye, including Ankara and Bolu, were distinct from those in eastern Türkiye populations, including Hakkari, Iran, Kars, and Tunceli. Also, Kars populations appear distinct from other Eastern populations.

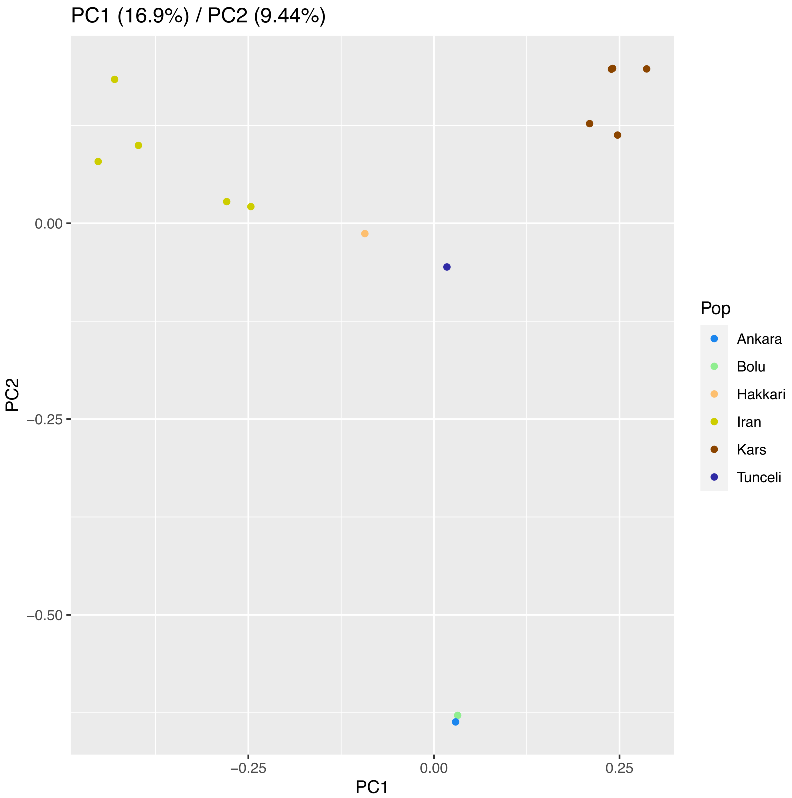


Figure 3.8: Principal Component Analysis (PCA) of the Türkiye (n=10) and Iran (n=5) populations

Chapter 4

CONCLUSION

Our findings contribute to the knowledge of the evolutionary history and genetic diversity of Middle Eastern brown bears. We used mtDNA and RAD sequences to compare different markers and expand our understanding of the phylogenetic patterns of brown bears. The phylogenetic trees and median-joining network based on mitochondrial DNA demonstrated a similar phylogeographical structure with previous studies. We found that there is a high level of genetic variation in Türkiye's populations, with 18 haplotypes and three main haplogroups. Additionally, our research confirmed genetic differentiation within Middle Eastern clade 7, and demonstrated the geographical distribution of clade 7 subclades. We found that Turkish brown bear bears dispersed into all three clades and formed distinct subgroups merging with other Middle Eastern populations in each clade. We also estimated the divergence times, and the split time of the main haplogroups was consistent with the previous estimates. However, we found the split times of subclades before the LGM glacial period. The hypotheses, including female philopatry and geographical boundaries, may contribute to genetic variation resulting in the emergence of subclades. Furthermore, phylogenetic analysis using RAD sequences of Turkey and Iran populations and whole genome sequences of other world populations showed the distinction between Eastern and Western world populations. PCA analysis supported this finding with the distinction between Eastern and Western regions in Türkiye. In conclusion, Türkiye and the Middle Eastern brown bears are distinct and have higher haplotype diversity than other world populations.

Our future objectives include using SNPs to generate a maximum likelihood tree using our RAD and whole genome sequences from GenBank and improving RAD-phylogeny trees using parsimony-informative sites. In addition, we aim to calculate effective population size and figure out how effective population size affected mutation rate and genetic diversity. This will allow us to acquire a complete picture of the evolutionary history and genetic differentiation of brown bears.

BIBLIOGRAPHY

A. Ruiz-Villalba, M. B. v. d. H., J. M. Ruijter. (2021). Use and misuse of cq in qpcr data analysis and reporting. *Life (Basel, Switzerland)*, *11*(1), 496. https://doi.org/10.3390/life11060496

Ambarli, H., Ertürk, A., & Soyumert, A. (2016). Current status, distribution, and conservation of brown bear (ursidae) and wild canids (gray wolf, golden jackal, and red fox; canidae) in turkey. *Turkish Journal of Zoology*, *40*(1), 944–956. 10.1111/j.1472-4642.2009.00586.x

Ashrafzadeh, M. R., Kaboli, M., & Naghavi, M. R. (2016). Mitochondrial dna analysis of iranian brown bears (ursus arctos) reveals new phylogeographic lineage. *Mammalian Biology*, *81*(1), 1–9.

Borowiec, M. L. (2016). Amas: A fast tool for alignment manipulation and computing of summary statistics. *PeerJ*, *4*, e1660.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). Beast 2: A software platform for bayesian evolutionary analysis. *Plos Computational Biology*, *10*(1), 496. https://doi.org/10.1371/journal.pcbi.1003537

Calvignac, S., Hughes, S., & Hänni, C. (2009). Genetic diversity of endangered brown bear (ursus arctos) populations at the crossroads of europe, asia and africa. *Diversity and Distributions*, *15*(5), 742–750.

Can, Ö., & Togan, İ. (2004). Status and management of brown bears in turkey. *Ursus*, *15*(1), 48–53.

Cariou, M., Duret, L., & Charlat, S. (2013). Is rad-seq suitable for phylogenetic inference? an in silico assessment and optimization. *Ecology and Evolution*, *3*(1), 846–852. doi:%2010.1002/ece3.512

Çilingir, F. G., Pekşen, Ç. A., Ambarlı, H., Beerli, P., & Bilgin, C. C. (2016). Exceptional maternal lineage diversity in brown bears (ursus arctos) from turkey. *Zoological Journal of the Linnean Society*, *176*(12), 463–477. https://doi.org/10.1111/zoj.12322

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools [giab008]. *GigaScience*, *10*(2). https://doi.org/10.1093/gigascience/giab008

Davison, J., Ho, S. Y., Bray, S. C., Korsten, M., Tammeleht, E., Hindrikson, M., Østbye, K., Østbye, E., Lauritzen, S.-E., Austin, J., et al. (2011). Late-quaternary biogeographic scenarios for the brown bear (ursus arctos), a wild mammal model species. *Quaternary Science Reviews*, *30*(3-4), 418–430.

de Jong, M. J., Niamir, A., Wolf, M., Kitchener, A. C., Lecomte, N., Seryodkin, I. V., Fain, S. R., Hagen, S. B., Saarma, U., & Janke, A. (2023). Range-wide whole-genome resequencing of the brown bear reveals drivers of intraspecies divergence. *Communications Biology*, *6*(1), 153.

Euronatur. (n.d.). *Profile: Brown bear (ursus arctos)*. https://www.euronatur.org/en/what-we-do/bear-wolf-lynx/bears-in-europe/profile-brown-bear-ursus-arctos

Hall, B. G. (2013). Building phylogenetic trees from molecular data with mega. *Molecular biology and evolution*, *30*(5), 1229–1235.

Hall, T. (2013). Bioedit, version 7.2. 5. *Ibis Biosciences, Carlsbad, CA, USA*.

Hassanin, A. (2015). The role of pleistocene glaciations in shaping the evolution of polar and brown bears. evidence from a critical review of mitochondrial and nuclear genome analyses. *Comptes rendus biologies*, *338*(7), 494–501.

Herrero, S., Peyton, B., & Servheen, C. (1999). Bears: Status survey and conservation action plan.

Hirata, D., Mano, T., Abramov, A. V., Baryshnikov, G. F., Kosintsev, P. A., Vorobiev, A. A., Raichev, E. G., Tsunoda, H., Kaneko, Y., Murata, K., et al. (2013). Molecular phylogeography of the brown bear (ursus arctos) in northeastern asia based on analyses of complete mitochondrial dna sequences. *Molecular biology and evolution*, *30*(7), 1644–1652.

Hohenlohe, P. A., Bassham, S., Etter, P. D., an E. A. Johnson, N. S., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced rad tags. *PLOS Genetics*, *6*(1), e1000862. https://doi.org/10.1371/journal.pgen.1000862

Kasamatsu, H., Robberson, D. L., & Vinograd, J. (1971). A novel closed-circular mitochondrial dna with properties of a replicating intermediate. *Proceedings of the National Academy of Sciences*, *68*(9), 2252–2257.

Kivisild, T. (2015). Maternal ancestry and population history from whole mitochondrial genomes. *Investigative genetics*, *6*, 1–10.

Kohn, M., Knauer, F., Stoffella, A., Schröder, W., & Pääbo, S. (1995). Conservation genetics of the european brown bear–a study using excremental pcr of nuclear and mitochondrial sequences. *Molecular Ecology*, *4*(1), 95–103. https://doi.org/10.1111/j.1365-294X.1995.tb00196.x

Korsten, M., Ho, S. Y., Davison, J., Pähn, B., Vulla, E., Roht, M., Tumanov, I. L., Kojola, I., Andersone-Lilley, Z., Ozolins, J., et al. (2009). Sudden expansion of a single brown bear maternal lineage across northern continental eurasia after the last ice age: A general demographic model for mammals? *Molecular ecology*, *18*(9), 1963–1979.

Kumar, S., Tamura, K., Jakobsen, I. B., & Nei, M. (2001). Mega2: Molecular evolutionary genetics analysis software. *Bioinformatics*, *17*(12), 1244–1245.

Kumar, V., Lammers, F., Bidon, T., Pfenninger, M., Kolter, L., Nilsson, M. A., & Janke, A. (2017). The evolutionary history of bears is characterized by gene flow across species. *Scientific Reports*, *7*(1), 46487.

Kutschera, V. E., Bidon, T., Hailer, F., Rodi, J. L., Fain, S. R., & Janke, A. (2014). Bears in a forest of gene trees: Phylogenetic inference is complicated by incomplete lineage sorting and gene flow. *Molecular biology and evolution*, *31*(1), 145–55. 10.1007/s00239-001-0063-4

Lan, T., Leppälä, K., Tomlin, C., Talbot, S. L., Sage, G. K., Farley, S. D., Shideler, R. T., Bachmann, L., Wiig, Ø., Albert, V. A., et al. (2022). Insights into bear evolution from a pleistocene polar bear genome. *Proceedings of the National Academy of Sciences*, *119*(24), e2200016119.

Lansman, R. A., Avise, J. C., & Huettel, M. D. (1983). Critical experimental test of the possibility of ”paternal leakage” of mitochondrial dna. *Proceedings of the National Academy of Sciences*, *80*(7), 1969–1971.

Larizza, A., Pesole, G., Reyes, A., Sbisà, E., & Saccone, C. (2002). Lineage specificity of the evolutionary dynamics of the mtdna d-loop region in rodents. *J Mol Evol.*, *54*(1), 145–55. 10.1007/s00239-001-0063-4

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows–wheeler transform. *bioinformatics*, *25*(14), 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Subgroup, 1. G. P. D. P. (2009). The sequence alignment/map format and samtools. *bioinformatics*, *25*(16), 2078–2079.

Lopez, J. V., Culver, M., Stephens, J. C., Johnson, W. E., & O’Brien, S. J. (1997). Rates of nuclear and cytoplasmic mitochondrial dna sequence divergence in mammals. *Molecular Biology and Evolution*, *14*(3), 277–286.

MacDonald, T., & Thanukos, A. (n.d.). What is an evolutionary tree? https://evolution.berkeley.edu/the-tree-room/evolutionary-trees-a-primer/what-is-an-evolutionary-tree/

Matsuhashi, T., Masuda, R., Mano, T., Murata, K., & Aiurzaniin, A. (2001). Phylogenetic relationships among worldwide populations of the brown bear ursus arctos. *Zoological Science*, *18*(8), 1137–1143. 10.2108/zsj.18.1137

Mclennan, D. A. (2010). How to read a phylogenetic tree. *Evolution: Education and Outreach*, *40*(1), 506–519. 10.1007/s12052-010-0273-6

Mizumachi, K., Spassov, N., Kostov, D., Raichev, E. G., Peeva, S., Hirata, D., Nishita, Y., Kaneko, Y., & Masuda, R. (2020). Mitochondrial haplogrouping of the ancient brown bears (ursus arctos) in bulgaria, revealed by the aplp method. *Mammal Research*, *65*, 413–421.

Molodtseva, A. S., Makunin, A. I., Salomashkina, V. V., Kichigin, I. G., Vorobieva, N. V., Vasiliev, S. K., Shunkov, M. V., Tishkin, A. A., Grushin, S. P., Anijalg, P., et al. (2022). Phylogeography of ancient and modern brown bears from eastern eurasia. *Biological Journal of the Linnean Society*, *135*(4), 722–733.

Murtskhvaladze, M., Gavashelishvili, A., & Tarkhnishvili, D. (2010). Geographic and genetic boundaries of brown bear (ursus arctos) population in the caucasus. *Molecular Ecology*, *19*(9), 1829–1841.

Nelson, R. A., Folk Jr, G. E., Pfeiffer, E. W., Craighead, J. J., Jonkel, C. J., & Steiger, D. L. (1983). Behavior, biochemistry, and hibernation in black, grizzly, and polar bears. *Bears: their biology and management*, 284–290.

Paetkau, D., & Strobeck, C. (1994). Microsatellite analysis of genetic variation in black bear populations. *Molecular ecology*, *3*(5), 489–495.

Pedersen, B. S., & Quinlan, A. R. (2018). Mosdepth: Quick coverage calculation for genomes and exomes. *Bioinformatics*, *34*(5), 867–868.

Rambaut, A. (2018). Figtree v1. 4.4, a graphical viewer of phylogenetic trees. 2014. java.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). Dnasp 6: Dna sequence polymorphism analysis of large data sets. *Molecular biology and evolution*, *34*(12), 3299–3302.

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X., & Rozas, R. (2003). Dnasp, dna polymorphism analyses by the coalescent and other methods. *Bioinformatics*, *19*(18), 2496–2497.

Rubin, B. E. R., Ree, R. H., & Moreau, C. S. (2012). Inferring phylogenies from rad sequence data. *PLOS ONE*, *11*(1). https://doi.org/10.1371/journal.pone.0033394

Saarma, U., Ho, S. Y., Pybus, O. G., Kaljuste, M., Tumanov, I. L., Kojola, I., Vorobiev, A. A., Markov, N. I., Saveljev, A. P., Valdmann, H., et al. (2007). Mitogenetic structure of brown bears (ursus arctos l.) in northeastern europe and a new time frame for the formation of european brown bear lineages. *Molecular Ecology*, *16*(2), 401–413.

Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., & Saccone, C. (1997). Mammalian mitochondrial d-loop region structural analysis: Identification of new conserved sequences and their functional and evolutionary implications. *Gene*, *205*(1-2), 125–140.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Molecular systems biology*, *7*(1), 539.

Stamatakis, A. (2006). Raxml-vi-hpc: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, *22*(21), 2688–2690.

Stenglein, J., De Barba, M., Ausband, D., & Waits, L. (2010). Impacts of sampling location within a faeces on dna quality in two carnivore species. *Molecular ecology resources*, *10*(1), 109–114.

Swofford, D. L., et al. (2002). Phylogenetic analysis using parsimony (* and other methods). *Version*, *4*, b10.

Taberlet, P., & Bouvet, J. (1994). Mitochondrial dna polymorphism, phylogeography, and conservation genetics of the brown bear ursus arctos in europe. *Proceedings: Biological Sciences*, *255*(1), 195–200. 10.1098/rspb.1994.0028

Talbot, S. L., & Shields, G. F. (1996). A phylogeny of the bears (ursidae) inferred from complete sequences of three mitochondrial genes. *Molecular phylogenetics and evolution*, *5*(3), 567–575.

Tammeleht, E., Remm, J., Korsten, M., Davison, J., Tumanov, I., Saveljev, A., Männil, P., Kojola, I., & Saarma, U. (2010). Genetic structure in large, continuous mammal populations: The example of brown bears in northwestern eurasia. *Molecular ecology*, *19*(24), 5359–5370.

Tao, M., You, C.-P., Zhao, R.-R., Liu, S.-J., Zhang, Z.-H., Zhang, C., & Liu, Y. (2014). Animal mitochondria: Evolution, function, and disease. *Current molecular medicine*, *14*(1), 115–124.

Tumendemberel, O., Zedrosser, A., Proctor, M. F., Reynolds, H. V., Adams, J. R., Sullivan, J. M., Jacobs, S. J., Khorloojav, T., Tserenbataa, T., Batmunkh, M., et al. (2019). Phylogeography, genetic diversity, and connectivity of brown bear populations in central asia. *PLoS One*, *14*(8), e0220746.

Waits, L., Paetkau, D., Strobeck, C., & Ward, R. H. (1998). A comparison of genetic diversity in north american brown bears. *Ursus*, 307–314.

Wehausen, J., Ramey, R. 2., & Epps, C. (2004). Experiments in dna extraction and pcr amplification from bighorn sheep feces: The importance of dna extraction method. *Journal of Heredity*, *95*(6), 503–509.

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). Astral-iii: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC bioinformatics*, *19*(6), 15–30.

Appendix A

APPENDIX

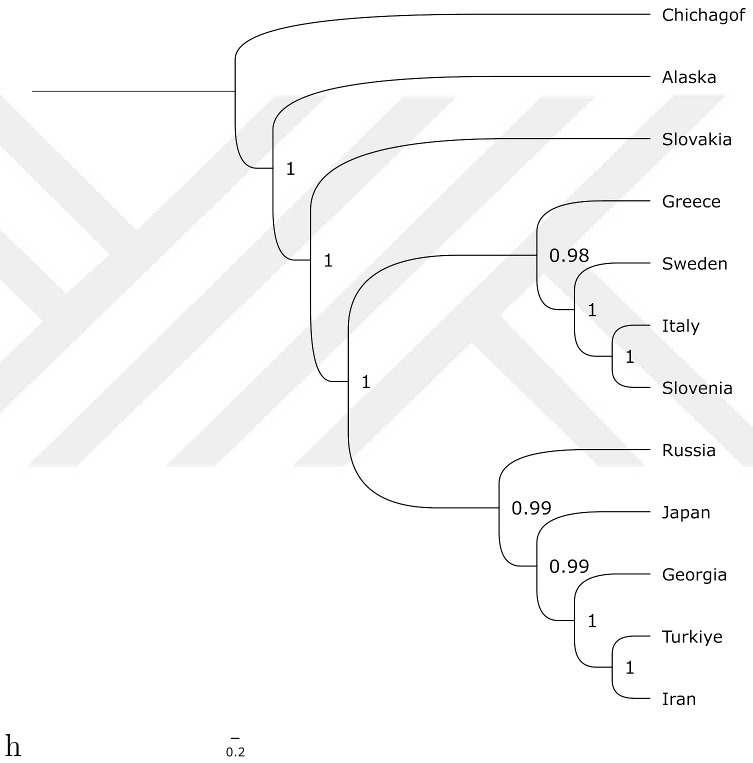
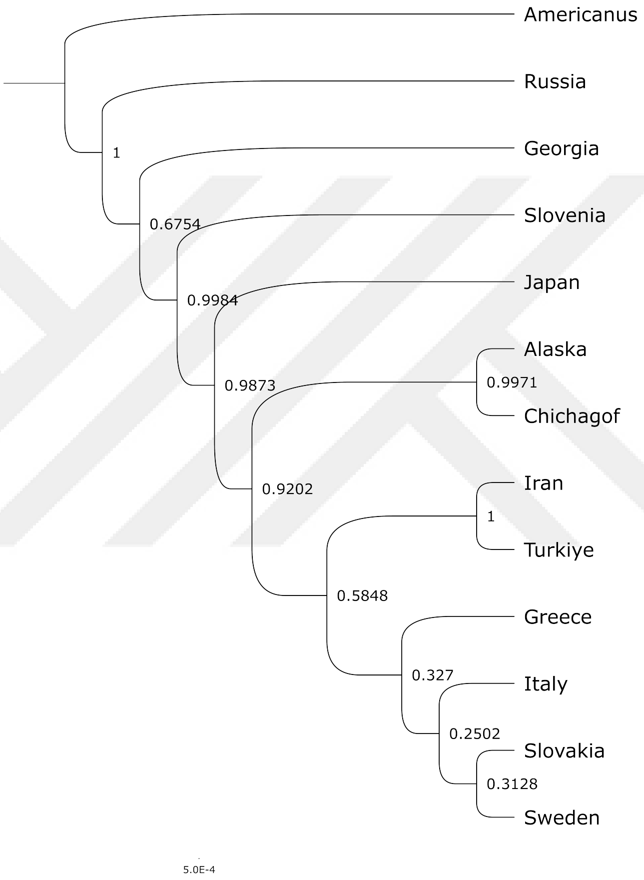
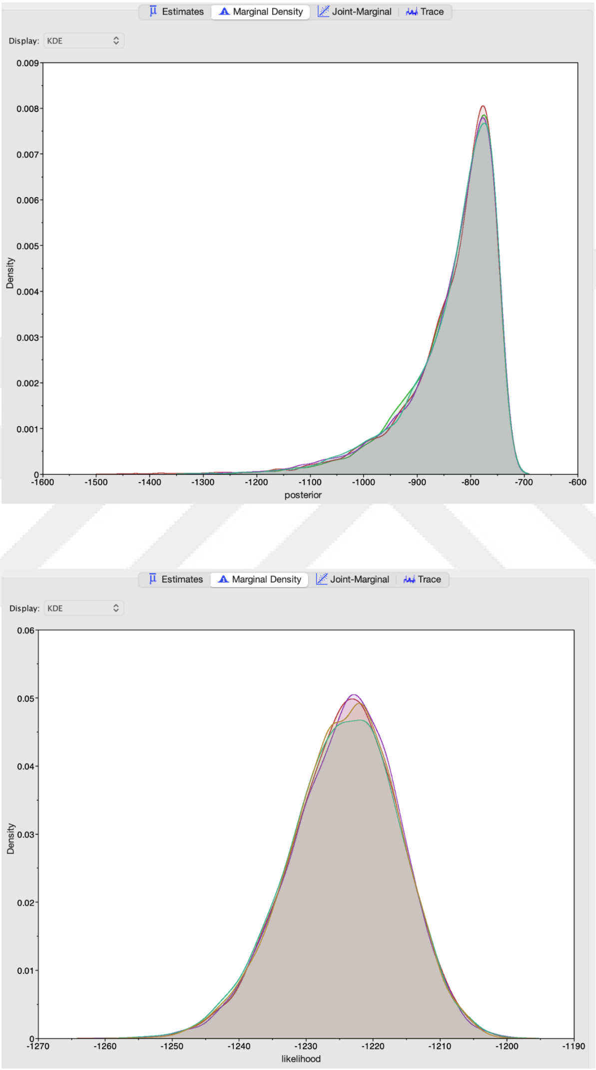


Figure A.1: Maximum likelihood of brown bears based on RAD-sequences using top 50 parsimony-informative locus



h

Figure A.2: Bayesian inference tree of brown bears based on RAD sequences from Iran and Turkey and whole genome sequences from world populations with outgroup Ursus americanus.



h

Figure A.3: Marginal density graphics of four replicates of Bayesian analysis for mtDNA sequences

Table A.1: Haplotype information for Türkiye brown bear samples

Haplotype ID	**Accession number**	**Location**	**References**
Hap01_Turkey	BTR001, BTR008, BTR018, BTR613	Kars	Our study
Hap01_Turkey	BTR342	Bayburt	Our study
Hap01_Turkey	BTR514	Şırnak	Our study
Hap01_Turkey	BTR591	Erzincan	Our study
Hap01_Turkey	BTR526, BTR529, BTR539, BTR540, BTR575	Hakkari	Our study
Hap01_Turkey	KT438638	Artvin	Çilingir et al., 2016
Hap02_Turkey	BTR017, BTR061, BTR064	Bilecik	Our study
Hap02_Turkey	BTR605	Yalova	Our study
Hap02_Turkey	KT438655.1	Bursa	Çilingir et al., 2016
Hap03_Turkey	BTR078, BTR085, BTR090	Erzurum	Our study
Hap03_Turkey	BTR130, BTR132, BTR204	Bolu	Our study
Hap03_Turkey	BTR140	Sakarya	Our study
Hap03_Turkey	BTR199, BTR225, BTR232	Ankara	Our study
Hap03_Turkey	BTR232, BTR235	Karabük	Our study
Hap03_Turkey	BTR251	Zonguldak	Our study
Hap03_Turkey	BTR310, BTR314	Tokat	Our study
Hap03_Turkey	BTR333	Erzincan	Our study
Hap03_Turkey	BTR344	Bayburt	Our study
Hap03_Turkey	BTR400, BTR404	Trabzon	Our study
Hap03_Turkey	BTR407	Tunceli	Our study
Hap04_Turkey	BTR102, BTR105, BTR111, BTR122	Antalya	Our study
Hap05_Turkey	BTR141	Mersin	Our study
Hap06_Turkey	BTR213, BTR214	Düzce	Our study
Hap06_Turkey	BTR253	Zonguldak	Our study
Hap06_Turkey	BTR388	Muğla	Our study
Hap07_Turkey	BTR245	Kastamonu	Our study
Hap07_Turkey	BTR281, BTR284, BTR285, BTR286	Çorum	Our study

Table A.2: Haplotype information for Türkiye brown bear samples

Haplotype ID	**Accession number**	**Location**	**References**
Hap08_Turkey	BTR250	Zonguldak	Our study
Hap09_Turkey	BTR294	Sinop	Our study
Hap10_Turkey_	BTR346	Erzurum	Our study
Hap10_Turkey_	BTR373, BTR374, BTR376, BTR377	Ardahan	Our study
Hap10_Turkey_	BTR395, BTR397, BTR403	Trabzon	Our study
Hap10_Turkey_	KT438621.1-KT438633.1, KT438643.1-KT438649.1	Artvin	Çilingir et al., 2016
Hap11_Turkey_	BTR354, BTR355, BTR359, BTR364	Artvin	Our study
Hap11_Turkey_	BTR371	Ardahan	Our study
Hap11_Turkey_	KT438640.1-KT438642.1	Artvin	Çilingir et al., 2016
Hap12_Turkey_	BTR381	Tunceli	Our study
Hap12_Turkey_	BTR454	Bingöl	Our study
Hap13_Turkey_	BTR386	Bursa	Our study
Hap14_Turkey_	BTR429, BTR435, BTR440, BTR446	Sivas	Our study
Hap14_Turkey_	BTR441, BTR443	Yozgat	Our study
Hap15_Turkey_	BTR438	Sinop	Our study
Hap15_Turkey_	BTR442	Yozgat	Our study
Hap16_Turkey_	BTR473 BTR477 BTR487	Muş	Our study
Hap16_Turkey_	BTR494 BTR501 BTR506 BTR507 BTR508	Bitlis	Our study
Hap17_Turkey_	BTR587	Tunceli	Our study
Hap17_Turkey_	BTR588	Erzincan	Our study
Hap18_Turkey_	BTR614	Muğla	Our study

Table A.3: Accession number of whole genome sequences used in phylogenetic analysis with RAD-sequences

Accession number	**Region**
SRR830213	Alaska
SRR830337	Alaska
SRR1693624	Chichagof
SRR1692419	Chichagof
SRR5878351	Greece
SRR5878355	Greece
SRR1693654	Sweden
SRR873880	Sweden
SRR873879	Sweden
SRR5878360	Italy
SRR5878350	Italy
DRR276777	Japan
DRR276778	Japan
ERR2678639	Georgia
ERR2678638	Slovenia
ERR2678640	Russia
SRR5878349	Slovakia
SRR518711	Alaska
SRR935626	Maritimus

Table A.4: Mitochondria D-loop sequences taken from GenBank

Organism	**Accession ID**	**Countries**	**Continent**
Ursus arctos	1004125984	Russia	Russia
Ursus arctos	1004125984	Azerbaijan	Middle East
Ursus arctos	1004125984	Armenia	Middle East
Ursus arctos syriacus	FN292973.1 - FN292982.1	Syriacus Zoo	Middle East
Ursus arctos	975866350	Turkey	Middle East
Ursus arctos	558762204	Russia	Russia
Ursus arctos	333601421	Croatia	Europe
Ursus arctos	558761416	Russia	Russia
Ursus arctos	333601421	Croatia	Europe
Ursus arctos	558761416	Russia	Russia
Ursus arctos	EU574910.1 - EU574911.1	Kenai Peninsula, south central Alaska	Alaska
Ursus arctos	AB041258.1	Kodiak	Alaska
Ursus arctos	732561212	Alaska	Alaska
Ursus arctos	AB013040.1 - AB013070.1	Hokkaido, Japan	Asia
Ursus arctos	X75875.1 - X75876.1	Slovakia	Europe
Ursus arctos	X75872.1 - X75873.1	Romania	Europe
Ursus arctos	X75878.1	Pyrenees-France	Europe
Ursus arctos	X75871.1	Norway	Europe
Ursus arctos	X75870.1	Greece	Europe
Ursus arctos	X75869.1	Estonia	Europe
Ursus arctos	X75867.1	Croatia	Europe
Ursus arctos	X75865.1 - X75866.1	Spain	Europe
Ursus arctos	X75864.1	Bulgaria	Europe
Ursus arctos	X75862.1	Abruzzo-Italy	Europe
Ursus arctos	LC349016.1	Japan	Asia
Ursus arctos	LC349015.1	Japan	Asia
Ursus arctos	933847690	Iran	Middle East
Ursus arctos	LC349016.1	Japan	Asia
Ursus arctos	LC349015.1	Japan	Asia
Ursus arctos	933847690	Iran	Middle East
Ursus arctos	GU057343.1 - GU057377.1	Georgia	Middle East

Table A.5: Mitochondria D-loop sequences taken from GenBank (continued)

Organism	**Accession ID**	**Countries**	**Continent**
Ursus arctos	MN477248.1 - MN477249.1	Spain: Asturias	Europe
Ursus arctos	MG254058.1	Greece	Europe
Ursus arctos	MG254054.1 - MG254057.1	Romania	Europe
Ursus arctos	MG254053.1	Poland	Europe
Ursus arctos	MG254052.1	Romania	Europe
Ursus arctos	MG254051.1	Ukraine	Europe
Ursus arctos	MG254049.1 - MG254050.1	Poland	Europe
Ursus arctos	LT962548.1	Greece	Europe
Ursus arctos	LT962542.1 - LT962546.1	Greece	Europe
Ursus arctos	LT962541.1	Albania	Europe
Ursus arctos	LT962540.1	Greece	Europe
Ursus arctos	LT962489.1 - LT962539.1	Greece	Europe
Ursus arctos	HG426390.1 - HG426391.1	Northwestern america	North america
Ursus arctos	HG426386.1 - HG426387.1	Montana	North america
Ursus arctos	HE657215.1 - HE657216.1	Norway	Europe
Ursus arctos	HE657213.1 - HE657214.1	Romania	Europe
Ursus arctos	HE657212.1	Sweden	Europe
Ursus arctos	HE657211.1	Romania	Europe
Ursus arctos	HE657208.1 - HE657209.1	Norway	Europe
Ursus arctos	HE657205.1 - HE657207.1	USA	America
Ursus arctos	HE657204.1	Canada	North America
Ursus arctos	HE657202.1	USA	America
Ursus arctos	HE657201.1	Bulgaria	Europe
Ursus arctos	HE657199.1 - HE657200.1	Romania	Europe
Ursus arctos	270160680	Russia, Finland, Estonia, Eurasia	Russia, Europe
Ursus arctos syriacus	FN292971.1	Lebanon	Middle East
Ursus arctos	MT152317.1	Russia	Russia

Table A.6: ESS values of tracer files of Bayesian inference tree based on mtDNA sequences

Statistic	**Mean**	**ESS**
posterior	-832.638	3634
likelihood	-1224.277	15386
prior	391.639	3637
treeLikelihood	-1224.277	15386
TreeHeight	4.656E-3	4430
kappa	9.434	16548
gammaShape	0.137	13708
clockRate	50.07	6519
popSize	1.006E-2	4443
CoalescentConstant	395.09	3637

Table A.7: List of mtDNA samples used in the study, with their locations and nanodrop value

Sample ID	**x**	**y**	**City**	**Nanodrop (ng/μl)**
BTR199	32.5807	40.4606	Ankara	122.0
BTR225	32.5413	40.4862	Ankara	32.2
BTR111	31.8133	37.3093	Antalya	27.3
BTR105	31.8171	37.3031	Antalya	25.9
BTR127	30.6398	36.9465	Antalya	24.1
BTR122	31.8213	37.2984	Antalya	23.3
BTR102	31.8167	37.3046	Antalya	23.9
BTR373	42.7834	41.4939	Ardahan	28.5
BTR376	42.7932	41.4952	Ardahan	26.2
BTR371	42.4511	41.0391	Ardahan	26.8
BTR377	42.7931	41.4947	Ardahan	23.2
BTR374	42.7832	41.4938	Ardahan	32.1
BTR354	41.7004	41.1329	Artvin	17.9
BTR364	42.1860	41.0538	Artvin	17.4
BTR359	42.1860	41.0537	Artvin	19.2
BTR355	41.6925	41.1159	Artvin	19.0
BTR342	39.7194	40.0456	Bayburt	20.2
BTR344	40.39862	40.05469	Bayburt	14.5
BTR017	30.3804	40.0145	Bilecik	18.8
BTR064	29.9456	40.0705	Bilecik	18.0
BTR061	29.9411	40.0767	Bilecik	21.2
BTR454	40.0509	39.2502	Bingöl	17.7
BTR508	42.2381	38.6438	Bitlis	5.6
BTR507	42.2381	38.6438	Bitlis	8.8
BTR501	42.2381	38.6438	Bitlis	17.9
BTR506	42.2381	38.6438	Bitlis	32.1
BTR494	42.2389	38.6450	Bitlis	27.7

Table A.8: List of mtDNA samples used in the study, with their locations and nanodrop value

Sample ID	**x**	**y**	**City**	**nanodrop (ng/μl)**
BTR132	31.43555	40.5764	Bolu	7.3
BTR130	31.38394	40.58974	Bolu	9.1
BTR204	31.5834	40.5357	Bolu	25.7
BTR386	29.2322	40.1238	Bursa	26.8
BTR286	34.3895	41.1680	Çorum	25.2
BTR281	34.3879	41.1685	Çorum	15.3
BTR284	34.3861	41.1684	Çorum	13.0
BTR285	34.3875	41.1684	Çorum	15.1
BTR213	31.2397	41.0838	Düzce	26.4
BTR214	31.2426	41.0783	Düzce	26.7
BTR591	39.6553	39.5702	Erzincan	30.7
BTR333	39.6901	39.5893	Erzincan	24.3
BTR588	38.5391	39.29363	Erzincan	20.8
BTR078	42.2596	40.5368	Erzurum	29.0
BTR085	42.2347	40.5320	Erzurum	33.7
BTR090	42.2304	40.5230	Erzurum	25.8
BTR346	41.4272	40.4016	Erzurum	23.4
BTR128	37.2958	37.0258	Gaziantep	16.4
BTR526	43.7223	37.5450	Hakkari	9.4
BTR529	43.7687	37.6592	Hakkari	7.9
BTR540	43.7237	37.7457	Hakkari	19.0
BTR575	44.0954	37.4239	Hakkari	79.5

Table A.9: List of mtDNA samples used in the study, with their locations and nanodrop value

Sample ID	**x**	**y**	**City**	**nanodrop (ng/μl)**
BTR232	32.4405	41.2739	Karabük	27.8
BTR235	32.4368	41.2653	Karabük	8.6
BTR008	42.6113	40.3016	Kars	33.2
BTR001	42.6113	40.3018	Kars	16.7
BTR018	42.6113	40.3018	Kars	18.2
BTR613	42.6113	40.3018	Kars	20.9
BTR245	33.1980	41.7229	Kastamonu	17.0
BTR141	32.7145	36.4081	Mersin	83.7
BTR143	33.04687	36.82107	Mersin	6.8
BTR614	28.3591	36.8359	Muğla	43.4
BTR388	28.6819	37.0224	Muğla	7.1
BTR473	41.5281	39.1868	Muş	14.8
BTR487	41.5866	39.1300	Muş	14.9
BTR477	41.5639	39.1959	Muş	22.7
BTR140	30.3393	40.6107	Sakarya	16.5
BTR438	37.0661	40.0199	Sinop	12.7
BTR294	34.5739	41.7484	Sinop	10.5
BTR514	42.8663	37.4788	Şırnak	15.9
BTR435	37.9926	40.0129	Sivas	12.1
BTR446	37.9662	39.2562	Sivas	36.2
BTR440	37.0901	40.0149	Sivas	6.0
BTR429	38.4391	40.0179	Sivas	16.9
BTR310	37.1723	40.3982	Tokat	30.0
BTR314	37.1688	40.4073	Tokat	45.6
BTR404	40.1671	40.5844	Trabzon	11.1

Table A.10: List of mtDNA samples used in the study, with their locations and nanodrop value

Sample ID	**x**	**y**	**City**	**nanodrop (ng/μl)**
BTR400	40.1402	40.5844	Trabzon	13.8
BTR395	39.6592	40.6943	Trabzon	21.1
BTR403	40.1656	40.5844	Trabzon	12.7
BTR397	39.6685	40.7002	Trabzon	30.9
BTR587	39.0301	39.0624	Tunceli	12.5
BTR407	39.5080	39.1681	Tunceli	12.1
BTR381	39.8492	39.4098	Tunceli	13.7
BTR605	28.99826	40.54808	Yalova	11.1
BTR443	35.8872	39.5534	Yozgat	15.0
BTR441	35.9031	39.5595	Yozgat	16.4
BTR442	35.8862	39.5495	Yozgat	13.0
BTR251	31.8231	41.2542	Zonguldak	13.4
BTR250	31.8463	41.2604	Zonguldak	24.2
BTR253	31.7082	41.1153	Zonguldak	20.2
BTR254	31.7047	41.1301	Zonguldak	20.6
BTR539	43.7238	37.7454	Hakkari	7.6
IBB12	53.1146	36.0324	İran	37.0
IBB2	48.0005	38.1748	İran	31.4
IBB13	55.5444	37.2312	İran	23.8
IBB8	48.0933	38.5216	İran	24.8
IBB10	53.3552	36.2603	İran	26.8