MTDNA BASED GENETIC DIVERSITY OF NATIVE SHEEP BREEDS AND ANATOLIAN MOUFLON (*OVIS GMELINII ANATOLICA*) IN TURKEY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

SEVGİN DEMİRCİ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
BIOINFORMATICS

MAY 2012

MTDNA BASED GENETIC DIVERSITY OF NATIVE SHEEP BREEDS AND ANATOLIAN MOUFLON (*OVIS GMELINII ANATOLICA*) IN TURKEY

Submitted by **SEVGİN DEMİRCİ** in partial fulfillment of the requirements for the degree of **Master of Science in Bioinformatics, Middle East Technical University** by,

Prof. Dr. Nazife BAYKAL ____________________
Director, **Informatics Institute**

Assist. Prof. Dr. Yeşim AYDIN SON ____________________
Head of Department, **Medical Informatics**

Prof. Dr. İnci TOGAN ____________________
Supervisor, **Biology, METU**

Examining Committee Members:

Assist. Prof. Dr. Yeşim AYDIN SON ____________________
Medical Informatics, METU

Prof. Dr. İnci TOGAN ____________________
Biology, METU

Assoc. Prof. Dr. Tolga CAN ____________________
Computer Engineering, METU

Prof. Dr. İrfan KANDEMİR ____________________
Biology, Ankara University

Dr. Evren KOBAN BAŞTANLAR ____________________
MAM, TUBİTAK

Date: 28.05.2012

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Sevgin Demirci

Signature: _______________

ABSTRACT

MTDNA BASED GENETIC DIVERSITY OF NATIVE SHEEP BREEDS AND ANATOLIAN MOUFLON (*OVIS GMELINII ANATOLICA*) IN TURKEY

Demirci, Sevgin
M.S., Bioinformatics Program
Supervisor: Prof. Dr. İnci Togan

May 2012, 131 pages

In the present study, history of domestic sheep were investigated by mitochondrial DNA (mtDNA) based haplogroups (HPG) of 628 samples and mtDNA control region (CR) sequences of 240 samples from 13 Turkish sheep breeds which were located in the hearth of the first domestication center. Also, 30 Anatolian wild sheep (*Ovis gmelinii anatolica*) mtDNA CR sequences were obtained to contribute to the scenarios on initial domestication stages of sheep. Haplogroup compositions of breeds were identified with SSCP method by using mtDNA ND2 region. The genetic diversity and relationship between haplogroups were calculated. Phylogenetic analyses of haplogroups such as median joining networks and neighbor joining trees were constructed for mtDNA CR, cytochrome *B* (*cytB*) and combined CR-*cytB* sequences with sequences from the present study together with sequences retrieved from NCBI (http://www.ncbi.nlm.nih.gov/).

Results of the present study showed that all previously observed haplogroups (HPG A-E) were present in Turkish sheep breeds. Two individuals from rare HPG D and eleven individuals from rare HPG E were detected and sequenced. With increased sample size, for HPG E, past population expansion was observed as was the case of HPG A, B and C with mismatch distributions and neutrality tests. Spatial autocorrelation analyses and synthetic map with respect to mtDNA (maternal) pattern revealed that Turkey was separated into two regions which may be attributed to the imprints of third migration of sheep associated with the arrivals of nomadic Turks to Anatolia nearly 1000 years before present. Finally, *Ovis gmelinii anatolica* samples exhibited two haplotypes; one of them belongs to HPG A (possibly feral domesticate), and the other one shows a distinct haplotype (close to HPG E and C) that was not observed before. Observed, low mtDNA diversity might be the result of isolation, fragmentation, extinction of fragments and bottlenecks. *Ovis gmelinii anatolica* can be part of the evolved descendants of the wild sheep which gave birth to the domestic sheep.

Keywords: Domestication, *Ovis gmelinii anatolica*, Turkish native sheep breeds, mtDNA haplogroups, DNA sequencing

ÖZ

TÜRKİYEDEKİ YERLİ KOYUN IRKLARININ VE ANADOLU MOUFLONUN (OVIS GMELINII ANATOLICA) MTDNA'YA DAYALI GENETİK ÇEŞİTLİLİĞİ

Demirci, Sevgin
Yüksek Lisans, Biyoinformatik Programı
Tez Yöneticisi: Prof. Dr. İnci Togan

Mayıs 2012, 131 sayfa

Bu çalışmada, evcil koyunların tarihi ilk evcilleştirmenin merkezinde yer alan 13 Türk koyun ırklarından 628 bireyin mitokondriyal DNA (mtDNA)'ya dayalı haplogrupları (HPG), ve bu bireylerin arasından seçilmiş 240 bireyin mtDNA kontrol bölgesi (CR) sekansları kullanılarak araştırılmıştır. Ayrıca, koyunun ilk evcilleştirme safhaları senaryosuna katkıda bulunmak için 30 Anadolu yaban koyunu (*Ovis gmelinii anatolica*) mtDNA CR sekansları elde edilmiştir. Irkların haplogrup kompozisyonları mtDNA ND2 bölgesi kullanılarak yapılan SSCP metoduyla tanımlanmıştır. Genetik çeşitlilik ve haplogruplar arasındaki ilişki hesaplanmıştır. Medyan birleştirme ağı ve Komşu birleştirme ağacı gibi haplogrupların filogenetik analizleri, mtDNA CR, sitokrom *B* (*cytB*) ve birleştirilmiş CR-*cytB* bölgelerine ait bu çalışmanın sekansları ve NCBI'dan (http://www.ncbi.nlm.nih.gov/) alınan sekanlar ile yapılmıştır.

Bu çalışmada sunulan sonuçlar, Türk koyun ırklarının şimdiye kadar bulunan bütün haplogrupları (HPG A-E) içerdiğini göstermiştir. Az olan HPG D'de 2 birey, az olan HPG E'de on bir birey gözlenmiş ve sekanslanmıştır. Artan örnek sayısı ile HPG E'nin, uyumsuzluk dağılımı ve nötralite testleriyle diğer yaygın haplogrouplar (HPG A, B ve C) gibi geçmişte büyüme gösterdiği gözlenmiştir. Göçer Türklerin Anadoluya yaklaşık 1000 yıl önce gelişi ile ilişkilendirilebilecek üçüncü koyun göçü sonucunda oluşabilecek Türkiyenin iki bölgeye bölünmesi, uzamsal otokorelasyon analizleri ve sentetik haritada gözlenen mtDNA'ya ilişkin (annesel) desen sayesinde ortaya çıkmıştır. Son olarak, *Ovis gmelinii anatolica* örnekleri iki haplotip olarak gözlenmiş, bunlardan bir tanesi HPG A içerinde gözlenmiş (muhtemelen yabana kaçan evcil), bir diğeri ise daha önce hiç gözlenmemiş ayrı bir haplotip (HPG E ve C'ye yakın) olarak görülmüştür. Gözlenen bu sınırlı çeşitlilik, izolasyon, parçalanma, parçaların yok olması ve darboğazın sonucu olabilir. *Ovis gmelinii anatolica,* evcil koyunu meydana getiren yaban koyunlarının evrilmiş soylarının bir parçası olabilir.

Anahtar kelimeler: Evcilleştirme, *Ovis gmelinii anatolica,* Türk yerli koyun ırkları, mtDNA haplogrupları, DNA sekanslaması

To my family,

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Prof. Dr. İnci Togan for her guidance and patience. She is always tolerant and supportive. Every time when I stumbled she encouraged me to go on. In the writing process of the thesis, she provided me a perfect working environment by sharing her own room. Besides her scientific insights her contributions to me are priceless.

Also, I would like to thank examining committee members for their valuable comments; Dr. Evren Koban Baştanlar for being like a co-advisor (I have learned most of the wet lab skills and DNA sequence analysis from her); Prof. Dr. H. Şebnem Düzgün and Dr. Serkan Kemeç for their comments on construction of the synthetic maps; Ulaş Canatalı for introducing me to the ArcGIS world; my lab mates: Eren Yüncü, N. Dilşad Dağtaş, Dr. Emel Özkan Ünal, Dr. S. Can Açan, İ. Cihan Ayanoğlu, Begüm Uzun, Hande Acar, H. Alper Döm, Dr. Havva Dinç, Dr. Fusun Özer, Dr. Arzu Karahan, Arif Parmaksız and Dr. Ceren Caner Berkman for gracious and the best lab environment ever; Dr. Emel Özkan Ünal and Teoman Ünal for opening their home when I needed; Dilek Çevik, İrem Çekel, Gülden Sayılan, Serdar Ayaz and Ö. Ozan Evkaya for their friendship; Hacer Nalbant for the wise advices and instructions to the academic life; and lastly, my family: Melek Demirci, Saniye Demirci and Mehmet Demirci for everything on my life and for any kind and infinite amount of support, I have relearned the meaning of being a family during the preparation of this thesis.

This study was supported by Scientific and Technical Research Council of Turkey (TÜBİTAK) as a part of the project In Vivo Conservation and Preliminary Molecular Identification of Some Turkish Domestic Animal Genetic Resources – I (TURKHAYGEN-I) under the grant number 106G115.

Ovis gmelinii anatolica samples were provided by the General Directorate of Nature Conservation and National Parks, Ministry of Forestry and Hydraulic Works.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

% : Per Cent

°C : Degrees Celsius

‰ : Per Mille

$AgNO_3$: Silver Nitrate

AIC : Akaike Information Criterion

AKK : Akkaraman

APS : Ammonium Per Sulfate

Arlequin : An Integrated Software Package for Population Genetics Data Analysis

BEAST : Bayesian Evolutionary Analysis by Sampling Trees

BIC : Bayesian Information Criterion

bp : Base Pair

BP : Before Present

BSA : Bovine Serum Albumin

BSP : Bayesian skyline plot

CI : Confidence Interval

CIC : Çineçaparı

CR: Control Region

cytB: cytochrome B

DAG : Dağlıç

DNA : Deoxyribonucleic Acid

dNTP : Deoxynucleotide Triphosphate

Dxy : Average number of nucleotide substitutions per site between populations

EDTA : Ethylene Diamine Tetra Acetic Acid

ESS : Effective Sample Size

EtBr : Ethidium Bromide

g : Generation time

GOK : Gökçeada

Hd : Haplotype Diversity

HEM : Hemşin

HER : Herik

HKY : Hasegawa Kishino Yano

HPG : Haplogroup

IVE : İvesi

KDE: Kernel Density Estimation

KIV : Kıvırcık

Km : Kilometer

KRG : Karagül

KRY : Karayaka

M : Molar

MCMC : Markov Chain Monte Carlo

MDS : Multidimensional Scaling

ME: Minimum Evolution

MEGA: Molecular Evolutionary genetics Analysis

mg : Milligram

$MgCl_2$: Magnesium Chloride

MJ : Median Joining

mM : Millimolar

MP : Maximum Parsimony

MRK : Morkaraman

mtDNA: Mitochondrial DNA

ND2 : NADH Dehyrogenase Subunit 2

ND4 : NADH Dehyrogenase Subunit 4

Ne : Effective Population Size

ng : Nanogram

NJ : Neighbor Joining

NOR : Norduz

OGA : *Ovis gmelinii anatolica*

PAGE : Polyacrylamide Gel Electrophoresis

PCA : Principle Component Analysis

PCR : Polymerase Chain Reaction

pmol : Picomoles

r : Spatial Autocorreltion Coefficient

RFLP : Restriction Fragment Length Polymorphism

SAK : Sakız

SSCP : Single Strand Conformational Polymorphism

Taq : Thermus aquaticus

TBE : Tris Borate EDTA

TEMED : Tetramethylethylenediamine

TURKHAYGEN-I : In Vivo Conservation and Preliminary Molecular Identification of Some Turkish Domestic Animal Genetic Resources – I

u : Unit

UV : Ultra Violet

V: Volt

μl : Microliter

π : Nucleotide Diversity

CHAPTER 1

INTRODUCTION

Domestication of plants and animals together with the change of human life style from hunting and gathering to farming is known as the ‘Neolithic Period’ in the history of humanity. Beginning of this period was subject to many researches (for instance see Zeder *et al.*, 2006 and Kuijt, 2002). The latest evidences indicate that the animal domestication, one of the essential components in the emergence of Neolithic period, resides in southeastern Anatolia (Peter *et al.*, 2007; Vigne *et al.*, 2003; Zeder *et al.*, 2006; Zeder, 2008). The present study; covering the molecular genetic studies, employing the tools of bioinformatics (editing the DNA sequences, retrieving DNA sequences from the gene banks, using various measures and statistical methods for the understanding of the data) contributes to answering some questions related with the beginning of this period. In particular, study reveals new insights about domestication of sheep and evolution of domestic sheep. As the study material, sheep DNA from modern specimens (wild and domestic) from Anatolia were used.

1.1 Nomenclature and distribution of wild sheep

Nomenclature of wild sheep, because of many revisions that were made, is quite complex. However in a recent publication (Rezaei *et al.*, 2010) evolution of the nomenclature is summarized. The summary is presented in Table 1.1.

Table 1.1 Evolution of nomenclature in relation to wild sheep. Table is taken from Rezaei *et al*.'s (2010) study.

Authors	Tsalkin (1951)	Haltenorth (1963)	Nadler *et al.* (1973)	Valdez (1982); Wilson and Reeder (1993); Shackleton and Lovari (1997)	Festa-Bianchet (2000)
Dall Sheep	*O. canadensis / O. nivicola*	*O. ammon*	*O. dalli*	*O. dalli*	*O. dalli*
Bighorn	*O. canadensis / O. nivicola*	*O. ammon*	*O. canadensis*	*O. canadensis*	*O. canadensis*
Snow Sheep	*O. canadensis / O. nivicola*	*O. ammon*	*O. nivicola*	*O. nivicola*	*O. nivicola*
Argali	*O. ammon*	*O. ammon*	*O. ammon*	*O. ammon*	*O. ammon*
Asiatic Mouflon	*O. ammon*	*O. ammon*	*O. orientalis*	*O. orientalis*	*O. gmelinii*
Urial	*O. ammon*	*O. ammon*	*O. vignei*	*O. orientalis*	*O. vignei*
European Mouflon	*O. ammon*	*O. ammon*	*O. musimon*	*O. orientalis musimon*	*O. orientalis musimon*

The distribution of all the seven wild sheep species listed in the last column of Table 1.1 is given on the map in Figure 1.1.

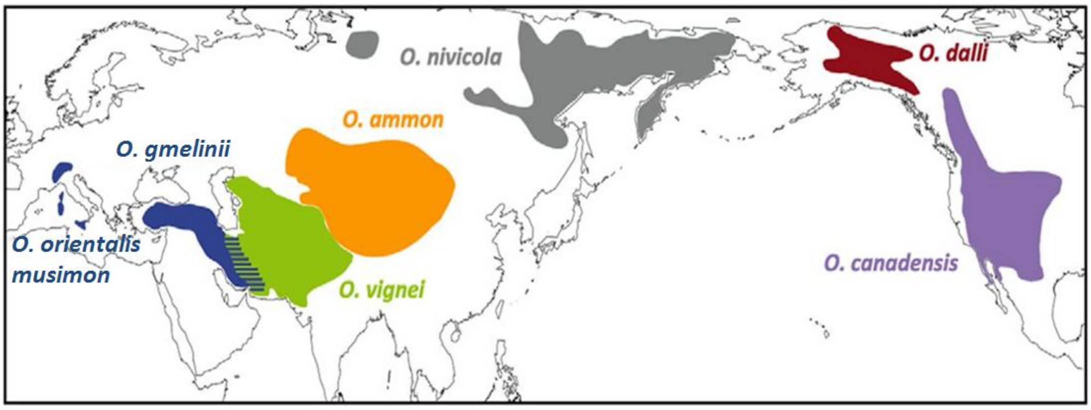


Figure 1.1 The worldwide distribution of wild sheep. The map is modified from Rezaei *et al.'s* (2010) study.

The first set of researches about the possible ancestor of domestic sheep was based on karyotypic comparisons (i.e. chromosome numbers) of different wild sheep species. Nadler *et al.* (1973) and Bunch *et al.* (1976) showed that the domestic sheep (n=54) most probably derived from *Ovis gmelinii* (Asiatic mouflon, *Ovis orientalis*) present in western Iran, southwestern Iran and Anatolia (Figure 1.1), which has 54 chromosomes. This proposition was confirmed by an early molecular genetics study (Hiendleder *et al.* 1998a) on comparative mtDNA RFLP analysis, which showed that *Ovis ammon* and *Ovis vignei* were not the ancestors of domestic sheep leaving *Ovis orientalis* as the putative ancestor.

However, hybrids of *Ovis vignei* (2n=55 to 57) and domestic sheep, or *Ovis ammon* and domestic sheep were also observed in the shaded areas as shown on Figure 1.1.

1.2 Domestication center of sheep, first and second migrations of domestic sheep

In the Merriam-Webster dictionary, domestication is defined as follows: "Domestication is the adaptation of an animal or a plant to live in intimate association with and to the advantage of humans". Archaeozoological findings showed that sheep is one of the first domesticated livestock species (Ryder, 1983). As was summarized by Zeder (2008), early archaeozoological studies considered size reduction in animal bones found in archaeological sites as a signature of the domesticated animals in the area. However, recently, the presence of relatively young males and old females in the archaeozoological finds, even in the absence of size reduction, were started to be considered as the signs of management of animal populations. In other words, these findings were the signs of first steps of the animal domestication process. In the light of this new accepted evidence, domestication centers for sheep, cattle, pig and goats were suggested to be residing largely in southeast Anatolia (Zeder, 2008). The map depicting the domestication centers for the four major livestock species (Zeder, 2008) was given in Figure 1.2.

The time for the first domestication and the wave of advancement of herding practice (as given in Figure 1.2) could be traced by the advanced dating procedures. The presence of wild forms in the domestication sites is a prerequisite for the start of domestication process. Wild sheep and goat were not present in Europe in Holocene (Clutton-Brock, 1999). The known geographical distribution of the *Ovis gmelinii* (Figure 1.1) suggests the presence of the wild sheep in the depicted domestication center (Figure 1.2).

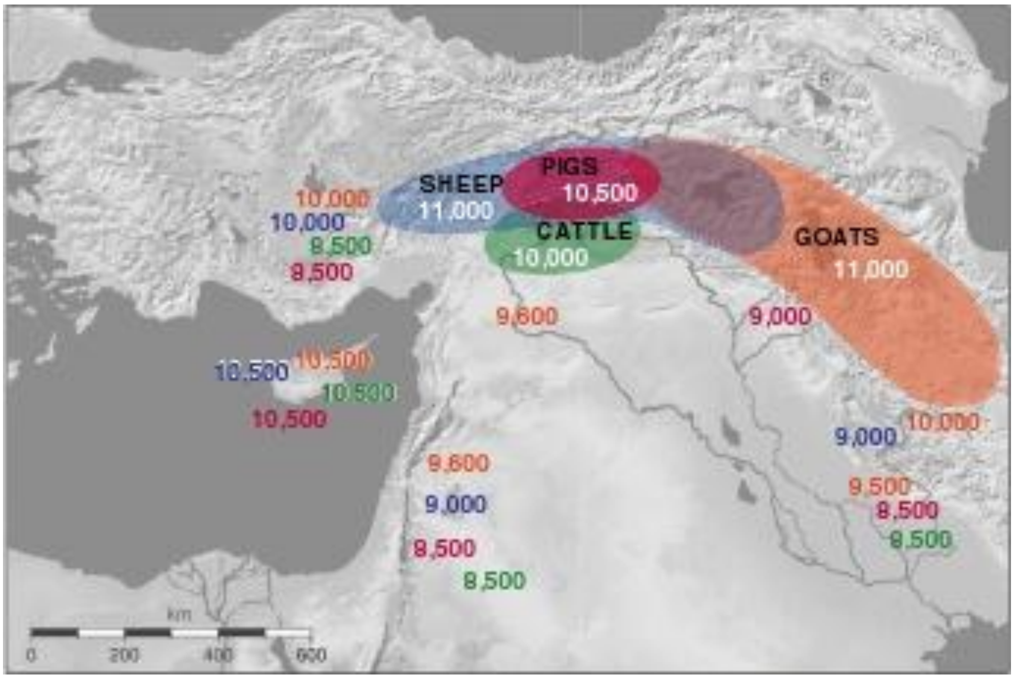


Figure 1.2 The domestication centers of sheep, goat, cattle and pigs in Fertile Crescent. Numbers on the figure shows the approximate time (in B.P.) of the beginning of domestication of animals to be seen in that region. Shaded areas shows the initial domestication area of sheep (blue), goats (orange), cattle (green), and pigs (fuscia). Figure was taken from Zeder's (2008) study.

What was the drive for animal domestication? It is still unknown. However, very recently, an archaeological site: Göbekli Tepe covered with 5 tons weighing erected stones were unearthed in Southeastern Anatolia. It might be the oldest temple of the humanity. It must have been built in the time of hunter-gatherers. The dating studies showed that the earliest level of the temple might be dating 12.000 B.P. (Schmidt, 2007). Wild animal figures such as scorpions, boars, crocodile like-animals, and ram's head were carved on the stones. It is believed that wild animals were offered during the ceremonies at Göbekli Tepe. Since, at that time animal domestication have not started yet, for the ceremonies wild animals must have been hunted and kept for some time near the temple before the ceremonies.

Göbekli Tepe is just 15 km away from Şanlıurfa province in Southeastern Anatolia region of Turkey (Schmidt, 2007) and it is at the heart of animal domestication centers shown in Figure 1.2. It is suggested that the first steps of domestication originated during these preparations for the ceremonies.

Another observation of interest is the sudden existence of wild sheep and goat of Cyprus around 10.500 BP suggesting that they were transported to the island; that is, the island was "colonized", by seafarers (Zeder, 2008). If samples of wild animals could be carried by boat about 60 km from the northern Levant, they could also be transported in and around the Fertile Crescent as was suggested for early managed herds of goat by (Naderi *et al*., 2008).

Migration of Neolithic farmers from the domestication center to the west is believed to follow two main routes: along the Mediterranean coast (Maritime route) and the Danube Valley (Dobney and Larson, 2006; Zeder, 2008; Pariset *et al.*, 2011). It is believed that through maritime (Mediterranean) route, sheep reached to western Europe. For instance, sheep reached to France by 7400 BP (Zilhão, 2001).

Mediterranean (European) mouflon (*Ovis orientalis musimon*) is considered as neolithic feral domesticate which were the descendants of firstly domesticated sheep returned to the feral. It is believed that *O. orientalis musimon* introduced to Corsica and Sardinia islands around 8000 BP (Poplin, 1979; Vigne, 1999; Hiendleder *et al.*, 2002).

Similarly, there were migrations of domestic sheep to North Africa (Barker, 2002) and Eurasia (Harris and Gosden, 1996; Price, 2000). In the present study, the migration of livestock, domesticated at the Neolithic time in the domestication center, into all directions was referred as "**the first migration of sheep".** The traces of first migration was believed to be seen among the breeds in the northern periphery of

Europe, and in the wild sheep (mouflon) of Sardinia, Corsica and Cyprus in a study based on retrovirus integration sites (Chessa *et al.*, 2009), but these markers of the first sheep migration were not observed in Asia.

In Chessa *et al.*'s (2009) study based on the distribution of 'retrovirus integration site combinations' (retrotypes), it was proposed that nearly 5000 years before the present a sheep highly praised for its secondary product such as wool (on the contrary of its primary product meat) was developed in the Middle East most likely between Iran and Iraq. This sheep went through an expansion. Consequently, another mass movement of sheep, "**the second migration of sheep**", towards all directions (from Middle East) was proposed. Furthermore, it was proposed that the extent of products of the second expansion of sheep was observed, almost exclusively, in modern sheep today.

Under these assumptions, it can be anticipated that, modern domestic sheep of Anatolia might be mostly the product of this second migration of sheep.

1.3 Populations of *Ovis gmelinii* with a special emphasize on Anatolian mouflon (*Ovis gmelinii anatolica*)

The distribution of wild sheep *Ovis gmelinii (*called as *Ovis orientalis* by some authors) is fragmented, in modern times. It exists in subpopulations. There are two subpopulations of *Ovis gmelinii* within Turkey (Figure 1.3) and at least 3 (3-5) populations in Iran (Figure1.4).

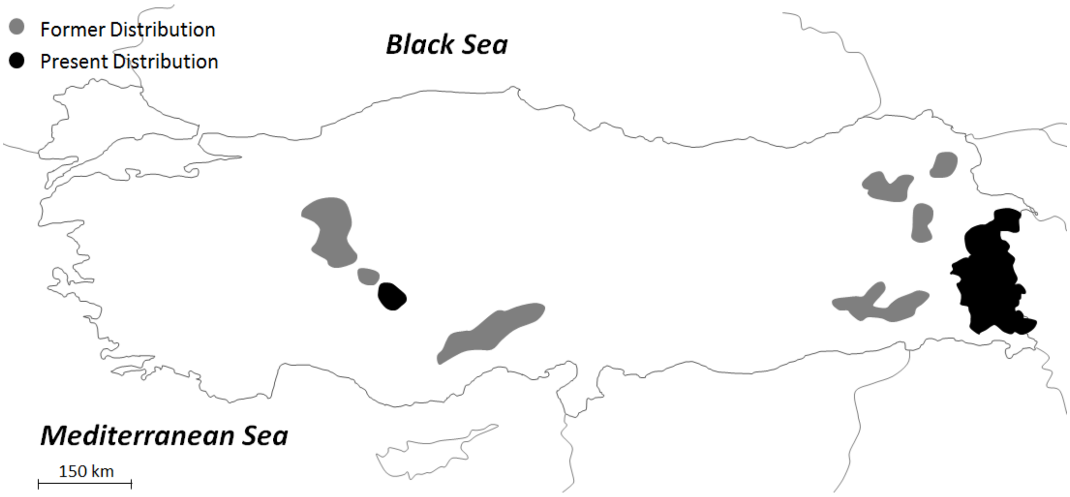


Figure 1.3 The former and present distribution of *Ovis gmelinii* populations in Anatolia. *Ovis gmelinii anatolica* (Bozdağ population) in the Central Anatolia and *Ovis gmelinii gmelinii* (Armenian mouflon) near the eastern border of Turkey. The map of Turkey is taken from Arıhan's (2000) study. Grey areas show the former, black areas show the present distributions of the populations.

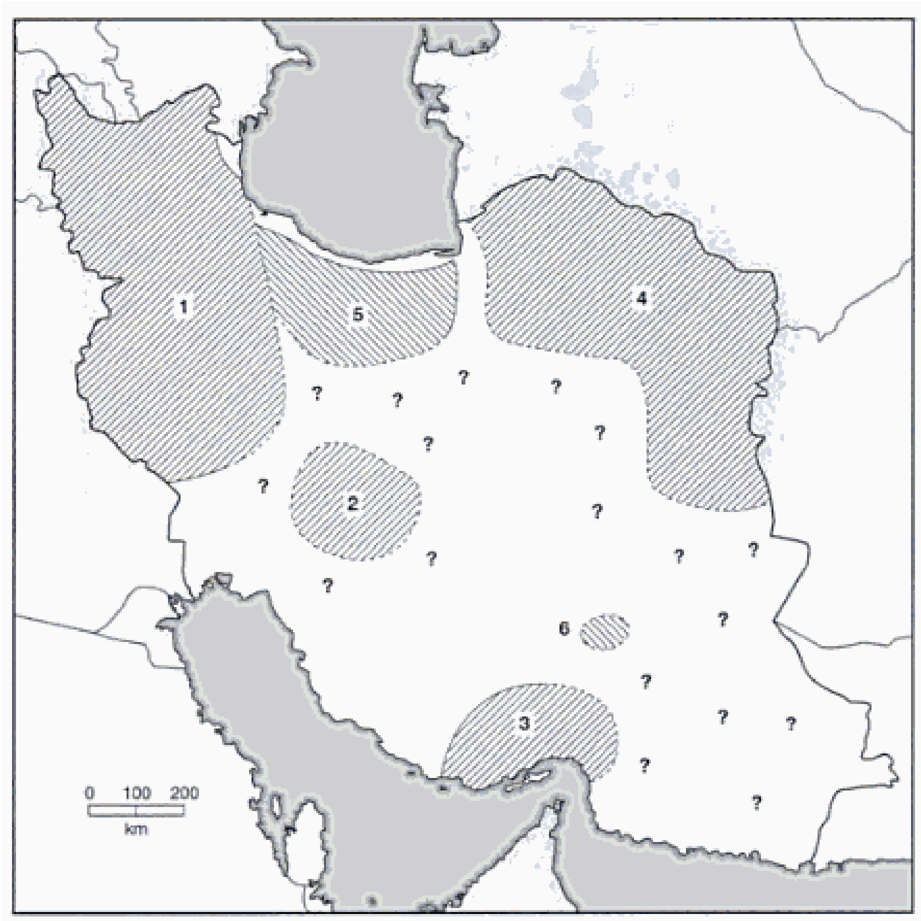


Figure 1.4 The map depicting wild sheep of Iran. 1) *Ovis gmelinii gmelinii* (Armenian Mouflon, *Ovis orientalis gmelinii*), 2) *Ovis gmelinii isphahanica (*Esfahan sheep, *Ovis orientalis isphahanica),* 3) *Ovis gmelinii laristanica* (Laristan sheep, *Ovis orientalis laristanica*), 4) *Ovis vignei* (Transcaspian urial, *Ovis orientalis arkal*) 5) Albroz red sheep (*Ovis gmelinii gmelinii* X *Ovis vignei arkali*) 6) Kerman sheep (*Ovis gmelnii laristanica* X *Ovis vignei blanfordi*). The map is taken from Ziaie's (1997) study. Population names were adapted.

Ovis gmelinii anatolica, the Bozdağ population in Central Anatolia, had a wider (Kaya *et al.*, 2004) range as was shown by the shaded areas on the map. Several factors like illegal hunting, predation, harsh weather conditions and increased pressure of domestic livestock caused *Ovis gmelinii anatolica* become extinct between 1940s and 1970s, except for the Bozdağ population (Arıhan, 2000; Sezen, 2000). The Bozdağ population was started to be protected in a wildlife protection area, which was established in 1966 in Konya-Bozdağ region with a starting population size of approximately 35 individuals (Sezen, 2000). The other population (Armenian mouflon) found in the easternmost part of Turkey and a part of the population found near the border of Iran makes seasonal migrations between the two countries (Sezen, 2000). The full change in the distribution area of *Ovis gmelinii anatolica* over the centuries is not known. Arıhan and Bilgin (2002) suggested that a barrier for the mouflon was occurred due to replacement of steppic corridor between central Turkey and western Iran by dry forest or woodland by about 6000 BP (Adams, 1997). Yet, the predicted distribution of sheep domestication center (Zeder, 2008), suggests that there was connection between *Ovis gmelinii anatolica* and *Ovis gmelinii gmelinii*. Still, "Were the *Ovis gmelinii* populations of Iran and Turkey connected with each other?" remained to be answered. However, *Ovis gmelinii anatolica* is the only *Ovis gmelinii* subspecies where ewes are hornless as it is in the domestic sheep. This is one of the reasons why it is suggested to be one of the ancestors of the domestic sheep. Furthermore, in the limits of the existing molecular data by 2002, *Ovis gmelinii anatolica* was again proposed as one of the ancestors of domestic sheep (Hiendleder *et al.*, 2002). The hypothesis remained to be attested by further molecular studies. In the present study, 30 samples from Bozdağ population were analysed based on mtDNA control region sequences.

1.4 Mitochondrial DNA (mtDNA) as an informative marker in phylogenetic and phylogeographic studies

mtDNA in particular control region (CR) shows high amount of variation in the intra-species level. Therefore, it exhibits high number of haplogroups (monophyletic group of sequences) and haplotypes (sequence types within the haplogroup) within the species. The mtDNA CR is the most widely used molecular marker in unraveling evolutionary history, phylogeny, of many animals (Bruford *et al.*, 2003). Since mtDNA is in haploid number it has low effective population size and therefore, it is subjected to genetic drift easily. Especially during migrations some haplotypes/haplogroups are lost or emerged and hence a genetic structure over the geography is generated. By this property of mtDNA, geographic patterns of genetic diversity can be easily visualized (for instance see Bruford and Townsend 2006). Geographic pattern in turn can be evaluated in the context of phylogenetics enabling phylogeographic interpretations. Furthermore, both haplogroup and sequence data obtained by different laboratories can be combined for joint analysis. There is an ample amount of mtDNA CR data from various parts of the distribution range of sheep in the online databases (for instance National Center for Biotechnology Information: NCBI, GenBank) that can freely be accessed. So, even if the collected data is from a limited region (such as Anatolia) results can be evaluated in almost the global scale (Olivieri *et al.*, 2012). If the time of divergence between the lineages is going to be estimated, some recurrent mutations may confound the estimations. Therefore, mtDNA cytochrome *B* (*cytB*) region is also used as an informative marker as its mutation rate is lower than that of CR. Diversity within this region is used in molecular clock analysis and/or to calculate the time to the most recent common ancestor (tMRCA) and in phylogenetic analysis to obtain more conservative, robust phylogenetic networks/trees.

According to mtDNA *cytB* region, Rezaei *et al.* (2010) draw a phylogenetic tree, in which *O. nivicola*, *O. dalli*, and *O. canadensis* diverged from other three wild species which were *O. ammon, O.vignei,* and *O. orientalis. Ovis gmelinii anatolica* (called as *Ovis orientalis anatolica* in Rezaei *et al.*'s (2010) study) clustered with *Ovis orientalis* samples as was predicted.

There are mtDNA markers other than the CR or *cytB* regions that can be employed for the haplogroup determination. For example, mtDNA ND2 and ND4 regions were used for haplogroup determination in domestic sheep in Guo *et al*.'s (2005) study. In a recent study (Meadows *et al*., 2011), whole mtDNA sequences were obtained from representatives of domestic and wild sheep populations to investigate the haplogroup discrimination ability of various mtDNA segments. Results indicated that the CR is the most useful marker (Meadows *et al.*, 2011). That study did not include the sequencing of *Ovis gmelinii* (*Ovis orientalis)*, but based on the results, it remained as the candidate progenitor of domestic sheep. Although, mtDNA *cytB* sequences of several *Ovis gmelinii anatolica* individuals have been reported no data is available for their mtDNA CR.

Of course, it must be remembered that, in driving at the conclusions about the evolutionary history of sheep domestication and domestic sheep, it is just one molecule and tells the story only from the female side as it is maternally inherited.

1.5 mtDNA haplogroups and their distributions in domestic sheep

Based on mtDNA, five haplogroups were observed in sheep, in previous studies. Firstly HPG A and HPG B were observed in sheep of New Zealand by Wood and Phua (1996). Then, Hiendleder *et al.* (1998a) confirmed the presence of these two haplogroups by control region sequences of the samples obtained from Russia, Kazakhstan and Germany. The third haplogroup, HPG C, was detected in sheep sampled from Turkey by Pedrosa *et al.* (2005), in China by Guo *et al*. (2005) and Chen *et al*. (2006), and in Portugal by Pereira *et al*. (2006). These three haplogroups, as they are widely present and high in frequency, considered as major haplogroups. One of the remaining two haplogroups, HPG D, was observed in north Caucasus in breed Karachai by Tapio *et al*. (2006) and in two individuals of breed Morkaraman by Meadows *et al.* (2007) in Turkey. The last haplogroup, HPG E was observed in Israel in breed Awassi (Turkish name is Ivesi) and in Turkey in breed Tuj by Meadows *et al.* (2007); in Turkey in breed Karayaka (at first denoted as C* as this lineage clustered near HPG C) by Pedrosa *et al.* (2005); in China in breeds Mongolian and Hu (at first as highly deviats of HPG C) by Guo *et al.* (2005). HPG D and E considered as rare haplogroups, as they are present only in the Middle East and Asia but with a higher occurance in Turkey among the studied geographic regions. Still, they are observed in very few samples: 3 for HPG D, 7 for HPG E.

In a seminal paper by Luikart *et al*. (2001) several haplogroups, rich in haplotypes, were observed for the domestic goats. It was argued that unless an unrealistically large wild goat population was sampled at the start of goat domestication, these haplogroups would have referred to the independent domestication events and then the time of expansions (presumably corresponding to the domestication times of haplogroups) for different haplogroups was estimated (Luikart *et al.*, 2001). However, observing that all the haplogroups seen in domestic goats are present in wild goats of the domestication site and that the presence of a haplogroup in France earlier than its predicted expansion time (Fernández *et al*., 2006). it is proposed that a single domestication event may involve expansion of more than one haplogroup (Naderi *et al.*, 2008).

Since, during the domestication process, few animals were used to obtain large domestic flocks signitures of this demographic event is seen in mtDNA. The population expansion of two major haplogroups, HPG A and HPG B were revealed by Meadows *et al*. (2007). Also for the third major haplogroup, HPG C, population expansion was first detected by Chen *et al.* (2006). The expansions of major haplogroups were expected for domestic animals. Demographic history of rare haplogroups could not be tested based on molecular data as the sample size of these haplogroups was not enough to carry out the analyses.

The domestic sheep breeds of today carries mixture of mtDNA haplogroups, but the HPG B dominancy in Europe and the HPG A dominancy in Asia (Tapio *et al*., 2006; Bruford and Townsend, 2006) are evident. The presence of relatively high frequency of HPG C in southwest Asia (Bruford and Townsend, 2006) as well as in Turkey (Pedrosa *et al.*, 2005) together with HPG D and E were observed (Meadows *et al*., 2007). The distribution of sheep haplogroup frequencies over the Euroasia must be shaped up by the domestication and expansion events, subsequent migrations, selections, admixtures and introgressions. A migration of sheep perhaps at the time of Andalucían's from Near East to Iberian Peninsula through Mediterranean route was well documented by the high diversity observed in mtDNA haplogroups. Incidentally, the distribution of fat tailed sheep almost exclusively found in Asia seemed to be associated with the existence of A and C haplogroups and thin tailed sheep of Europe seemed to be associated with HPG B. Although, both mtDNA CR (Pedrosa *et al*., 2005; Meadows *et al*. 2007,; Peter *et al*., 2007) and *cytB* region (Meadows *et al*., 2007) of native sheep breeds of Turkey have been examined extensively, the studies did not include all the breeds of Turkey. For example, Ivesi (Awassi) breed closest to the domestication and the second expansion sites of sheep, and the two of the thin tailed breeds of northwestern Turkey, namely Kıvırcık and Gökçeada, have not been examined yet.

1.6 Objectives of the study

Presence of wild sheep *Ovis gmelinii* in Anatolia, yet lack of information from its mtDNA CR (which is the most informative region on mtDNA), the location of predicted sheep domestication center in the southeastern Anatolia, proximity of the Anatolian native breeds to this center and to the origin of the second migration of sheep, makes the molecular research of both domestic and wild sheep of Anatolia very interesting.

Overall the purpose of the study is to contribute to the history of sheep domestication and domestic sheep.

Objectives of the study can be summarized as follows:

i) In the present study 628 samples from 13 sheep breeds, mostly native breeds, covering whole Turkey will be examined with respect to their haplogroup compositions where haplogroups will be determined based on mtDNA ND2 region with single strand conformational polymorphism method (SSCP) with the following purposes:

    1) To understand the spatial pattern of haplogroup diversity of domestic sheep in Turkey. It is expected that the pattern will display the signs of first and second migrations of sheep. The pattern can be used in conservation studies of domestic sheep in Turkey.

    2) To understand the demographic properties of E and D. All the individuals who are possibly belonging to HPG E and D will be sequenced to increase the number of observed sequences for HPG E and HPG D.

    3) To identify the individuals to be sequenced so that by few sequences haplogroups from all of the breeds will be presented

ii) To find out the goodness of fit between the haplogroup determination of mtDNA CR sequences and ND2 SSCP method; ND2 SSCP based haplogroups of selected samples will be confirmed by mtDNA CR sequences. Information can be useful to understand if these two methods can be used interchangeably.

iii) mtDNA CR of *Ovis gmelinii anatolica* (n=30) and individuals (n=240) of diverse haplotypes from 13 domestic sheep breeds of Anatolia will be sequenced (4 *Ovis gmelinii anatolica* individuals will also be sequenced for their *cytB* region)

The results will be used:

1) To find the genetic relatedness of *Ovis gmelinii anatolica*, the Anatolian wild sheep, to the domestic sheep samples from Anatolia.

2) For the comparative analysis of the *Ovis gmelinii anatolica* and domestic sheep samples from Anatolia with the data retrieved from the database on a worldwide geographical scale.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

Blood samples from 628 sheep of 13 sheep breeds were collected within the framework of the project TURKHAYGEN-1 (www.turkhaygen.gov.tr; Project no: 106G115). The collected breeds were Sakız (n=49), Karagül (n=50), Hemşin (n=48), Çine Çaparı (n=40), Norduz (n=46), Herik (n=49), Dağlıç (n=50), Morkaraman (n=50), Kıvırcık (n=45), Karayaka (n=50), Gökçeada (n=50), İvesi (n=51), and Akkaraman (n=50). To represent gene pools of the breeds, 3-5 samples (if possible) from each flock for each breed were sampled. The map of Turkey in Figure 2.1 shows the 106 locations of the sampled flocks and the map in Figure 2.2 shows the provinces where breeds were sampled from.

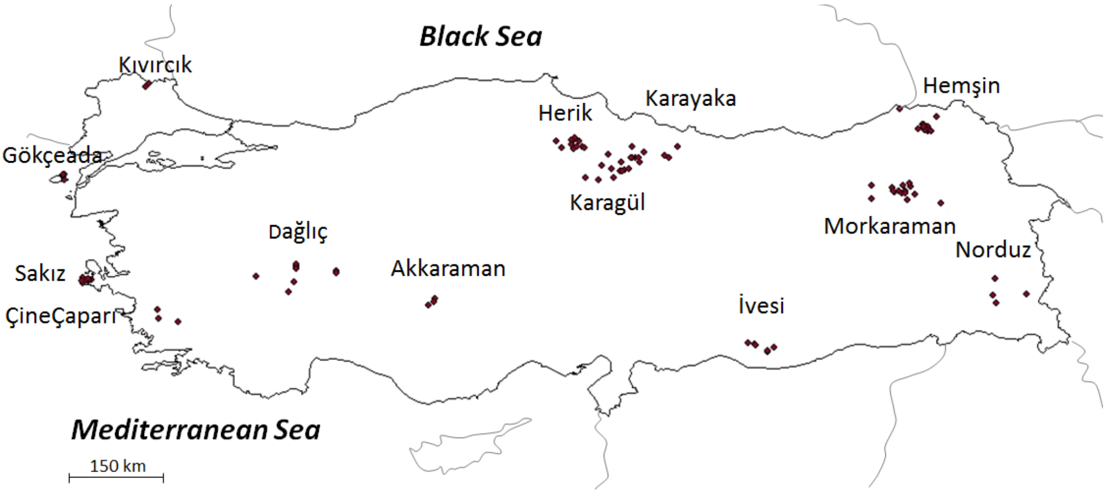


Figure 2.1 Locations of sampled flocks on the map of Turkey. Map was constructed with ArcMap as a product of ArcGIS 10 software (www.esri.com).

Figure 2.2 Sampling sites of breeds in relation to provinces on the map of Turkey. Picture taken from Doğan (2009).

30 Anatolian mouflon (*Ovis gmelinii anatolica*) blood samples were collected from Konya-Bozdağ region of Central Anatolia by the General Directorate of Nature Conservation and National Parks, Ministry of Forestry and Hydraulic Works.

In the the present study, some of the previously deposited mtDNA control region (CR) sequences were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/) to be used as reference sequences for the haplogroups. Their accession numbers, breed identities, the haplogroups that they represent and the references of these sequences were given in Table 2.1.

Table 2.1 mtDNA CR sequences retrived from GenBank. A-E represent sheep mtDNA haplogroups.

Accession numbers	Sample name	Haplogroup	Country	Breed	Reference
HM236183	tj6	E	Turkey	Tuj	Meadows *et al*., 2011
HM236182	aw25	E	Israel	Awassi	Meadows *et al*., 2011
HM236181	mk9	D	Turkey	Morkaraman	Meadows *et al*., 2011
HM236180	mk3	D	Turkey	Morkaraman	Meadows *et al*., 2011
HM236179	mk4	C	Turkey	Morkaraman	Meadows *et al*., 2011
HM236178	kk12	C	Turkey	Karakas	Meadows *et al*., 2011
HM236177	kk2	B	Turkey	Karakas	Meadows *et al*., 2011
HM236176	kk1	B	Turkey	Karakas	Meadows *et al*., 2011
HM236175	R359	A	England	Romney	Meadows *et al*., 2011
HM236174	cl122	A	Germany	Merino	Meadows *et al*., 2011
DQ242212	krc6	D	north Caucasus	Karachai	Tapio *et al.,* 2006
AY829404	MG024	E	China	Y G HOU	Guo *et al.,* 2005
AY829385	HSK018	E	China	CH R HU	Guo *et al.,* 2005
DQ097468	KAR15	E	Turkey	Karayaka	Pedrosa *et al.*, 2005
HM042760	TR0005	E	Turkey	-	Koban *et al.*, unpublished
HM042761	TR0006	E	Turkey	-	Koban *et al.*, unpublished
HM042785	TR0030	E	Turkey	-	Koban *et al.*, unpublished
HM042838	TR0083	E	Turkey	-	Koban *et al.*, unpublished
AY091490	Ovis vignei bochariensis 1	Outgroup/ Wild sheep	Turkmenistan	Urial	Hiendleder *et al*., 2002
AY091491	Ovis vignei bochariensis 2	Outgroup/ Wild sheep	Turkmenistan	Urial	Hiendleder *et al*., 2002
AF242347	Ovis ammon ammon	Outgroup/ Wild sheep	Mongolia	Argali	Hiendleder *et al*., 2002
AF242348	Ovis ammon darwini	Outgroup/ Wild sheep	Mongolia	Argali	Hiendleder *et al*., 2002

2.2 Laboratory experiments

2.2.1 DNA extraction from blood samples

Blood samples from 628 domestic sheep from 13 breeds and from 30 *Ovis gmelinii anatolica* were isolated with standard phonol: chloroform DNA extraction protocol according to modified (Koban, 2004) version of Sambrook et al.'s (1989) study. The extraction procedure was given in Acar's (2010) study.

The concentration and quality of extracted DNA were measured with Thermo Scientific NanoDrop 2000c Spectrophotometer.

2.2.2 Conditions for Polymerase Chain Reaction (PCR)

2.2.2.1 Amplification of a part of mitochondrial DNA (mtDNA) ND2 region

A 205 base pair (bp) long part of sheep mtDNA ND2 gene (AF010406 positions 4053 to 4258) was amplified by OV6 primers which were used by Guo *et al.*'s (2005) study. Sequences of these primers were as given below.

OV6-Forward: 5'-CAACCCACGAGCCACAGAAG-3'

OV6-Reverse: 5'-CTGGGACTCAGAAGTGGAATGG-3'

Table 2.2 presents the concentrations of the PCR mixture and Table 2.3 presents the PCR reaction conditions for the mtDNA ND2 region.

Table 2.2 PCR mixture for the mtDNA ND2 region.

PCR Buffer	1X
$MgCl_2$	2.5mM
dNTP	0.2mM
Primer	10pmol
Taq Polymerase	0.5u
DNA	50ng-100ng
Total Volume	**15µl**

Table 2.3 PCR conditions for the ND2 region.

Step	**Temperature**	**Duration**	**Number of Cycles**
Denaturation	94°C	3 minutes	1
Denaturation	94°C	30 seconds	30
Annealing	57°C	30 seconds	
Extension	72°C	75 seconds	
Final Extension	72°C	15 minutes	1

2.2.2.2 Amplification of mtDNA control region (CR)

A 1501 bp long part of sheep mtDNA (AF010406 positions 15271 to 156), encompassing control region (CR), 1180 bp long (AF010406.1 positions 15437 to 16616), was amplified by primers designed by Evren Koban using Hot-start PCR. Sequences of these primers are given below.

Forward (CRF): 5'- CATCGAAAACAACCTCCTCAA -3'

Reverse (CRR):5'- GATTCGAAGGGCGTTACTCA -3'

Table 2.4 presents the concentrations of thePCR mixture and Table 2.5 presents the PCR conditions for the mtDNA CR.

Table 2.4 PCR mixture for the mtDNA CR.

PCR Buffer	1X
$MgCl_2$	2mM
dNTP	0.2mM
Primer	10pmol
BSA	48ng/ml
Taq Polymerase	1u
DNA	50ng-100ng
Total Volume	**25µl**

Table 2.5 PCR conditions for the mtDNA CR.

Step	**Temperature**	**Duration**	**Number of Cycles**
Denaturation	95°C	3 minutes	1
Denaturation	95°C	30 seconds	35
Annealing	52°C	45 seconds	
Extension	72°C	90 seconds	
Final Extension	72°C	15 minutes	1

2.2.2.3 Amplification of mtDNA cytochrome *B* (*cytbB*) region

A part of sheep mtDNA, 1272 bp long (AF010406.1 positions 14078 to 15349), encompassing cytochrome *B* (*cytbB*) region, 1140 bp long (AF010406.1 positions 14159 to 15298), was amplified by *CytB* primers which were used by Meadows *et al.*'s (2005) study. Sequences of these primers are given below.

CytB-F 5'-GTCATCATCATTCTCACATGGAATC-3'

CytB-R 5'-CTCCTTCTCTGGTTTACAAGACCAG-3'

Table 2.6 presents the concentrations of thePCR mixtureand Table 2.7 presents the PCR conditions for the mtDNA *cytB* region.

Table 2.6 PCR mixture for the mtDNA *cytB* region.

PCR Buffer	1X
$MgCl_2$	1mM
dNTP	0.2mM
Primer	10pmol
Taq Polymerase	0.5u
DNA	50ng-100ng
Total Volume	**20µl**

Table 2.7 PCR conditions for the mtDNA *cytB* region.

Step	Temperature	Duration	Number of Cycles
Denaturation	95°C	3 minutes	1
Denaturation	94°C	50 seconds	35
Annealing	54°C	60 seconds	
Extension	72°C	60 seconds	
Final Extension	72°C	5 minutes	1

2.2.3 Single Strand Conformational Polymorphism (SSCP) analysis of mtDNA ND2 region

Single Strand Conformational Polymorphism (SCCP) method is based on identifying new conformations of single-stranded DNA, based on the mutations on the sequences defining their haplogrups, after separating double strands of the DNA (Klug and Cummings, 2000).

Guo *et al.*'s (2005) study was followed for the SSCP analysis. Equal amounts of amplified ND2 region PCR product was mixed with formamide loading dye (98% formamide, 10 mM $EDTANa_2$, 0.02% xylene cyanol, 0.02% bromophenolblue and adjusted by distilled water) and the mixture was incubated in 98ºC for 10 minutes to separate double strands. After incubation, mixture was placed on ice to stop the denaturing reaction. Results were visualized with Polyacrylamide Gel Electrophoresis (PAGE).

Non-denaturing polyacrylamide gel was prepared by dissolving 10% acrylamide with bisacrylamide (39:1) in 1X TBE. 30µl tetramethylethylenediamine (TEMED) and 60µl 10% ammonium persulfate (APS) were added to the 10ml polyacrylamide solution as an activator and an initiator respectively. Gel was poured between vertical glasses and waited for polymerization. After polymerization of the polyacrylamide, gel samples were loaded to the wells. Gel was run with 75V for 14-16 hours in 1X TBE by vertical electrophoresis. Bands were visualized by silver staining.

In silver staining method (Bassam *et al*., 1991), first, gel was incubated in 5% acetic acid solution for 10 minutes and then 10% ethanol solution for 5 minutes. Immediately after this step, gel was incubated in 1X silver nitrate solution ($AgNO_3$) for 25 minutes. Then polyacrylamide gel was washed with deionized water for 90 seconds and incubated in developer solution (0.75M NaOH and 37% formaldehyde) until the bands were visualized. Photograph was taken under the UV light, by a gel imaging system.

2.2.4 Sequencing of mtDNA CR and mtDNA *cytB* region

After checking quality and quantity of amplified mtDNA CR and *cytB* fragments, to remove any remaining unused PCR master mix components (i.e. salts, primers, Taq Polymerase and genomic DNA), the fragments were purified by using Roche PCR Purification Kit with the provided protocol.

Purified fragments were sequenced using Sanger *et al*.'s (1973) chain termination method. For sequencing, Forward (CRF) and Reverse (CRR) primers as well as SOAD (Wood and Phua, 1996) and HC2 (Townsend, 2000) internal primers were used. The sequencing reactions were carried out using ABI Prism BigDye Terminator and the data were collected using ABI Prism 3100^{TM} DNA Analyser by the company called RefGen (http://refgen.com.tr/).

The sequences obtained from each primer were assembled and the consensus sequence of each sample was exported in FASTA format by using Chromas Pro version 1.5 http://www.technelysium.com.au/ChromasPro.html). After all samples were obtained, progressive multiple sequence alignment was performed on these sequences by ClustalW algorithm (Thompson *et al.*, 1994), which was implemented within Bioedit version 7.1.3 (Hall, 1997-2011).

2.3 Statistical data analyses

2.3.1 Identification of nucleotide substitution model

Identifying a nucleotide substitution model (also referred as sequence evolution model) is essential to accomplish a reliable topology between sequences. In the present study, nucleotide substitution models for mtDNA CR and *cytB* region were determined by jMODELTEST (Posada, 2008) which identifies the appropriate model according to five criteria including Akaike Information Criterion (AIC) (Akaike, 1974) and Bayesian Information Criterion (BIC) (Schwarz, 1978). First, jMODELTEST uses Pyhml software (Guindon and Gascuel, 2003) to estimate maximum likelihood phylogenies for a given sequence set. Likelihoods were calculated for 88 models which include 6 core substitution models with/without invariable site (+I), gamma distributed rate heterogeneity (+G) options and variable number of parameters. These core substitution models are JC (Jukes and Cantor, 1969), HKY (Hasegawa, Kishino and Yano, 1985), TN (Tamura and Nei, 1993), K81 (Kimura, 1981), TIM (Posada, 2003) and GTR/REV (Tavaré, 1986). Then selection criteria applied to substitution models according to formulae were given below.

AIC formula is as follows:

$$AIC = -2l + 2K \tag{2.1}$$

where l is the maximum log-likelihood value of the data under that model and K is the number of free parameters in the model, including branch lengths if they were estimated *de novo*. The smallest AIC preferred for that model since AIC was thought as the amount of information lost when a specific model was used to approximate the actual process of molecular evolution.

BIC formula is as follows:

$$BIC = -2l + K \log n \tag{2.2}$$

where l is the maximum log-likelihood value of the data under that model, K is the number of free parameters in the model and n is the sample size.

2.3.2 Neighbor joining (NJ) tree

Neighbor joining (NJ) method is a fast clustering algorithm developed by Saitou and Nei (1987). It produces an unrooted tree as it does not require the assumption of a constant rate of evolution. In the tree constructed by NJ method, an outgroup taxon is essential for defining the root. NJ method is also regarded as simplified version of minimum evolution (ME) method (Rzhetsky and Nei 1992), which uses distance measures that correct for multiple hits at the same sites. In other words, the correct topology was chosen according to the smallest value of the sum of all branches (S) in ME method. Since ME method evaluates S values for all topologies, it is time-consuming and with increasing number of possible topologies, examining all topologies could be more difficult. However, in NJ method, the S value is not computed for all or many topologies. Instead the examination of different topologies is embedded in the algorithm, so that only one final tree is produced.

In the present study, trees reconstructed by using MEGA 5 (Tamura *et al.*, 2011) based on NJ algorithm. Bootstrap values set as 1000 and the nucleotide substitution model determined by jModeltest was used.

2.3.3 Sequence based haplogroup diversity measures

2.3.3.1 Nucleotide diversity (π)

Nucleotide diversity, π, also called the average number of nucleotide differences per site between two sequences is a measure of genetic diversity of a population. π is simply heterozygosity at nucleotide level in randomly mating populations (Nei, 1987).

The formula of nucleotide diversity (Nei 1987, equation 10.6) was shown in equations 2.3.

$$\pi = \sum_{i<j} \pi_{ij} / \binom{n}{2} \tag{2.3}$$

where n and π_{ij} are sample size and the proportion of nucleotide differences per nucleotide site between the i^{th} and j^{th} sequences, repectively. Combinatorial (n, 2) is the number of pairwise comparisons. As it is calculated upon proportions, it is a sequence length independent measure. By using the π value, comparisons among different parts of a particular DNA region can be possible.

The standard deviation of π is the square root of the variance and the formula for variance of π (Nei 1987, equation 10.7) were given in equation 2.4.

$$V(\pi) = \frac{4}{n(n-1)} \begin{bmatrix} (6-4n)\left(\sum_{i<j} x_i x_j \pi_{ij}\right)^2 \\ +(n-2)\sum x_i x_j x_k \pi_{ij}\pi_{ik} + \sum_{i<j} x_i x_j \pi_{ij}^2 \end{bmatrix} \tag{2.4}$$

where $x_i$and x_j are the respective frequencies of i^{th} and j^{th} haplotypes.

In the present study, Nucleotide diversity and its standard deviation were calculated using DnaSP v5 (Librado and Rozas, 2009) software for each mtDNA based haplogroup.

2.3.3.2 Haplotype diversity (Hd)

Haplotype diversity (Hd) simply measures the probability of occurrences of different haplotypes of two sequences when they were selected randomly in a data set (population). The level of heterozygosity in a population can be estimated by this measure.

The formulas of Hd (Nei 1987, equation 8.4) and its variance (Nei 1987, equation8.12) were given (with a replacement of 2n by n) in equations 2.5 and 2.6, respectively. The standard deviation of Hd was calculated by taking the square root of the variance.

$$Hd = \frac{n}{n-1}(1 - \sum x_i{}^2) \tag{2.5}$$

$$V(Hd) = \frac{1}{n(n-1)}\left\{\begin{matrix} 2(n-2)\left[\sum x_i^3 - \left(\sum x_i^2\right)^2\right] \\ +\sum x_i^2 - \left(\sum x_i^2\right)^2 \end{matrix}\right\} \quad (2.6)$$

where x_i is the population frequency of a haplotype and n is the sample size.

In present study, Haplotype diversity and its standard deviation were calculated using DnaSP v5 (Librado and Rozas, 2009) software for each mtDNA haplogroup.

2.3.3.3 Average number of nucleotide substitutions per site between populations (Dxy)

Average number of nucleotide substitutions per site between populations (Dxy) is used to estimate the amount of DNA divergence between population (Nei, 1987).

The formula of Dxy (Nei 1987, equation 10.20) was shown in equation 2.8.

$$D_{xy} = \sum_{ij} \hat{x}_i \hat{y}_j d_{ij} \quad (2.8)$$

where $\hat{x}_i$and $\hat{y}_j$are the population frequencies of i^{th} and j^{th} haplotype for populations X and Y respectively, and d_{ij} is the nucleotide substitutions between the i^{th} haplotype from population X and j^{th} haplotype from population Y.

2.3.4 Median joining (MJ) network

Median joining (MJ) network (Bandelt *et al.* 1999) simply demonstrates the closest haplotypes connecting each other in a measure of mutation number. More theoretically, MJ method combines features of two algorithms: Kruskal's (1956) algorithm and Farris's (1970) maximum-parsimony (MP) heuristic algorithm; the former finds minimum spanning trees by favoring short connections, the latter sequentially adds new vertices called ''median vectors'' (in this case sequences). Unlike MP method, MJ method does not resolve the ties. MJ method uses recombination-free population data (such as mtDNA).

In the present study, MJ networks were reconstructed by NETWORK 4.6.1.0 (http://www.fluxus-engineering.com/) using the default parameters, which are ε=0; weight of the sites = 10. Also, star contraction method (Forster *et al.*, 2001) were applied prior to MJ algorithm to simplify the network where sample size is large (for example, n > 100).

2.3.5 Mismatch distribution

The distribution of the number of pairwise differences between sequences (also called as mismatch distribution) (Harpending *et al.*, 1998) is an informative method from the aspect of both population diversity and population demographic history. The shape of distribution provides information about population history, particularly about population expansion (Jobling *et al.*, 2004). A smooth bell-shaped distribution indicates a sudden population expansion while the multimodal, ragged distribution indicates a constant population size. The mean number of pairwise differences in mismatch distribution is a measure of genetic diversity for that population. Moreover, by comparing means of pairwise distributions of populations showing a bell-shaped distribution (i.e. rapid population growth), the relative expansion times of populations can be estimated when the mean values were calculated from same DNA region. The larger value of mean corresponds to the earlier population expansion.

In the present study, demographic population expansion of haplogroups were investigated and the mean of pairwise differences of haplogroups were calculated by Arlequin 3.11 (Excoffier, 2005).

2.3.6 Neutrality tests

Neutrality tests, also called tests for selection, are used extensively to reveal the past population history. These tests compare the observed diversity of a population to the expected under neutral evolution. In the present study Tajima's *D* (1989) and Fu's F_S (Fu, 1997) tests were applied to reveal population history of haplogroups.

2.3.6.1 Tajima's *D*

Tajima's *D* test (Tajima, 1989) is based on the infinite-site model without recombination, appropriate for short DNA sequences. It compares two estimators of the mutation parameter theta ($\theta = 2Mu$, with M=2N in diploid populations or M=N in haploid populations of effective size N). The test statistic *D* is then defined as

$$D = \frac{\hat{\theta}_{\pi} - \hat{\theta}_S}{\sqrt{Var(\hat{\theta}_{\pi} - \hat{\theta}_S)}} \quad (2.9)$$

where $\hat{\theta}_{\pi} = \hat{\pi}$, $\hat{\theta}_S = S / \sum_{i=0}^{n-1}(1/i)$ and S is the number of segregating sites in the sample. The limits of confidence intervals around *D* may be found in Table 2 of Tajima's (1989) study for different sample sizes.

In the present study, Tajima's *D* statistics were calculated by Arlequin 3.11 (Excoffier, 2005). The significance of the *D* statistic was tested by generating random samples (sample number, n=10000) under the hypothesis of selective neutrality and population equilibrium, using a coalescent simulation algorithm adapted from Hudson (1990).

Although Tajima's *D* statistics is a measure of selective neutrality, the significant *D* values can be due to factors like population expansion, bottleneck, or heterogeneity of mutation rates (Tajima, 1993; Aris-Brosou and Excoffier, 1996; Tajima, 1996).

2.3.6.2 Fu's F_S

Like Tajima's *D* (Tajima, 1989) test, Fu's F_S test (Fu, 1997) is based on the infinite-site model without recombination, and thus appropriate for short DNA sequences. In Fu's F_S test, first, the probability of observing a random neutral sample with a number of alleles similar or smaller than the observed value given the observed number of pairwise differences, which were taken as an estimator of θ, were assessed and formulated as

$$S' = Pr\left(K \geq k_{obs} \middle| \theta = \hat{\theta}_{\pi}\right). \tag{2.10}$$

Then F_S statistic defined as the logit of S'

$$F_S = \ln\left(\frac{S'}{1-S'}\right). \tag{2.11}$$

In the present study, Fu's Fs statistics were calculated by Arlequin 3.11 (Excoffier, 2005). To determine the statistical significance 10000 random samples were used. As in the case of Tajima's *D* significance, the significance of the F_S statistic is tested by generating 10000 random samples. The P-value of the F_S statistic is then obtained as the proportion of random F_S statistics less or equal to the observation.

The F_S statistic was very sensitive to population demographic expansion, which generally leads to large negative F_S values (Fu, 1997).

2.3.7 Bayesian skyline plot (BSP)

Bayesian skyline plot (BSP) (Drummond *et al.*, 2005) demonstrates changes in effective population size through time. The analysis calculates coalescent-based tree with specified number of discrete changes (m) in the population history. It will then estimate a demographic function that has the specified number of steps integrated over all possible times of the change-points and population sizes within each step to calculate a function of Ne through time (Drummond *et al.*, 2005).

In the present study, BSPs were generated by BEAST v1.6 (Drummond and Rambaut, 2007). As tree prior, Coalescent: Bayesian Skyline; as a substitution model, gamma distributed HKY model were set and other parameters left as default, which include m=10, fixed clock rate=1. By setting clock rate as 1, the time axis of BSP was represented as the number of mutations per nucleotide site. Results are analyzed with Tracer v1.4 (Rambaut and Drummond, 2007).

2.3.8 Spatial autocorrelation

Spatial autocorrelation analysis investigates the relationship between geographic distance and genetic distance among given samples/populations. It simply measures whether genetic data were correlated with distance, in a given distance class. Euclidean distance was used to calculate distance between locations of samples.

The spatial autocorrelation coefficient (*r*), defined by Smouse and Peakall (1999)

$$r = \frac{2\sum x_{ij}c_{ij}}{\sum x_{ij}c_{ii} + \sum x_{ij}c_{jj}} \qquad (2.12)$$

where c_{ij}, $c_{ii}$and c_{jj} are the respective elements of the covariance matrix, and x_{ij}, x_{ii} and x_{jj} are the respective elements of X matrix. Each distance class is represented by its own X matrix, with elements having the value 1 when that specific pairwise comparison falls within the distance class, all other elements have the value 0. The covariance matrix is derived from the genetic distance matrix (following Smouse and Peakall 1999) by the formula:

$$C_{ij} = \frac{1}{2}\left[-{d_{ij}}^2 + \frac{1}{N}\left(\sum_{i=1}^{N} {d_{ij}}^2 + \sum_{j=1}^{N} {d_{ij}}^2\right) - \frac{1}{N^2}\left(\sum_{i \neq j}^{N} {d_{ij}}^2\right) \right] \qquad (2.15)$$

where d_{ij} is genetic distance between samples.

The coordinates for each sample, which were used for spatial autocorrelation analysis, were given in Appendix A.

Spatial autocorrelation analyses were calculated with GENALEX 6 (Peakall and Smouse, 2006) from individual mtDNA haplogroup relations of each sample.

2.3.9 Construction of synthetic map

Synthetic maps are used to visualize the genetic trends of populations on a geological map. The construction of a synthetic map in the present study requires three steps: 1) Calculation of pairwise genetic distance between populations (in this case, breeds) 2) Reducing the pairwise matrix data into a vector (one-valued data for each breed) 3) Construction of synthetic map from these values by using either interpolation or density estimation methods. This stepwise approach was given in Pereira *et al.*'s (2006) study to construct similar map for Portugal sheep breeds.

2.3.9.1 F_{ST} values

In the first step, pairwise F_{ST} values (Wright, 1965) for populations (breeds) were calculated since F_{ST} values are the measures of the proportion of total variance in allele (haplogroup) frequencies that occurs between populations (breeds). If breeds were differentiated (due to genetic drift) the proportion will be high, if gene flow between breeds were present, total variance in meta-population (all breeds) comes from populations and results a low value of proportion (F_{ST}) (Jobling *et al.*, 2004). As it is a measure of proportion F_{ST} values are in range of 0 and 1.

The proportions (F_{ST}) can be calculated by following formula:

$$F_{ST} = \frac{(H_T - H_S)}{H_T} \qquad (2.16)$$

where H_T is expected heterozygosity under Hardy-Weinberg theorem (Wright, 1951) and H_S is the mean expected heterozygosity across populations.

In the present study pairwise F_{ST} distances for breeds were calculated from haplogroup frequencies by Arlequin 3.11 (Excoffier, 2005).

2.3.9.2 Multidimensional scaling

In the second step, pairwise differences were summarized into one-dimension by multidimensional scaling (MDS) analysis (Gower, 1966) in which m dimensional data (a matrix) can be summarized into n dimensions where $n<m$. Unlike Principle component analysis (PCA), MDS summarizes all variance in the matrix as possible as could be without partitioning the variance. Since, like PCA, MDS gives the eigen values, the most variance assumed to be present in first dimension.

In the present study, metric MDS analysis performed by cmdscale command in R 2.14 (http://www.r-project.org/).

2.3.9.3 Kernel density estimation (KDE)

In the third step, Kernel density estimation (KDE) method was used to construct synthetic maps from values of first dimension of MDS analysis performed from pairwise F_{ST} values.

KDE is used to explore spatial point patterns (analysis of a set of point locations). It calculates the density of a circle (location s) in a given bandwidth (τ) (radius of an event s) by using a quadratic kernel function, k (). The representations of each variable were shown on a study region (R) in Figure 2.3.

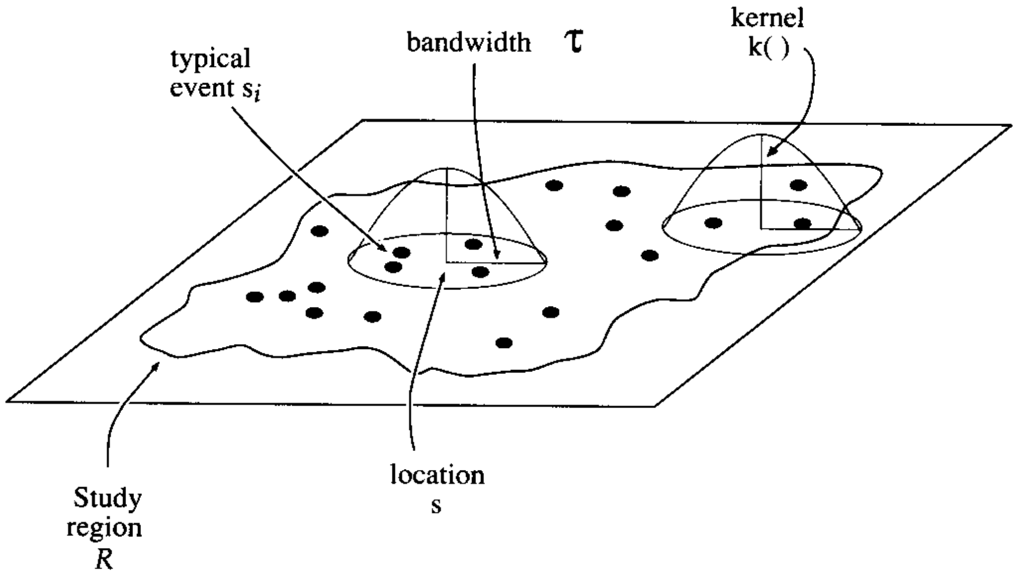


Figure 2.3 Kernel estimation of a point pattern. The picture taken from Gatrell *et al.*'s (1996) study.

The center of location s is the event s_i, the most contributed point in a location and also the highest value in kernel function.

Hence, the density for location s, $\hat{\lambda}_\tau(s)$ is estimated by the formula (Gatrell *et al.*, 1996, equation 3) given in equation 2.17.

$$\hat{\lambda}_\tau(s) = \sum_{i=1}^{n} \frac{1}{\tau^2} k\left(\frac{(s-s_i)}{\tau}\right) \tag{2.17}$$

where k, s, τ, and *n* are the kernel function, defined location, bandwidth, and sample size, respectively.

The volume under the kernel surface for each point (event s_i) is the value of that point and defines the number of times to count the point in the estimation process.

In present study, KDE was performed with centroids of collection sites of breeds (as points) and values in the first dimention of MDS representing each breed in a genetic level. For construction, Kernel Density method implemented in the Spatial Analyst Tool of ArcMap within ArcGIS Desktop 10 (www.esri.com) was used. To visualize major trend in Turkey, bandwidth was set as 3. In a map of KDE, the same color represents the kernel surfaces exhibiting the same density. So, genetically similar breeds are shown with the same color appearing as patches.

CHAPTER 3

RESULTS

3.1 DNA extraction and amplification

DNA extraction from blood samples was achieved by standard phenol:chloroform DNA extraction protocol (Sambrook *et al.*, 1989). The quality and quantity of DNAs were checked by micro-volume spectrophotometer (Nanodrop).

The amplification of DNA fragments was completed by Polymerase Chain Reaction (PCR). Amplified ND2, cytochrome *B* (*cytB*) and control region (CR) DNA fragments were run on 2%, 1.5% and 1.5% agarose gel containing EtBr, respectively, and visualized by Vilber Lourmat CN-3000WL displaying device. Negative controls were used to observe if there was a DNA contamination in the master mix causing false positives. DNA ladders were used to understand the size of amplified DNA fragments and to check the size of possible non-specific DNA fragments. Figure 3.1, Figure 3.2 and Figure 3.3 show the gel photographs of ND2, CR and *cytB*, respectively.

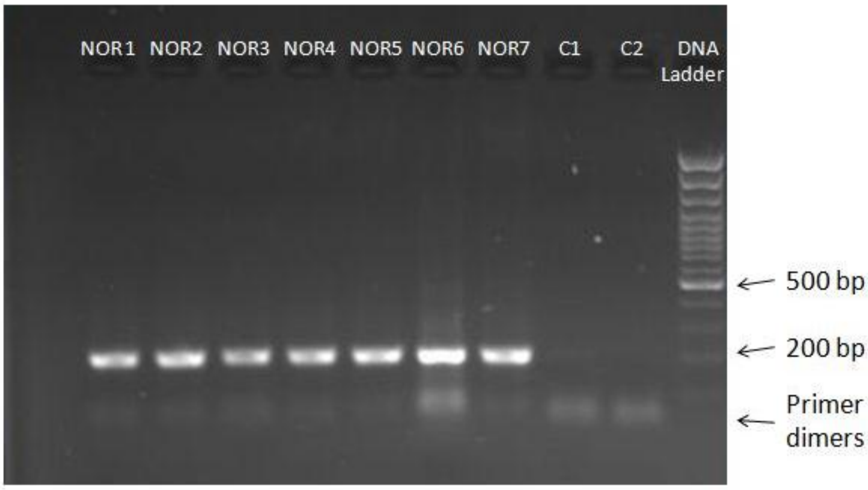


Figure 3.1 Visualization of a part of mtDNA ND2 gene DNA fragment. First seven wells contain the samples from breed Norduz (NOR), next two wells contain negative controls (C1 and C2), the last well contains DNA ladder (GeneRuler 100bp Plus DNA Ladder). 'bp' stands for base pair.

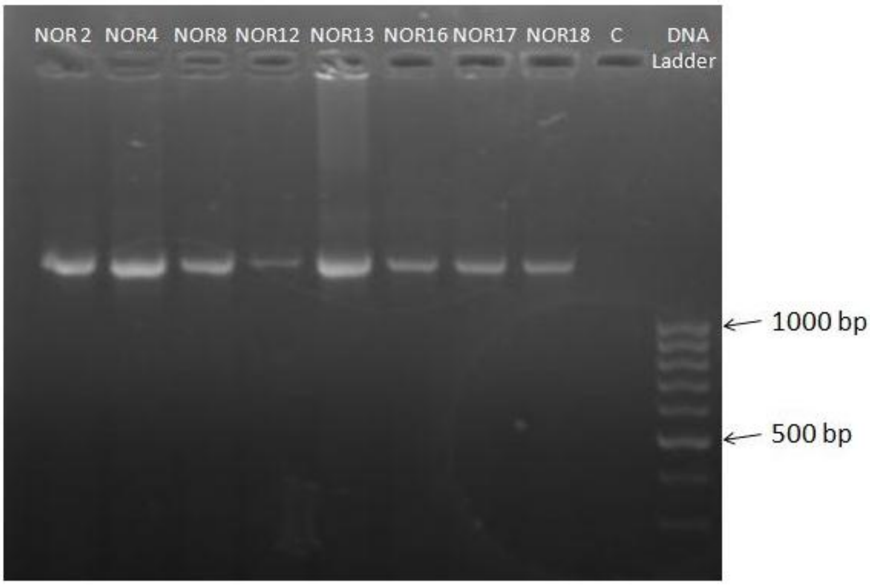


Figure 3.2 Visualization of the mtDNA CR region DNA fragment. First eight wells contain the samples from breed Norduz (NOR), the next well contains negative control (C), the last well contains DNA ladder (GeneRuler 100bp DNA Ladder).'bp' stands for base pair.

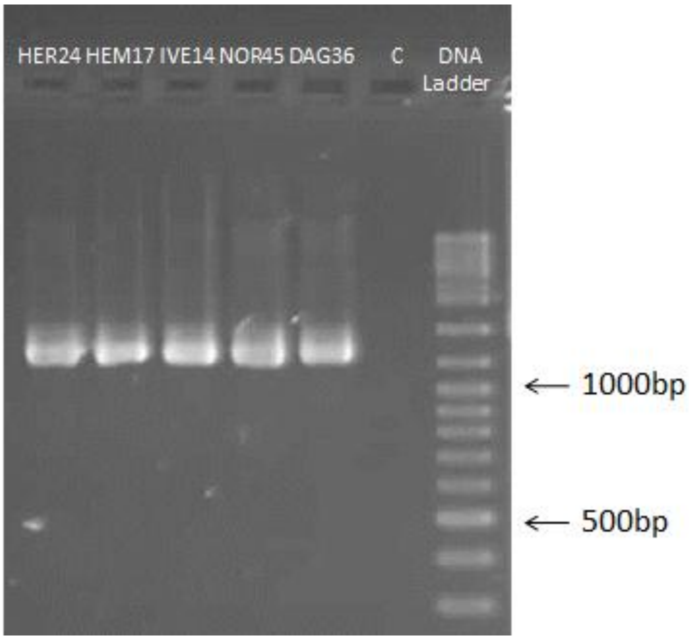


Figure 3.3 Visualization of the mtDNA *cytB* region DNA fragment. Wells contain the samples from breed Herik (HER), Hemşin (HEM), İvesi (IVE), Norduz (NOR), Dağlıç (DAG), negative control (C), and DNA ladder (GeneRuler 100bp plus DNA Ladder).‘bp’ stands for base pair.

3.2 Mitochondrial DNA (mtDNA) polymorphism in Turkish sheep breeds

Polymorphism in mtDNA was examined in different depths. In search for haplogroup (HPG) based polymorphism ND2 region SSCP and for within haplogroup and nucleotide polymorphisms CR sequencing were used.

3.2.1 mtDNA haplogroup (HPG) identification based on ND2 single strand conformational polymorphism (SSCP) analysis

Haplogroup identification of individuals was achieved by Single Strand Conformational Polymorphism (SSCP) analysis for the segment of mtDNA ND2 gene (Guo *et al.*, 2005). To identify banding patterns which are associated with the specific haplogroups, samples with known haplogroups (for HPG A, B, and C Yüncü’s (2009) and for HPG D and E Koban *et al*.’s (unpublished) samples) were employed.

The conformations of single-stranded DNA fragments which are specific for haplogroups were visualized by silver staining, after Polyacrylamide Gel Electrophoresis (PAGE). Photographs of gels were taken by Vilber Lourmat CN-3000.WL displaying device. Figure 3.4 shows the ND2 SSCP band patterns for each haplogroup.

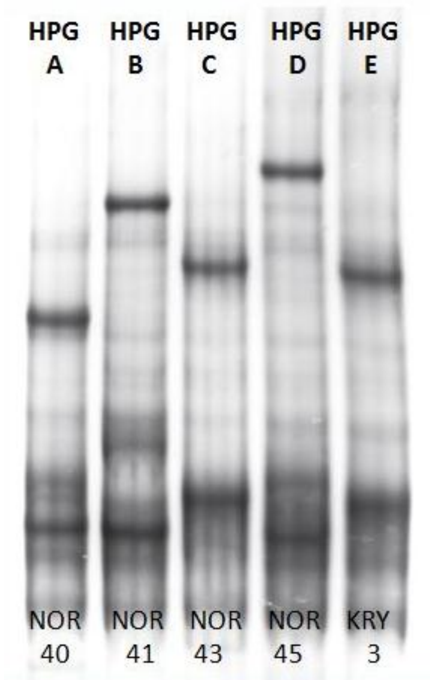


Figure 3.4 ND2 SSCP band patterns of each haplogroup.

As it can be seen from Figure 3.4 A, B and D haplogroups each had a distinct band patterns but C and E revealed identical hence a non-discriminating pattern for these two haplogroups. Therefore, the pattern is indicated by C/E afterwards.

ND2 SSCP analyses were carried out for 628 sheep in 13 Turkish sheep breeds. For each individual identified haplogroups were given in the Appendix A.

3.2.2 mtDNA CR sequencing

After identifying the haplogroups of each individual with respect to HPG A, B and C/E, individuals were selected for mtDNA CR sequencing. For this, 5 individuals per haplogroup were chosen from each population if possible. Since the observed individuals with HPG E are very rare (≤7) in the world and yet it is mainly observed in Turkey and Israel so far, to be able to increase the number and hence to facilitate the studies to be done on HPG E all of the C/E individuals were sequenced to obtain all those individuals which belong to HPG E.

A mtDNA fragment of 1501 bp long encompassing mtDNA CR (AF010406.1 positions 15271 to 156) were sequenced both in forward and reverse directions. Sequences were obtained from 240 domestic sheep from 13 Turkish sheep breeds. The chromatograms were checked and assembled by Chromas Pro version 1.5 http://www.technelysium.com.au/ChromasPro.html). Obtained consensus sequences from assembled primers were exported in fasta format. Figure 3.5 shows the part of assembled chromatograms of HC2 and CRF primer sequences.

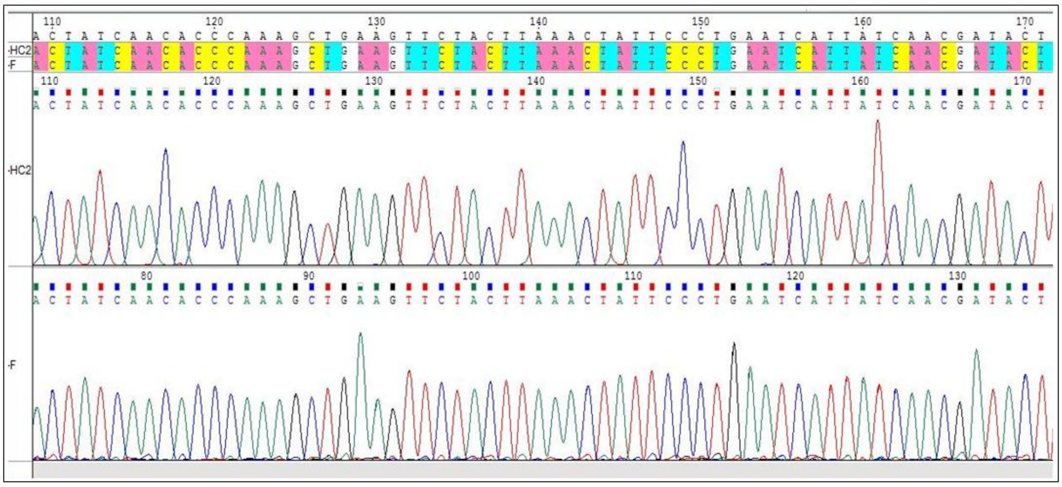


Figure 3.5 The part of assembled chromatograms of sequences obtained by HC2 and CRF primers from sample Morkaraman 17. At the top, the consensus sequence and under the consensus sequence, sequences obtained by HC2 and CRF primers and their corresponding chromatograms were represented, respectively.

Multiple sequence alignment of sequences was performed by ClustalW algorithm (Thompson *et al.*, 1994) implemented within Bioedit version 7.1.3 (Hall, 1997-2011). Alignment was checked by eye for the case of misplaced gaps. Also singletons are double checked from the chromatograms. Figure 3.6 shows the part of aligned sequences.

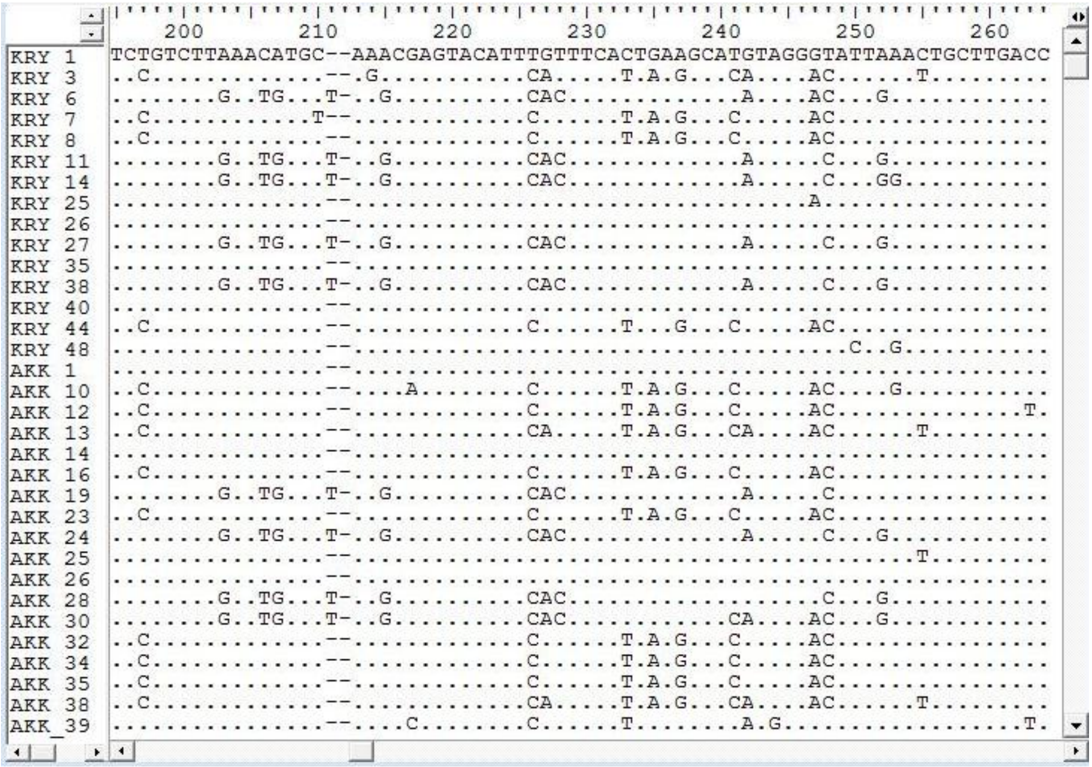


Figure 3.6 Snapshot was taken for a 70 base pair (bp) part of aligned sequences of Karayaka and Akkaraman breeds from Bioedit software.

Table 3.1 displayed the distribution of sequenced samples into breeds and haplogroups.

Table 3.1 Haplogroup distribution of sequenced samples.

	HPG A	HPG B	HPG C	HPG D	HPG E	Total
KARAYAKA	5	6	3	-	1	15
AKKARAMAN	6	4	7	1	3	21
GOKCEADA	5	6	-	-	-	11
DAGLIC	6	7	8	-	1	22
IVESI	8	7	11	-	-	26
HERIK	5	8	2	-	-	15
KARAGUL	7	8	3	-	2	20
CINECAPARI	4	7	5	-	-	16
SAKIZ	2	13	4	-	-	19
NORDUZ	5	5	11	1	-	22
HEMSIN	8	7	3	-	-	18
KIVIRCIK	-	7	2	-	-	9
MORKARAMAN	9	3	10	-	4	26
Total	70	88	69	2	11	240

After the sequences were aligned using Bioedit version 7.1.3 (Hall, 1997-2011), the location of 75 base pairs (bp) tandem repeats (3-5 repeat is present in each sample) was detected as was first reported by Hiendleder *et al.* (1998b). The distribution of the number of 75 bp tandem repeats within the breeds and haplogroups was given in Table 3.2.

Table 3.2 The distribution of the number of 75bp tandem repeats within breeds and haplogroups.

	4 repeat					3 repeat		5 repeat	
Breeds	HPG A	HPG B	HPG C	HPG D	HPG E	HPG A	HPG B	HPG C	n
Karayaka	5	5	3	-	1	-	1	-	15
Akkaraman	5	4	7	1	3	1	-	-	21
Morkaraman	8	3	10	-	4	1	-	-	26
Gökçeada	5	6	-	-	-	-	-	-	11
Dağlıç	6	7	7	-	1	-	-	1	22
Kıvırcık	-	7	2	-	-	-	-	-	9
İvesi	7	7	11	-	-	1	-	-	26
Herik	5	8	2	-	-	-	-	-	15
Karagül	7	8	3	-	2	-	-	-	20
Hemsin	7	7	3	-	-	1	-	-	18
Çine Çaparı	4	7	5	-	-	-	-	-	16
Sakız	1	12	3	-	-	1	1	1	19
Norduz	5	5	10	1	-	-	-	1	22
Total	65	86	66	2	11	5	2	3	240

Table 3.2 shows that mtDNA CR sequences with 4 tandem repeats were the most frequent among breeds. Sequences with 5 tandem repeats were observed only within HPG C, and there was no sequence with 3 tandem repeats within HPG C. All sequences belong to rare haplogroups (HPG D and E) had 4 tandem repeats. Since the tandem repeat number can change due to heteroplasmy (Hiendleder *et al.*, 1998b) it was removed from the sequences to prevent miscalculations/ overestimations. The remaining 879-889 bp (with gaps) segment of sequences was used for further analyses.

3.2.2.1 mtDNA haplogroups based on CR sequences by using neighbor joining (NJ) tree

The neighbor joining (NJ) tree based on 240 mtDNA CR sequences was reconstructed using MEGA 5 (Tamura *et al.*, 2011) software to reveal mtDNA haplogroup relations of samples. The Tamura-Nei (Tamura and Nei, 1993) model, was used instead of HKY model since HKY model was not implemented in MEGA 5 and Tamura-Nei model was the nearest comprehensive model of HKY model. Also, the Tamura-Nei model, which has 5 free parameters, covers the TPM1uf+I+G model (which was the results of AIC), in terms of the number of free parameters. In this way, the differential effects of selecting either AIC or BIC result was minimized.

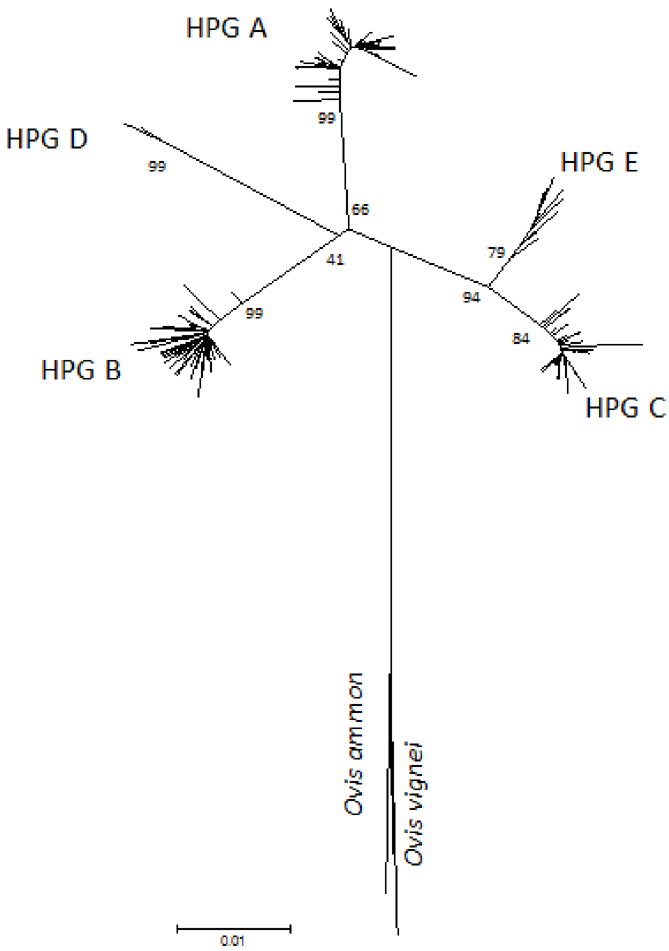


Figure 3.7 Neighbor joining (NJ) tree of mtDNA CR haplotypes. Bootstrap values for main divisions were shown on the branches of the tree. *Ovis vignei* (AY091490, AY091491) and *Ovis ammon* (AF242347, AF242348) sequences were from Hiendleder *et al.*'s (2002) study. Reference sequences (HM236174-83) for each haplogroup were from Meadows *et al.*'s (2011) study.

In Figure 3.7, six distinct clusters were observed: one of them represents wild sheep (*Ovis ammon* and *Ovis vignei*), five of them represent domestic sheep haplogroups

(HPG A-E). Haplogroup identifications of samples were made in accordance with the reference sequences retrieved from Meadows *et al.*'s (2011) study.

NJ tree showed that mtDNA CR sequences of the present study (n=240) were composed of 70 HPG A; 88 HPG B; 69 HPG C; 2 HPG D and 11 HPG E sequences. Sequences classified with respect to haplogroups were presented in Appendix A.

NJ tree indicated that HPG A, HPG D and HPG B are on one side of the node where wild sheep merged with the domestic sheep. HPG E and HPG C seemed to be relatively close to each other and two together were different than the other tree (HPG A, HPG B and HPG D).

3.2.2.2 Incompatibility in the identified haplogroups

NJ tree of mtDNA CR sequences revealed that nine of the samples belong to different haplogroup than that of identified by SSCP-ND2 analyses. These were given in the Table 3.3. Over all 3.75 % (9/240) difference was detected between two different haplogroup assignment methods. Highest discordance 5/83 was observed for C/E pattern of ND2 region.

Table 3.3 Samples that have different haplogroup relations according to SSCP-ND2 analysis and NJ tree of mtDNA CR sequences.

Samples	Haplogroup relation acc. to SSCP-ND2 analysis	Haplogroup relation acc. to NJ tree of CR sequences
DAG12	A	C
DAG32	C/E	B
MRK44	B	A
HER31	C/E	A
HER39	C/E	B
HEM7	C/E	B
HEM21	B	C
HEM43	C/E	B
CIC27	A	B

3.2.3 mtDNA *cytochrome B* (*cytB*) sequencing

Samples to be sequenced for *cytB* region were selected from each haplogroup based on mtDNA CR NJ tree. Two sequences from each of HPG A, D and E; 3 sequences from HPG C, and 1 sequence from HPG B were obtained. Sample names of sequences were as follows: HER24 (A) from breed Herik, DAG21 (A) and DAG36 (E) from breed Dağlıç, KRG17 (E) from breed Karagül, CIC23 (C) from breed Çine çaparı, IVE14 (C) from breed İvesi, AKK39 (D) from breed Akkaraman, NOR18 (C) and NOR45 (D) from breed Norduz, HEM17 (B) from breed Hemşin.

CytB sequences were used to generate a NJ tree together with the wild sheeps of Reazei *et al*.'s (2010) study and a MJ network together with the domestic sheep of examined by several previous studies (Meadows *et al*., 2005; Meadows *et al*., 2007;Pardeshi *et al*., 2007; Meadows *et al*., 2011) to construct similar network as it was in Olivieri et al.'s (2012) study. NJ tree and MJ network results were given in section 3.4 of the present chapter (Chapter 3).

3.2.4 Sequence based haplogroup diversities

The number of polymorphic sites, nucleotide and haplotype diversities of haplogroups were calculated with Polymorphism tool within DnaSP v.5 software (Librado and Rozas, 2009). These statistics were given in Table 3.4. For the rare haplogroups (HPG D and E), available sequences from the literature which were compatible with the sequenced region were used in order to facilitate the comparisons between the haplogroups or to detect the properties (such as for observing the signature of expansion) of the haplogroup.

HPG A was composed of 70 sequences; HPG B of 88; HPG C of 69; HPG D of 4 (2 of them from Meadows *etal*.'s (2011) study (HM236180, HM236181)); HPG E of 15 (2 of them from Meadows *etal*.'s (2011) study (HM236182, HM236183) and 2 of them from Guo *etal*.'s (2005) study (AY829385, AY829404) which were previously defined as HPGC) sequences.

Table 3.4 Number of sequences per haplogroup, variable sites (S), number of haplotypes (h), Haplotype diversity (Hd) and Nucleotide diversity (π).

	HPG A	HPG B	HPG C	HPG D	HPG E
Number of sequences	70	88	69	4	15
Variable sites (S)	61	82	47	6	27
Number of haplotypes (h)	48	66	43	3	13
Haplotype diversity (Hd)	0.975 ±0.011	0.987 ±0.006	0.947 ±0.020	0.833 ±0.222	0.981 ±0.031
Nucleotide diversity (π) (‰)	3.89 ±0.29	5.50 ±0.32	2.89 ±0.26	3.60 ±1.57	6.62 ±0.83

HPG B and E have the highest haplotype diversities among the haplogroups. Considering that haplotype diversity is a sample size dependent diversity measure (Jobling *et al.*, 2004), HPG E exhibiting the second highest haplotype diversity after HPG B, and with a small sample size relative to common haplogroups, indicated that HPG E was at least as divergent as HPG B. On the other hand, the least haplotype diversity was observed in HPG D as expected in relation to its minimum sample size.

Nucleotide diversity which is not a sample size dependent measure (unlike haplotype diversity) indicated that HPG E and HPG B had the highest values (also by considering the standard deviation (std) ranges) among the other haplogroups, Whereas HPG C exhibited the least nucleotide diversity among others. It was not reliable to compare HPG D with HPG C and A as its std range was very high. But still HPG D nucleotide diversity amount was not high as HPG E and B.

Number of sequences per haplogroup, variable sites and number of haplotypes observed were given to give a general idea about haplogroups and as supportive information for further analyses (i.e MJ network).

To analyze the relative genetic relatedness of the haplogroups Average number of nucleotide substitutions per site between haplogroups (Dxy) was calculated by using DnaSP v.5 software (Librado and Rozas, 2009).

Table 3.5 Dxy between haplogroups. The same individuals of Table 3.4 were used to represent the haplogroups. Dxy was explained in the text.

	HPG A	HPG B	HPG C	HPG D
HPG A				
HPG B	0.03275			
HPG C	0.03681	0.03586		
HPG D	0.03331	0.03392	0.03319	
HPG E	0.03298	0.03727	0.01030	0.03323

Average number of nucleotide substitutions per site between haplogroups (Dxy) was highest between HPG B- HPG-E and HPG A- HPG C but the minimum between HPG C –HPG E.

3.2.4.1 Median joining (MJ) network analysis

Median joining (MJ) network (Bandelt *et al.*, 1999) for 240 mtDNA CR sequences was drawn with NETWORK 4.600 (www.fluxus-engineering.com) to illustrate haplotype variation and to illustrate the relationships between haplotypes of the haplogroups.

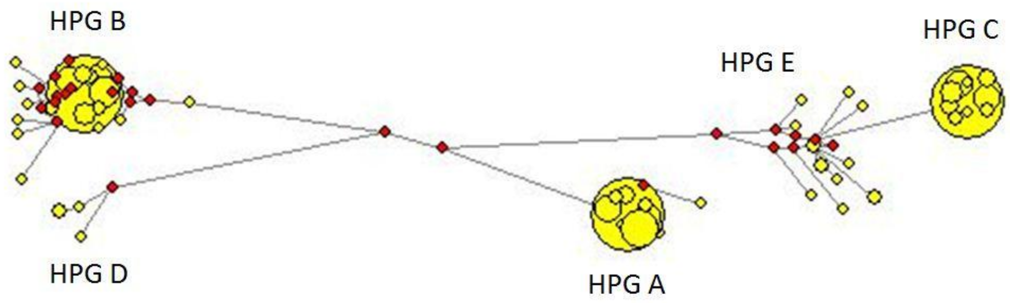


Figure 3.8 Median joining (MJ) network analyses of 886 bp mtDNA CR sequences. Yellow nodes indicate samples, red nodes indicate medians, and node sizes are proportional to the number of samples. Network was drawn after star contraction (Forster *et al.*, 2001) with maximum star radius of 10 (mutations). Branch lengths are proportional to number of mutations. Haplogroup compositions were the same as in Table 3.4.

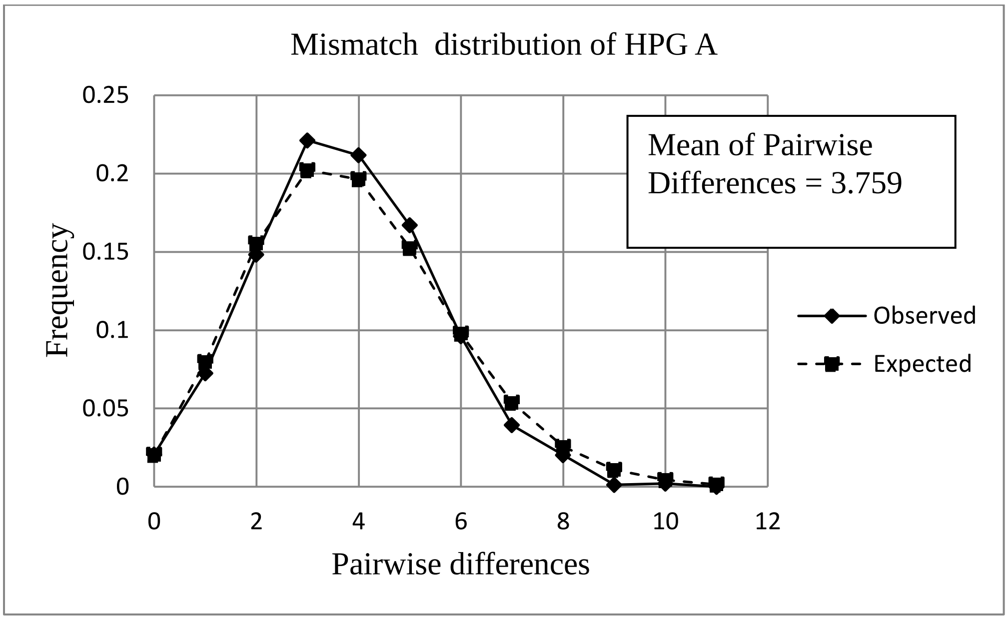
In the Figure 3.8, HPG B and E were observed as the clusters harboring the highest haplotype diversities in parallel to the observations made by the statistics given in Table 3.4 and Table 3.5. Conclusions on HPG D can not be made because of its small sample size. Regarding branch lengths, least distance was observed between HPG C and E.

3.2.5 Demographic history of haplogroups

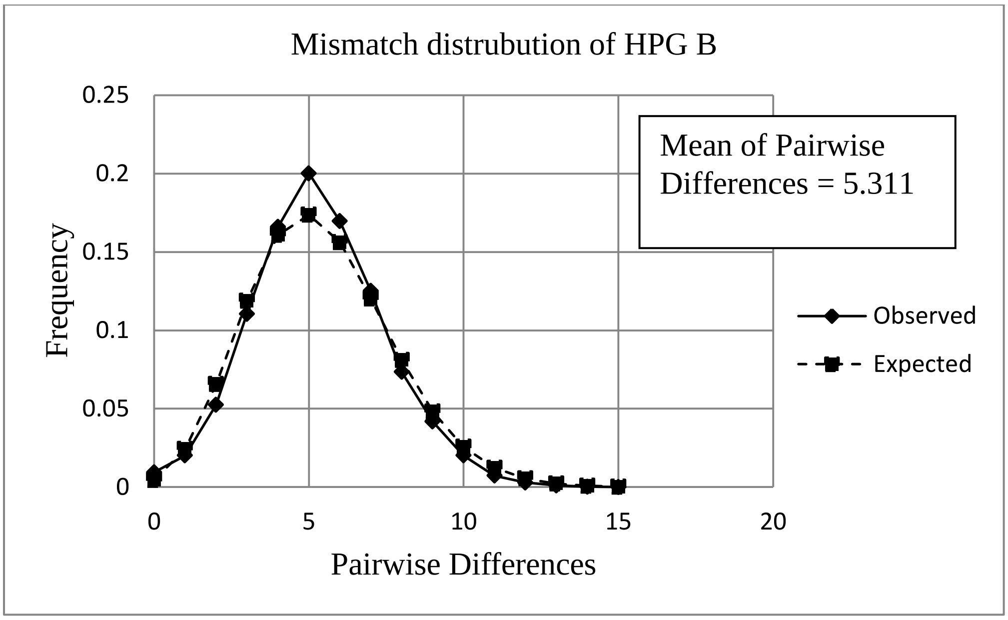
3.2.5.1 Mismatch distributions

To analyze whether or not haplogroups went through a population expansion in the past, the mismatch distributions (also called the distribution of the observed pairwise nucleotide site differences) for each haplogroup were examined. The observed and expected pairwise differences under sudden expansion model were calculated by Arlequin 3.11 (Excoffier, 2005). The graphics were visualized by using MS Office Excel. Figure 3.9 shows the graph of frequency versus pairwise differences for each haplogroup. Mean values of pairwise differences were also given on the graph.

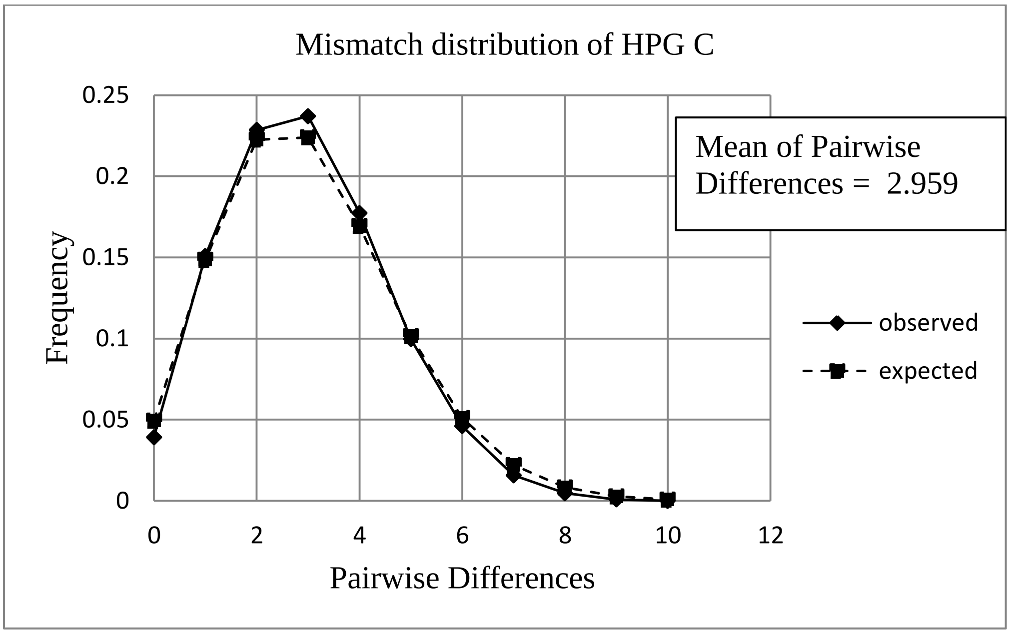
a)


Mismatch distribution of HPG A
Mean of Pairwise Differences = 3.759
Frequency
0.25
0.2
0.15
0.1
0.05
0
0
2
4
6
8
10
12
Pairwise differences
Observed
Expected

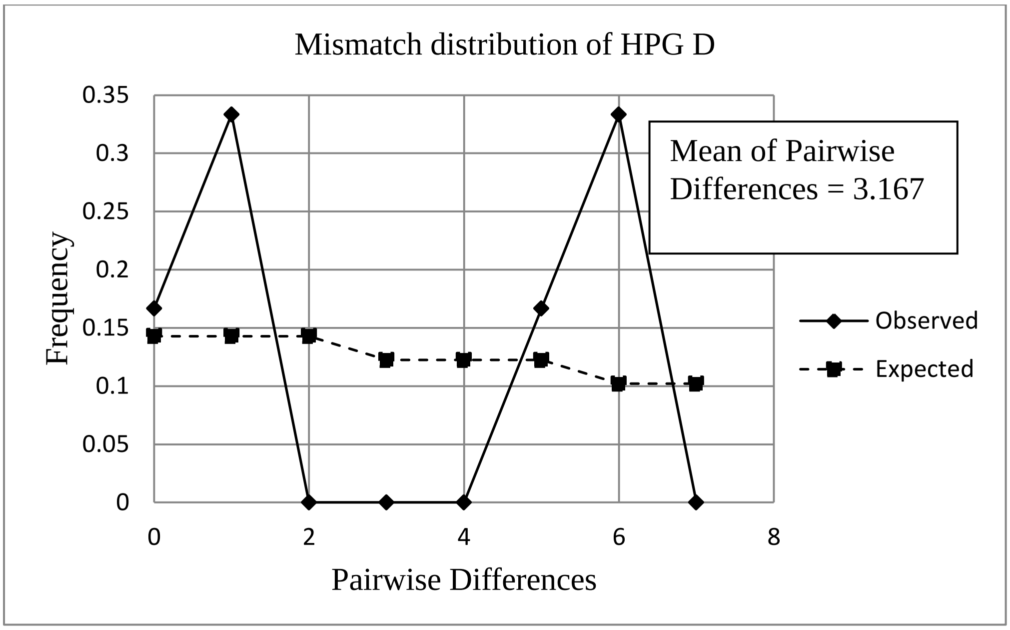
b)


Mismatch distrubution of HPG B
Mean of Pairwise Differences = 5.311
Frequency
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
Pairwise Differences
Observed
Expected

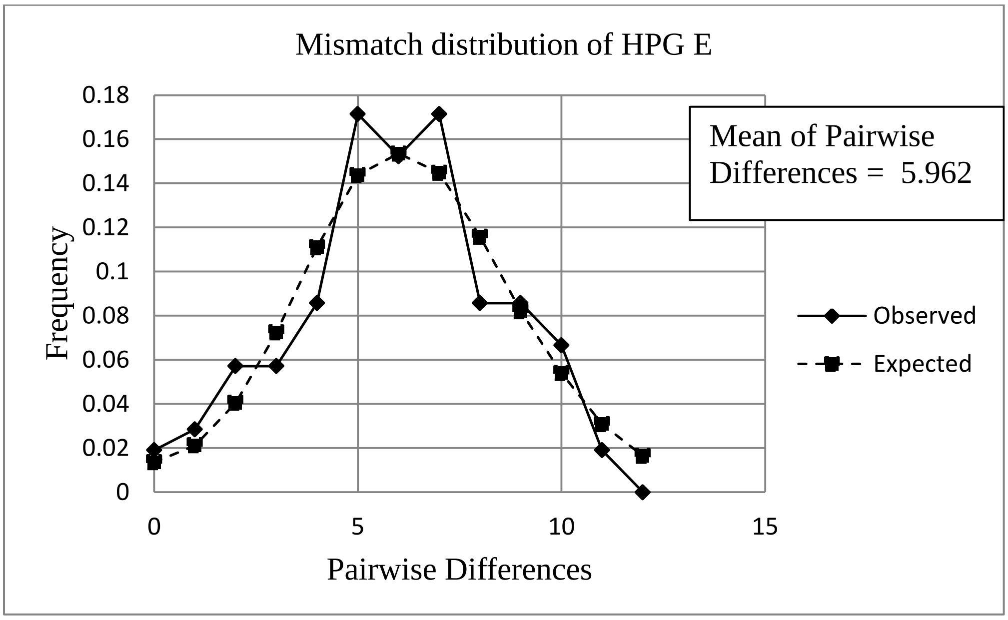
c)


Mismatch distribution of HPG C
Mean of Pairwise Differences = 2.959
Frequency
0.25
0.2
0.15
0.1
0.05
0
0
2
4
6
8
10
12
Pairwise Differences
observed
expected

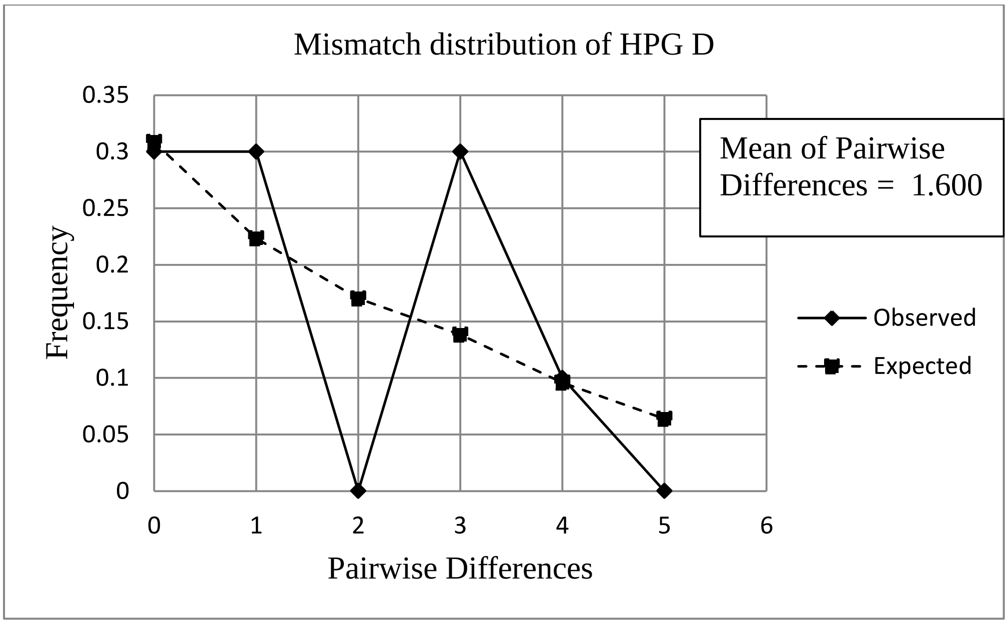
d)


Mismatch distribution of HPG D
Mean of Pairwise Differences = 3.167
Frequency
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
2
4
6
8
Pairwise Differences
Observed
Expected

e)


Mismatch distribution of HPG E
Mean of Pairwise Differences = 5.962
Frequency
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
Pairwise Differences
Observed
Expected

f)


Mismatch distribution of HPG D
Mean of Pairwise Differences = 1.600
Frequency
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
1
2
3
4
5
6
Pairwise Differences
Observed
Expected

g)

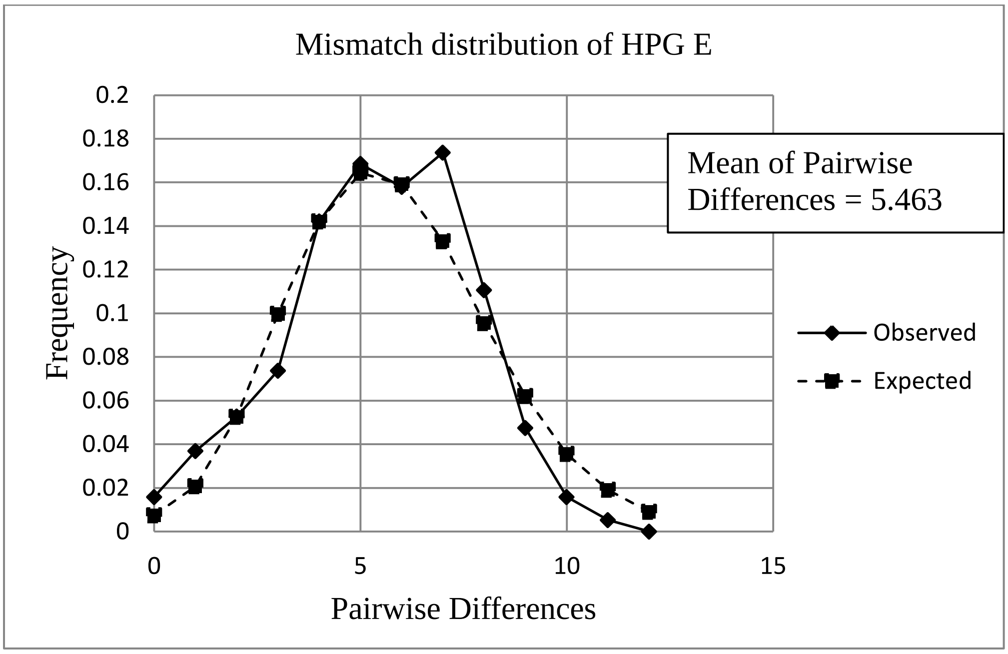


Figure 3.9 Mismatch distribution analyses of haplogroups over mtDNA CR (886 bp). a) HPG A (n=70); b) HPG B (n=88); c) HPG C (n=69); d) HPG D (n=4); e) HPG E (n=15); f) HPG D (n=5); g) HPG E (n=20). Haplogroup compositions for a-e were same as in Table 3.4. One sequence from Tapio et al.'s (2006) study (DQ242212) was retrieved and added to HPG D in f, and the sequence length was reduced to 420bp after including this sample. One sequence from Pedrosa *etal.*'s (2005) study (DQ097468) and four sequence from Koban *et al.*'s (unpublished) study (HM042760-61, HM042785, HM042838) were retrieved and added to HPG E in g, and the sequence length was reduced to 665bp when these samples were included.

In the Figure 3.9, the observed values of a, b, c, and g revealed bell-shaped mismatch distribution supporting population expansion model. For the domestic animals it is expected as the signature of expansion (Meadows *et al.*, 2007).

Mean number of pairwise differences of haplogroups, which fitted to sudden expansion model, were compared to determine the relationships of the time elapsed since population expansion of haplogroups. If the population sizes were the same for the haplogroups the highest value of the mean corresponds to the longest time elapsed since the population expansion. Under the assumptions of equal population sizes and equal rate of growth and no introgression from the wild HPG B (5.311) expanded earlier than HPG A (3.759) and HPG C (2.959). Among the tree HPG C seemed to went through the most recent expansion.

For the first time, in the presented study, with the additional samples from HPG E, the expansion of HPG E was confirmed (p = 0.77). In the light of this information, we can say that HPG E with the highest mean (5.962) represented the most ancient sheep population expansion.

3.2.5.2 Neutrality tests

Haplogroups were subjected to neutrality tests, to understand if the expansions of haplogroups were significant. Results of the tests were given in Table 3.6.

Table 3.6 Neutrality tests with regard to haplogroups. Haplogroups were represented by the same members as was used in Figure 3.9. Complete mtDNA CR sequences (879-889 bp) were used. Tajima's D (Tajima, 1989) and Fu's FS (Fu, 1997) statistics were calculated with Arlequin 3.11 (Excoffier *et al.*, 2005).

Neutrality Tests		
	Tajima's D	Fu's FS
HPG A (n=70)	-2.38970***	-26.04012***
HPG B (n=88)	-2.32174***	-25.40942***
HPG C (n-69)	-2.39903***	-26.47375***
HPG D (n=4)	-0.31446 NS	0.81143 NS
HPG E (n=15)	-1.17961***	-5.64201***
HPG E (n=20)	-1.37528***	-10.23346***

*** p < 0.001, NS Not Significant

Except Haplogroup D, all haplogroups had significant negative values for neutrality tests, supporting the sudden expansion model.

3.2.5.3 Bayesian skyline plots (BSP)

Bayesian skyline plots (BSP), that show population size changes as a function of time, in different combinations of haplogroups were summarized in Figure 3.10. The same 240 mtDNA CR sequences were used in BSP analysis. Plots were generated by BEAST v1.6 (Drummond and Rambaut, 2007). The analyses were done according to HKY nucleotide substitution model with invariable sites and gamma distribution as prior. Group number was left as default (10). MCMC iteration number was selected according to BEAST v1.4 manual (Drummond *et al.*, 2007) which states that iteration number can be increased until effective sample size (ESS) for each parameter exceed 200. MCMC iteration number was set as 200 million and samples were recorded in every 1000 iteration. BSPs were first drawn for each haplogroup to observe the existence of population expansion and to determine the beginning time of expansion. However, possibly due to low mutational differences within the samples of the haplogroups meaningful BSPs of individual haplogroups could not be obtained. Therefore, BSPs for only combined haplogroups and the whole data (HPG A, B, C, D and E) were shown on Figure 3.10.

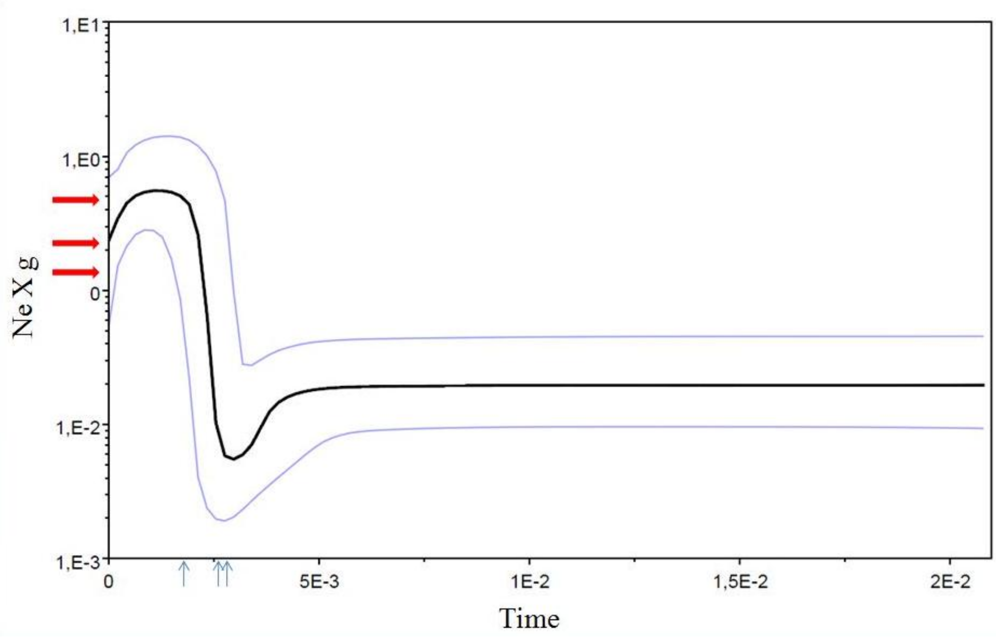


Figure 3.10 Bayesian skyline plot of 240 domestic sheep mtDNA CR sequences (HPG A-B-C-D-E). Time axis scale corresponds to number of mutations per nucleotide site. The arrows in Time-axis show the beginning of expansion of haplogroups in doublets which were grouped as A-C (n=139), B-C (n=157), A-B (n=158) left to right, respectively. The arrows in Ne X g (Effective population size X generation time) axis show the maximum values of doublets which were represented as A-B, B-C, A-C from top to down, respectively.

Haplogroup A-C doublet has least population size and has latest expansion time, whereas haplogroup A-B doublet has highest population size and earliest expansion time.

As a summary in relation to demographic histories of haplogroups mismatch distributions and neutrality tests confirms expansion of Haplogroup A, B, C and E. Bayesian skyline plots and the means of mismatch distributions suggested that expansion times can be in the order of B>A>C. However, the rare haplogroup E which is closest to C haplogroup might be the earliest expanded haplogroup among all of the tested.

3.2.6 Spatial distribution of haplogroups and sequences

3.2.6.1 Distribution of the haplogroups across the breeds

The mtDNA ND2 region based haplogroup frequencies of breeds calculated from 628 samples (Table 3.7) were summarized by pie charts and placed on the map of Turkey according to centroids of the collection sites of breeds (Table 3.8). Figure 3.11 shows the haplogroup compositions of breeds on the map of Turkey. Pie charts were generated by using MS Office Excel 2007 (Microsoft Cooperation).

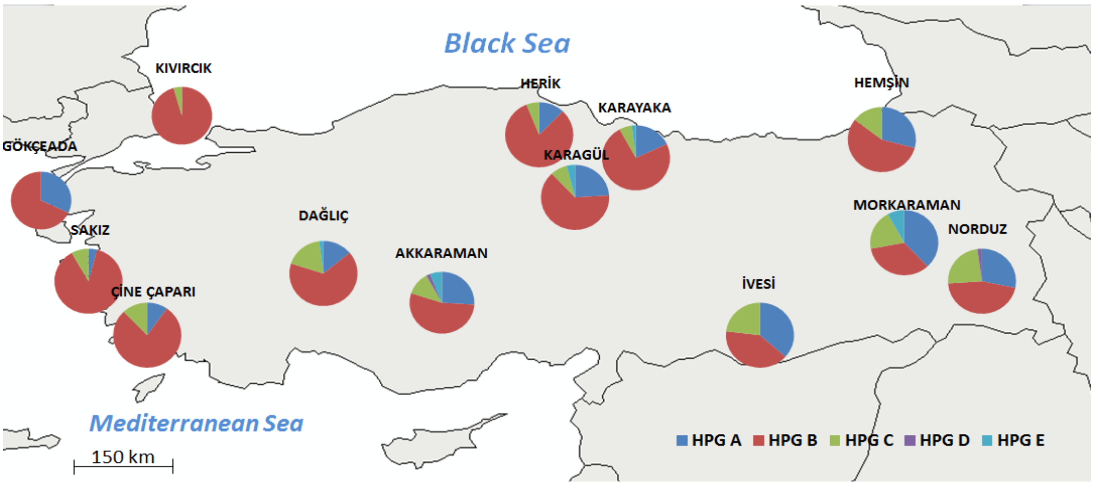


Figure 3.11 Demonstration of haplogroup compositions of breeds on the map of Turkey. Locations of the pie charts were on the centroids of the collection sites as given by Table 3.7.

Table 3.7 Haplogroup frequencies of breeds and sample sizes (n) of breeds.

	HPG A	HPG B	HPG C	HPG D	HPG E	n
KARAYAKA	0.1800	0.7400	0.0600	0.0000	0.0200	50
AKKARAMAN	0.2600	0.5200	0.1400	0.0200	0.0600	50
GOKCEADA	0.3200	0.6800	0.0000	0.0000	0.0000	50
DAGLIC	0.1400	0.6800	0.1600	0.0000	0.0200	50
MORKARAMAN	0.4000	0.3000	0.2200	0.0000	0.0800	50
KIVIRCIK	0.0000	0.9556	0.0444	0.0000	0.0000	45
IVESI	0.3725	0.4118	0.2157	0.0000	0.0000	51
HERIK	0.1225	0.8163	0.0612	0.0000	0.0000	49
KARAGUL	0.2600	0.6400	0.0600	0.0000	0.0400	50
HEMSIN	0.3333	0.5625	0.1042	0.0000	0.0000	48
CINECAPARI	0.1000	0.7750	0.1250	0.0000	0.0000	40
SAKIZ	0.0408	0.8776	0.0816	0.0000	0.0000	49
NORDUZ	0.2826	0.4565	0.2391	0.0218	0.0000	46
Total	0.2200	0.6430	0.1160	0.0030	0.0180	628

Table 3.8 Location of breeds given as centroids of the collection sites.

Breed Name	Lattitude	Longitude
AKKARAMAN	37.916625	32.86665
CINE CAPARI	37.62666667	27.83333333
DAGLIC	38.41444444	30.44777778
HEMSIN	41.19721364	42.03051364
HERIK	40.82033529	35.49661176
IVESI	37.07	39.02833333
KARAGUL	40.394	36.349125
GOKCEADA	40.20446	25.93656
KARAYAKA	40.46729286	36.63696429
KIVIRCIK	41.86985	27.50695
MORKARAMAN	39.841484	41.827072
NORDUZ	38.075	43.51333333
SAKIZ	38.2956625	26.3508125

In the Figure 3.11, a pattern of cline where there is a gradual reduction in HPG A frequency from east to west with an exception of Gökçeada breed was observed.For the HPG C frequencies the highest values at southeast Anatolia; intermediate values in the middle and southwest side and lowest values at northwest side of Turkey were observed.

3.2.6.2 Median joining analysis for two regions of Turkey

To analyze the difference between western and eastern regions of Turkey, five breeds were selected to represent Turkish native breeds, they were relatively high in population sizes hence were assumed not to be affected by the random genetic drift. As eastern breeds, Morkaraman and İvesi; as western breeds, Karayaka, Sakız and Kıvırcık were used. To increase the sample sizes, Morkaraman, İvesi, Karayaka and Sakız samples from Meadows *et al.*'s (2007) study were included to the analyses. Final sample sizes were 65 for both eastern and western regions of Turkey. However, length of mtDNA CR region was reduced to 526 bp upon the inclusion of samples from another study.

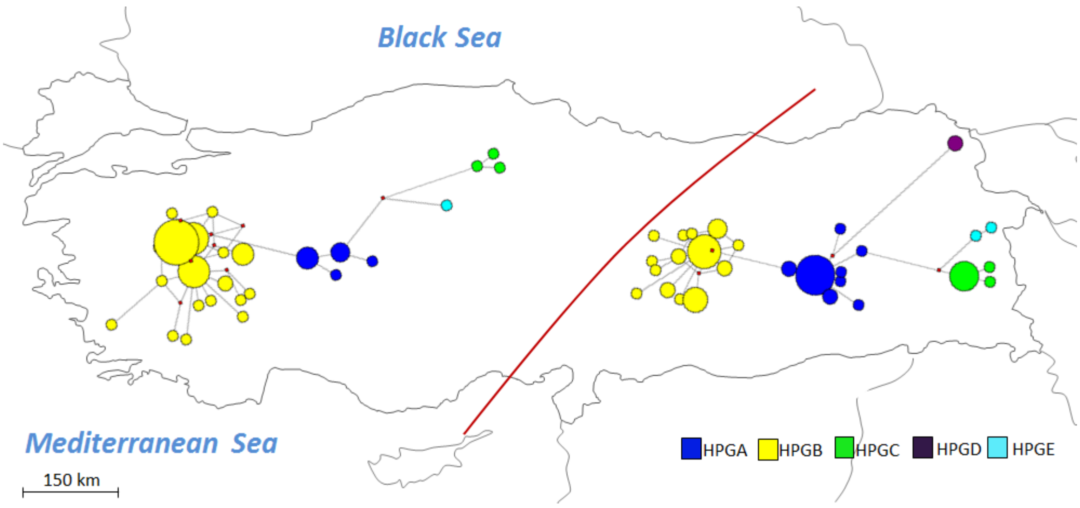


Figure 3.12 Median joining (MJ) networks of western and eastern regions of Turkey. Nodes are proportional to sample sizes and they represent haplotypes; red dots represent medians, branch lengths are proportional to mutational difference.

In the MJ network of the western part of Turkey, the samples were from the breeds Karayaka, Sakız and Kıvırcık, also the samples from the breeds Karayaka and Sakız (Meadows *et al.*, 2007) were retrieved from GenBank (DQ852107-12, DQ852190-215) and added to this data set. In the MJ network of the eastern part of Turkey, the samples were from the breeds İvesi and Morkaraman, also the samples from same breeds (Meadows *et al.*, 2007) were retrieved from GenBank (DQ852113-17, DQ852216-225, DQ852266-67, DQ852277, DQ852287-91, DQ852140-50, DQ852251, DQ852274-75) and added to this data set. Sequence length was 526 bp.

In Figure 3.12, eastern part of Turkey showed a higher diversity in terms of haplogroup composition, which was as follows: 32.3% HPG A, 47.7% HPG B, 13.8% HPG C, 3.1% HPG D, and 3.1% HPG E. For the western part of Turkey, haplogroup composition was as follows: 13% HPG A, 80% HPG B, 4.6% HPG C, and 1.5% HPG E.

3.2.6.3 Spatial autocorrelation analysis

Since, there was an obvious gradual decrease in frequencies of HPG C (southeast-to-northwest) and HPG A (east-to-northwest) as can be seen on the figures 3.11 and 3.12, Spatial autocorrelation analysis was carried out (Figure 3.13) to unravel the significance of the spatial pattern based on mtDNA haplogroups of Turkey.

The spatial correlogram was generated from mtDNA ND2 based haplogroups of each sample (n=628) and sample locations (Appendix A) by GenAlEx 6.41 (Peakall and Smouse, 2006). In this analysis, Spatial autocorrelation coefficients (r) were computed as in Smouse and Peakall's (1999) study and the distance classes were set as 150 km. The correlogram obtained from the Spatial autocorrelation analysis was given in Figure 3.13.

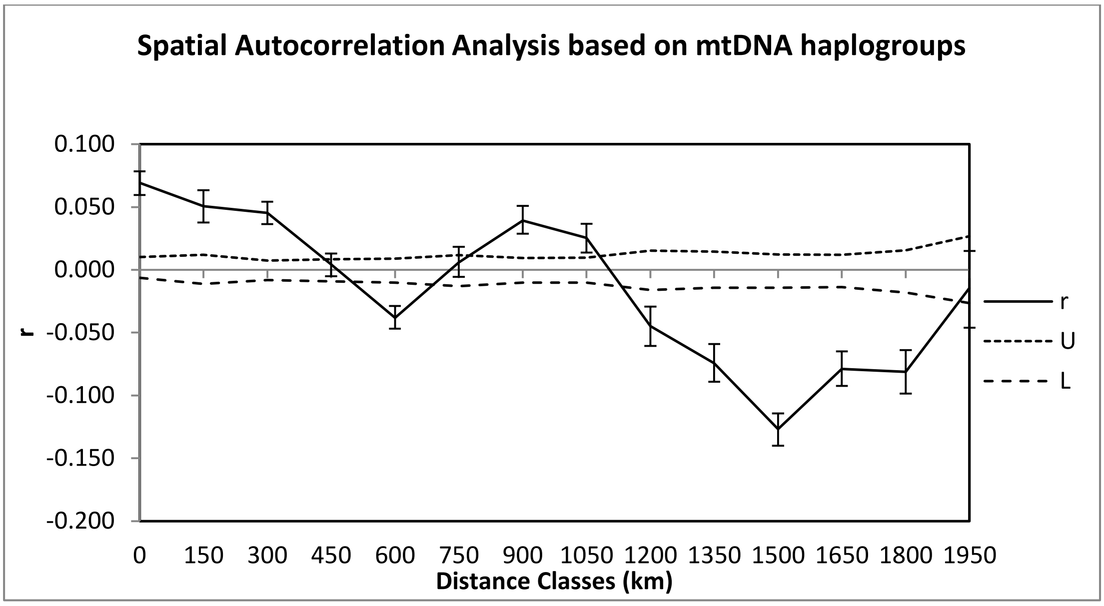


Figure 3.13 Spatial autocorrelation analysis based on mtDNA haplogroups. Abscissa represents distance class in km, ordinate represents spatial autocorrelation coefficients. The dashed lines represent the 95% confidence interval (CI) (U: Upper limit; L: Lower limit) for random distribution of genotypes in space from 1000 random permutations. Vertical bars indicate 95% CI for defined distance class from 1000 bootstrap trials. Any vertical bar for each defined distance class outside of 95% CI indicates significant ($p \leq 0.001$) deviation, from the expectation of random distribution.

For most of the r (spatial autocorrelation coefficient) highly significant ($p \leq 0.001$) values were observed. Correlogram indicated that, in general, up to 1500 km distance class there was a decrease in the similarity of CR haplogroup compositions suggesting the exhibition of long distance differentiation pattern (Barbujani *et al.*, 1994) with intrusion (Sokal, 1979) However, there may be two or more homogenous regions in Turkey. This proposed pattern would fit to the non-significant differences observed in the middle ranges of the correlogram.

3.2.6.4 Synthetic maps

After observing a significant correlogram in spatial autocorrelation analysis, which then called as intrusion model, to detect the regions which were affected by intrusion, the spatial trend wanted to be visualized on the map of Turkey by Kernel density estimation (KDE) method which is available in the Spatial Analyst tool of ArcGIS 10 software (esri.com). The centroids of collection sites for each breed in Table 3.8 were used as the location of breeds. The synthetic map obtained was shown on Figure 3.14.

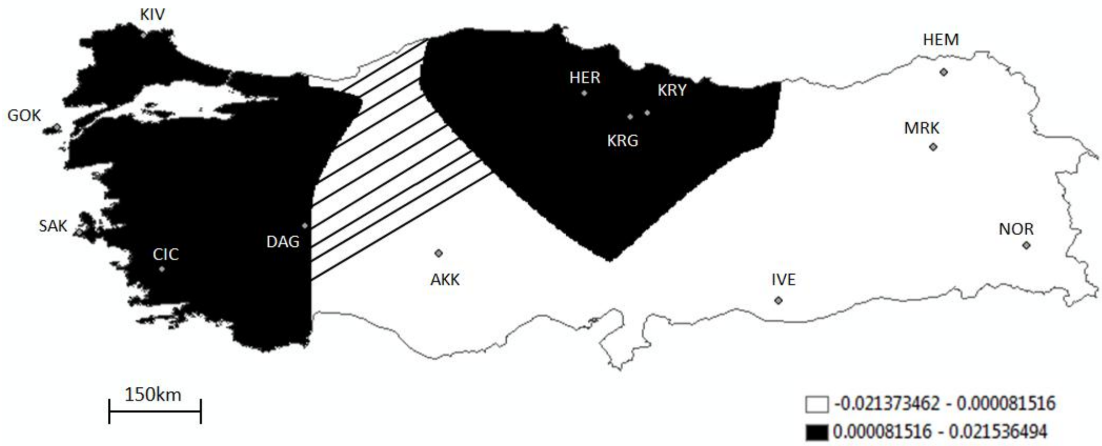


Figure 3.14 Synthetic map drawn by first dimension values of Multidimensional Scaling obtained from pairwise Fst distances between breeds based on mtDNA haplogroup frequencies. Kernel Density Estimation method build under the Spatial Analyst Tool in ArcGIS 10 (esri.com) were used to generate the synthetic map. Breeds were Kıvırcık (KIV), Gökçeada (GOK), Sakız (SAK), Çineçaparı (CIC), Dağlıç (DAG), Akkaraman (AKK), Herik (HER), Karagül (KRG), Karayaka (KRY), İvesi (IVE), Hemşin (HEM), Morkaraman (MRK), Norduz (NOR).

As seen on the synthetic map, Turkey was divided into three patches: two black regions and a white region. If two black regions considered as genetically similar and area between them (shaded area on Figure 3.14) was disregarded due to absence of data points, then presence of a barrier between eastern and central-western part of Turkey can be assumed supporting the difference between western and eastern part of Turkey as observed in MJ networks.

3.3 mtDNA analyses of *Ovis gmelinii anatolica*

3.3.1 mtDNA CR sequences and diversities.

In the present study, mtDNA CR sequences of 30 *Ovis gmelinii anatolica* (*Oga*) samples were also obtained. These sequences were used to construct NJ tree together with those of 240 domestic sheep samples of the present study and the tree was shown in Figure 3.15.

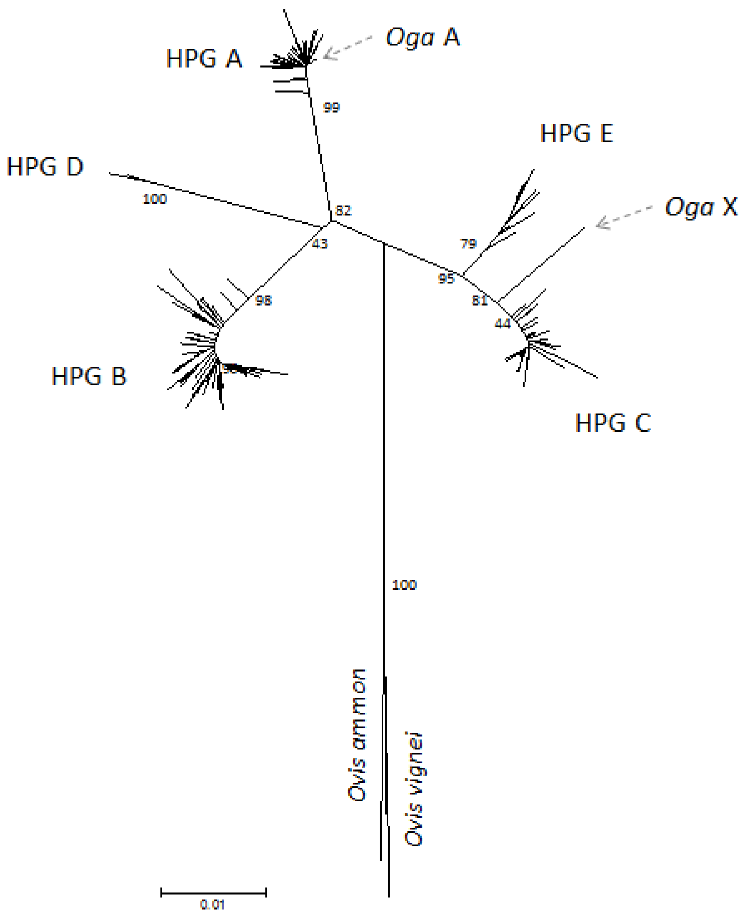


Figure 3.15 Neighbor joining (NJ) tree of mtDNA CR sequences with *Ovis gmelinii anatolica* samples. Bootstrap values for main divisions were shown on the branches of the tree. *Ovis vignei* (AY091490, AY091491) and *Ovis ammon* (AF242347, AF242348) sequences were from Hiendleder *et al.*'s (2002) study. Reference sequences (HM236174-83) for each haplogroup were from Meadows *et al.*'s (2011) study.

Ovis gmelinii anatolica samples were clustered with domestic sheep haplogroups instead of other wild sheep. They were separated into two parts revealing two different haplotypes, one of them consists of 8 samples located on the NJ tree in between HPG C and E (called as *Oga* X in the remaining text); the other haplotype consists of 22 samples located on the NJ tree in the middle of domestic HPG A.

In relation to the tandem repeats that *Oga* sequences had: One sample with 5 tandem repeats (75bp) within cluster X, 3 samples with 3 tandem repeats within HPG A were observed. Remaining 26 samples had 4 tandem repeats like most of the domestic sheep as seen in Table 3.2.

3.3.2 mtDNA *cytB* sequences

The 1041 bp long part of mtDNA fragment encompassing *cytB* region (NC_001941.1 positions 14180 to 15221) were obtained for two samples from both A and X group of *Ovis gmelinii anatolica* (*Oga*). The sample names were as follows: OGA14, OGA21 from Haplotype X and OGA9, OGA18 from Haplotype A. *CytB* sequences were used on the tree and the network defined in Section 3.2.3 of the present Chapter.

3.3.3 Comparative diversity estimates

Average number of nucleotide substitutions per site between populations (Dxy) was calculated between *Ovis gmelinii anatolica*, *Ovis musimon* (mouflon) from Germany and domestic haplogroups by using DnaSP v.5 (Librado and Rozas, 2009).

Table 3.9 Dxy between *Ovis gmelinii anatolica*, *Ovis orientalis musimon* (mouflon) from Germany and domestic haplogroups. *Ovis musimon* samples (n=5) consist of 2 samples from Hiendleder *et al*.'s (2002) study (AY091487-88), 1 sample from Hiendleder *et al*.'s (1998a) (AF039579) and 2 samples from Meadows *et al.*'s (2011) study (HM236184-85). For the haplogroups of domestic sheep, the ones given in Table 3.4 were used. *Ovis gmelinii anatolica* (*Oga*) samples (n=30) belong to the present study.

Sequences	HPG A	HPG B	HPG C	HPG D	HPG E
Oga X	0.03805	0.04134	0.01486	0.03932	0.02288
Oga A	0.00317				
O. orientalis musimon		0.00749			

Ovis gmelinii anatolica (Oga) X samples shared more nucleotides with haplogroup C and haplogroup E than others as was visually observed in NJ tree (Figure 3.15). Also, nucleotide substitutions between *Oga* A and HPG A (Dxy= 0.00317) were less than those nucleotide substitutions within HPG A (π= 0.00389) (Table 3.4 and Table 3.9). Moreover, *O. orientalis musimon* shared more nucleotides with HPG B (Dxy= 0.00749) than *Oga* X did with HPG C (Dxy= 0.01486). HPG C was the closest to *Oga* X among all of the haplogroups. These observations together with the observation that nucleotide substitutions between *Oga* X and HPG C (Dxy= 0.01486) is higher than nucleotide substitutions between HPG C and HPG E (Dxy= 0.01030) (Table 3.5) supports that *Oga* X cannot be called as HPG C.

3.4 Meta analyses

3.4.1 Median joining (MJ) network analysis of concatenated mtDNA CR and *cytB* (mtCR-*cytB*) sequences

To analyze the relationship based on mtCR-*cytB* sequences between *Ovis gmelinii anatolica* (*Oga*) and few (1-3) samples from the present study from each haplogroup and the haplotypes observed from all over the world, a MJ network similar to the one presented in Olivieri *et al.*'s (2012) study was constructed with NETWORK 4.6.1.0 (fluxus-engineering.com).

The concatenated sequences of *cytB* (NC_001941.1 positions 14214 to 15180) and CR (NC_001941.1 positions 16093 to 16615) (which were the common parts of the sequences of the present study and of the retrieved sequences), named as mtCR-*cytB*, were used to construct the MJ network. To construct MJ algorithm, the weights for variable sites (characters) were set as inverse proportional to the number of mutations observed for each site. The number of mutations observed for each site was calculated by statistics option given when MJ network was constructed from the same data by setting all weights equal (NETWORK 4.6 manual).

Figure 3.16 shows the MJ network of mtCR-*cytB* sequences, Figure 3.17 and Figure 3.18 shows the detailed version of Figure 3.16 in two parts; one containing HPG A, D and the outgroup and the other containing HPG B, C and E.

In Figure 3.16, MJ network was composed of 348 sequences, 334 of them retrieved from GenBank (Accesion numbers and corresponding haplotype labels in MJ network were given in Appendix B) including samples from Meadows *et al.*'s (2005), Pardeshi *et al*.'s (2007), Meadows *et al.*'s (2007), Meadows *et al*.'s (2011) and Olivieri *et al.*'s (2012) studies, 4 of them from *Ovis gmelinii anatolica* (sample names were given in section 3.3.2), 10 of them from domestic sheep breeds (sample names were given in section 3.2.3). The copper age sheep sequence was generated from NC_001941 according to nucleotide polymorphisms described in Table 1 of Olivieri *et al*.'s (2012) study.

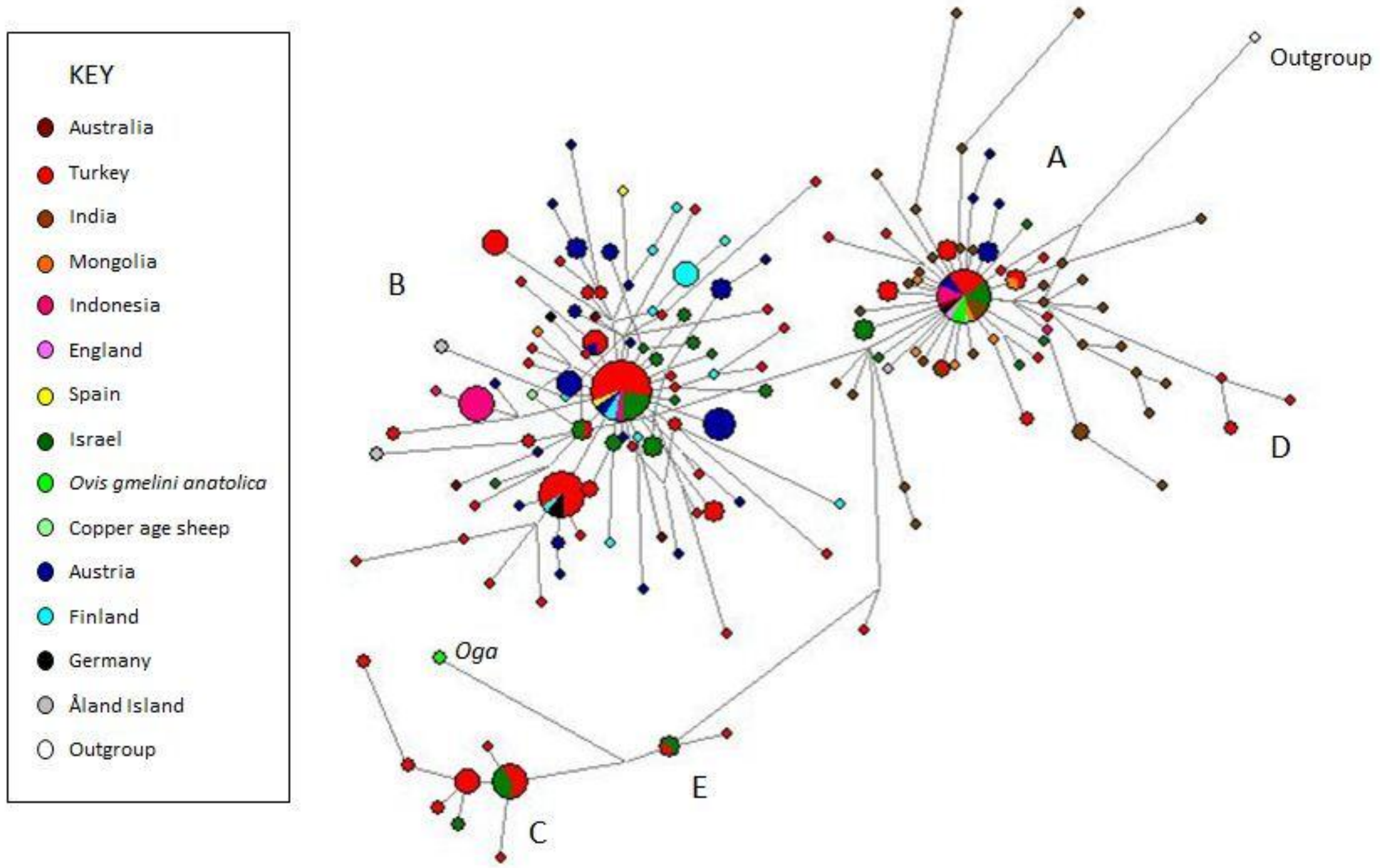


Figure 3.16 Median joining (MJ) network of mtCR-*cytB* sequences.Nodes representing the haplotypes are proportional to sample sizes. HPG names were given near the clusters.

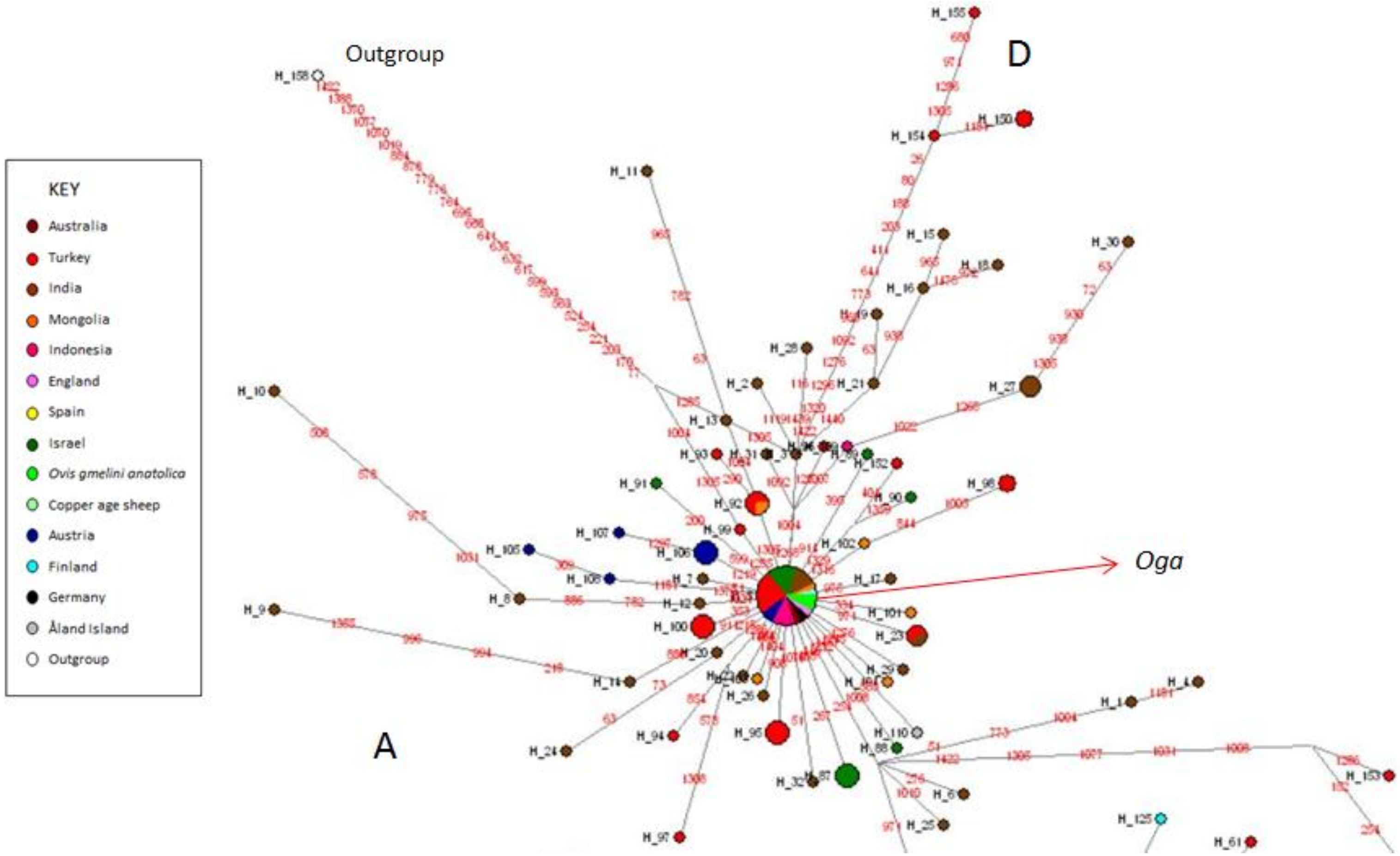


Figure 3.17 A part of MJ network showing HPG A, D and Outgroup. Nodes representing the haplotypes are proportional to sample sizes. The numbers on the links shows the variable sites on mtCR-*cytB* sequences (1-967 *cytB*, 968-1491 CR). Haplotype names were given near the nodes. The accession numbers of sequences used in MJ network and their respective haplotypes were given in Appendix B.

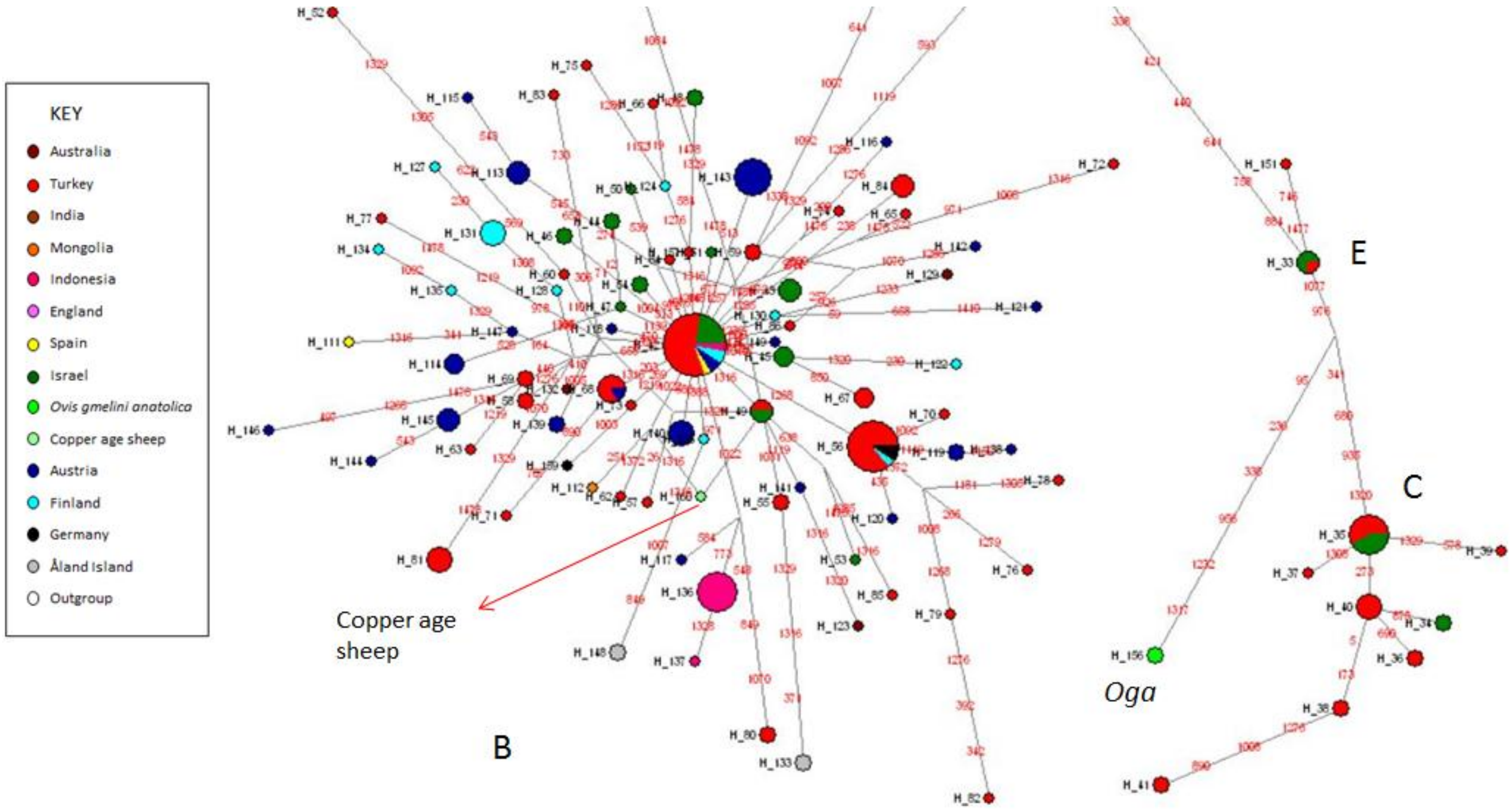


Figure 3.18 A part of MJ network showing HPG B, C and E. Nodes representing the haplotypes are proportional to sample sizes. The numbers on the links shows the variable site position on mtCR-*cytB* sequences (1-967 *cytB*, 968-1491 CR). Haplotype names were given near the nodes. The accession numbers of sequences used in MJ network and their respective haplotypes were given in Appendix B.

All the samples which were previously associated with the haplogroups based on the mtDNA CR sequences were clustered in the same haplogroups on the new reference frame defined by mtCR-*cytB*, exhibiting a consistent haplogroup definition for along mtDNA.

Ovis gmelinii anatolica (*Oga*) samples which clustered with HPG A in NJ tree also clustered with HPG A in MJ network and observed in central node of HPG A which was the most frequent haplotype in HPG A (25/99). Similarly, *Ovis musimon* samples (HM236184-85, Meadows *et al*., 2011) which were considered as feral domesticates (Hiendleder *et al.*, 2002) were observed in second central node (20/208) of HPG B. When the two central nodes of HPG B in which samples from Turkey were in high frequency (18/31 and 17/20) were analyzed in breed level, 6 samples from Karakas breed, 1 from Morkaraman, 2 from Karayaka, 2 from Çineçaparı, 3 from Norduz, and 4 from Karya breed were observed in the first central node (31/208). In the second node there were 15 samples from Sakız breed, 1 from Karayaka, and 1 from Karya breed.

Oga samples which were in between the HPG E and C were again linked with the central nodes of HPG C and E, differing in 6 mutations from median node (the node that intercepts the link between central nodes of HPG E and C to the link to *Oga*). In addition, the copper age sheep (5350-5100 years before present) were linked to central node of HPG B (31/208) with 2 mutations.

3.4.2 Median joining (MJ) network based on mtDNA *cytB*

To analyze the relationship based on mtDNA *cytB* sequences of *Ovis gmelinii anatolica* (*Oga*) with those of the other wild *Ovis gmelinii* populations, the data from a recent study was used (Rezaei *et al.*, 2010). In that study, wild sheep *Ovis gmelinii anatolica* from Turkey, *Ovis gmelinii gmelinii* from Armenia and Iran, *Ovis gmelinii isphahanica* from Central Iran, *Ovis gmelinii laristanica* from south of Iran, *Ovis vignei* and hybrids of *Ovis vignei* and *Ovis orientalis* from eastern and central eastern Iran were presented, respectively. By the way, in describing those sequences the new nomenclature (IUCN/SSC, 2000; Festa-Bianchet, 2000) was used (for instance instead of *Ovis orientalis gmelenii, Ovis gmelinii gmelinii* was used) Few (1-3) samples of the present study representing each of the haplogroups observed for the Turkish sheep, 4 *Oga* samples were employed to construct, a MJ network by NETWORK 4.6.1.0 (fluxus-engineering.com) where they were visualized together with the wild sheep samples used by Rezaei *et al.* (2010). The MJ network was given in Figure 3.20.

Comparisons of the wild sheep sequences and the haplogroups of the domestic sheep suggested that only one *O. gmelinii gmelinii* sample (1/37) which was observed in Iran near the border of Turkey share a haplotype with HPG B. Also, 7 (7/37) *O. gmelinii gmelinii* samples and 1 *Ovis gmelinii laristanica* sample (1/4) which shared a haplotype with HPG E were observed in the northwest region of Iran and southeast region of Iran, respectively. In the HPG A case, *Ovis gmelinii anatolica* samples from the present study, which clustered with HPG A based on mtDNA CR sequences, clustered in HPG A based on mtDNA *cytB* sequences, too. In that haplotype (H_7), 2 *O. gmelinii anatolica* (2/4) samples, 2 *O. orientalis isphahanica* (2/7) samples, and 4 *O. gmelinii gmelinii* (4/37) samples were present. In the HPG C case, only one *O. gmelinii* (1/37) sample shared a haplotype with 2 HPG C samples. Also, it is worth to say, the nearest wild sheep haplotype to HPG D was found to be *O. gmelinii isphahanica* which had 7 different nucleotides from haplotype of HPG D.

The X haplotype of *O. gmelinii anatolica* observed in NJ tree, was also present in *O. gmelinii anatolica* samples (H_12) of Rezaei *et al*.'s (2010) study which were sampled from Central Anatolia region of Turkey.

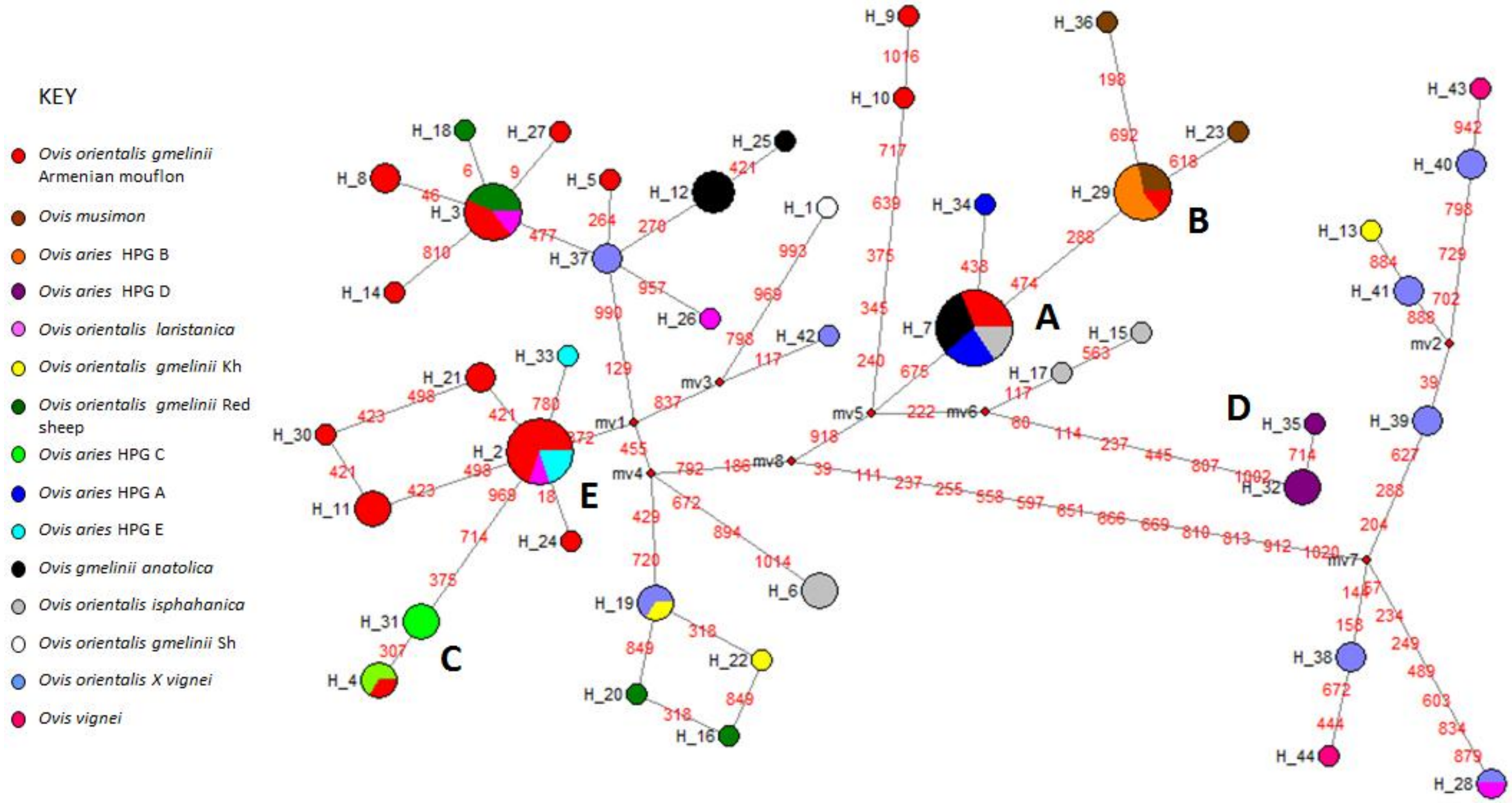


Figure 3.19 MJ network of mtDNA *cytB* region.Nodes representing the haplotypes are proportional to sample sizes. The numbers on the links shows the variable sites on mtDNA *cytB* sequences (1-1042 corresponds to NC_001941 positions 14180-15221). Haplotype names were given near the nodes. The accession numbers of sequences used in MJ network and their respective haplotypes were given in Appendix C.

CHAPTER 4

DISCUSSION

In the present study, to contribute to the understanding of sheep domestication and then to the domestication history of sheep; the distribution of genetic diversity among domestic sheep based on mtDNA were investigated. The inclusion of the samples *Ovis gmelinii anatolica* from Konya Bozdağ region and sequencing their respective mtDNA regions made the discussions more comprehensive. In this chapter, results were evaluated and discussed comparatively with the results reported in the literature.

4.1 Recent publications and changing paradigms in the evolution of domestic sheep

Since the present study was first conceived in 2005 in the context of a large scale national project with the acronym TURKHAYGEN-I, generally accepted models in relation to sheep domestication and evolutionary history of domestic sheep have changed.

In the first half of the last decade the working hypothesis was: "Sheep was domesticated in the Middle East than dispersed in different parts of the world. Anatolian mouflon (*Ovis gmelinii anatolica)* might be the ancestors of the HPG B sheep". Hence, the implicit understanding was that the global diversity pattern of the sheep in the old continents was set by the migrations of the domestic sheep perhaps with the confounding effects of local migrations. For this time interval, the most significant studies concerning the evolution of domestic sheep can be summarized as given below.

I) The observations contributing to the prevailing paradigm of the first half of the last decade in relation of sheep domestication were as follows:

1) Archaeological studies indicated that sheep was domesticated in Fertile Crescent around southwest Anatolia (Vigne *et al.*, 2003; Peter *et al*., 2007)

2) Sheep had many mtDNA haplogroups (Wood and Phua, 1996; Hiendleder *et al.*, 1998a; Pedrosa *et al*., 2005; Tapio *et al*., 2006; Meadows *et al*., 2007)

3) The wild sheep of Sardinia and Corsica (European Mouflon) was considered as the feral form of the domesticated sheep (because it was not different than mtDNA of domestic sheep and there were no wild sheep in Europe in Holocene)

4) Anatolian wild sheep being closest to the European sheep, which is predominantly exhibiting haplogroup B, was suggested as the candidate wild sheep population to be the ancestor of HPG B (Hiendleder *et al*., 2002).

5) As it was argued for goats, different mtDNA haplogroups refered to different domestication events (Luikart *et al*., 2001).

6) mtDNA haplogroups were observed as first two then as three (Tapio *et al.*, 2006). The third mtDNA haplogroup seemed to be present in Turkey (Pedrosa *et al.*, 2005; Meadows *et al*., 2007) and in Portugal (Pereira *et al*., 2006) the latter due to the migration through meditterenean perhaps during the Andulician’s.

7) The high number of haplogroups, and the higher diversity for the haplogroups in the Middle East (Bruford *et al*., 2003; Bruford and Townsend, 2006) was congruent with the argument that sheep domestication center was in the Middle East.

However, further studies in the last 5-6 years have either strengthened the early beliefs or have modified some of them. Yet, there were some new realizations also during this period.

II) Major studies and their contributions to the working hypothesis of sheep domestication in the last 5-6 years.

1) The belief that the Middle East was a genetic hot spot for sheep was reinforced as well as evidenced by the presence of high number of mtDNA haplogroups (Meadows *et al.*, 2007), also by the observation of high genetic diversity by microsattelites (Lawson-Handley *et al.,* 2007; Peter *et al*., 2007; Uzun *et al*., 2006).

2) As the domestication center of sheep, a more confined area within the Anatolia from Central Anatolia towards the northern Zagros Mountains was suggested (Zeder, 2008).

3) By reviewing archaeological studies, one of the interesting inference made was: At the down of domestication (~ 12.000 years before present) animals such as sheep and goat before exhibiting morphological signatures of domestication must be transported even to islands from the continent, presumably by hunter gatherers (Zeder, 2008).

4) There was a new realization and observations about the possible explanation of observing multiple mtDNA haplogroups in domestic sheep:
   i) In goat it was suggested that more than one haplogroup can be involved in one domestication event (Naderi *et al.*, 2008).

ii) Wild goat of northeastern Anatolia harbors almost all of the mtDNA haplogroups that exist in domestic goats (Naderi *et al.*, 2008). Hence it was suggested that with one domestication event more than one haplogroup from the domestication site (for instance from northeastern Turkey) can be captured in domesticated sheep (Naderi *et al.*, 2008)

5) MtDNA *cytB* sequences from wild sheep populations existing in Anatolia and Iran became available in the literature (Rezaei *et al.* 2010).

6) In an attempt to identify male mediated, Y chromosome dependent, sheep haplotypes (Meadows *et al.*, 2006; Meadows and Kijas, 2008), it was observed that H6 was widely distributed haplotype in rams.H4 with low frequency was observed among domestic sheep of Asia; similarly H5 was (low frequency) present in the sheep of Europe. H12 was private haplotype of Sakız rams (Meadows and Kijas, 2008).

7) In relation to domestic sheep evolution, with the help of nuclear markers "retrotypes" revolutionizing observations were made (Chessa *et al.*, 2009). In this study it was suggested that dispersion of the sheep to the world was not through a single large scale migration (after the domestication) but a second large scale migration took place, perhaps starting from Middle East and around 5^{th} Millennium before present and in majority of the modern day sheep genetic diversity was set by "the second expansion of sheep" (Chessa *et al.*, 2009).

8) A reference frame covering mtDNA CR and *cyt B* segments of sheep from many parts of the world was given on which sequences of domestic and wild sheep as well as an ancient DNA can be placed. With the provided network (Olivieri *et al.*, 2012) genetic relatedness of the newly acquired sequences to the already existing ones can be obtained or the distinctness of the old sheep DNA frozen about 5000 years before present can be used

as a measure to be employed as a range of expected difference between the DNA's of wild and domestic sheep.

4.2 Sampling and sheep mtDNA haplogroups in Turkey

The main genetic marker used in the presented study, mtDNA, is an informative, widely used marker to identify genetic diversity of domestic sheep (Bruford *et al.*, 2003; Toro *et al.*, 2009). The distribution of haplogroups defined by mtDNA may provide information about geographic patterns and migration ways in the maternal aspect (Groeneveld *et al.*, 2010) and as a result domestication centers of species may be assessed (Naderi *et al.*, 2007).

In spatial analyses, presentation of the breeds and studied area is very important and to obtain a reliable pattern sampling should be diverse enough. In the present study, sample collection was done by staff from Ministry of Food Agriculture and Livestock as the requirement of TURKHAYGEN-I project (www.turkhaygen.gov.tr). Sampling of 2-5 individuals from each flock (when possible) was carried out; hence samples used in this study were not close relatives. There were two previous studies, where data was collected from sheep of Turkey based on the mtDNA region. However, they were not aiming to reveal spatial patterns. They were representing the mtDNA diversity of sheep from Turkey as a whole. First, Pedrosa *et al.* (2005) studied on samples of Turkey with 79 individuals from 5 breeds. Then; Meadows *et al*. (2007) studied on samples of sheep from Turkey with 140 individuals from 8 breeds by taking samples from one flock per breed. The two major breeds: northwestern breed Kıvırcık and southeastern breed İvesi were not covered in Meadows *et al*.'s (2007) study. With 628 individuals from 13 breeds (also covering all of the major breeds and the breeds from the edges of the distribution), together with information of collection sites it is hoped that a good presentation of the mtDNA based spatial genetic diversity could be captured and presented in this study.

Sequencing all the 628 individuals of the present study would have been costly. Therefore, in the first stage of the study, a quick and cheap method to determine mtDNA haplogroups of these individuals was searched. It was assumed that, a large number of sheep (n=628) collected from many breeds (n=13) across the geographic region where sheep domestication and subsequent migrations were witnessed sets an appropriate environment to test the power of the HPG screening method. MtDNA based haplogroups were identified employing different regions of mtDNA as sugessted by some authors. For instance, Bruford and Townsend (2006) identified the haplogroups based on the control region (CR) by RFLP method and Guo *et al.* (2005) identified them by employing ND2 or ND4 regions by SSCP method. After, the power of ND2-SSCP in discriminating mtDNA based haplogroups was revealed by Guo *et al.* (2005), haplogroup relations of 628 samples were investigated by this method. In addition to Guo *et al.*’s (2005) study, the discriminating ability of ND2-SSCP compared to HPG D method was detected. As a result, four banding patterns for HPG A, B, C/E, and D was reported by this method in the present study. In total, the haplogrups of 240 individuals were identified by both methods, and the discordance was low (3.75 %), meaning results of the two approaches can be used interchangeably with a minor discrepancy.

The difference between haplogroups defined by different regions of mtDNA with comparing ND2-SSCP, ND4-SSCP and CR-RFLP methods was studied by Yüncü *et al.* (unpublished) (or TURKHAYGEN-I project Final report, turkhaygen.gov.tr). The study suggested that ND2-SSCP method should be used to haplogroup identification among three methods when the sequencing was not possible since ND2-SSCP method can discriminate HPG D in addition to HPG A and B and has high concordance with sequence results in comparison to others. The study also revealed high concordance between these three methods (91.72 %) indicating that results obtained from different haplogroup identification methods can be used together to understand the haplogroup distribution worldwide. With this information, haplogroup compositions of geographic regions taken from the previous studies (Guo *et al.*, 2005; Tapio *et al.*, 2006; Meadows *et al.*, 2007; and Econogene project) and from the presented study were shown on Eurasia map in Figure 4.1.

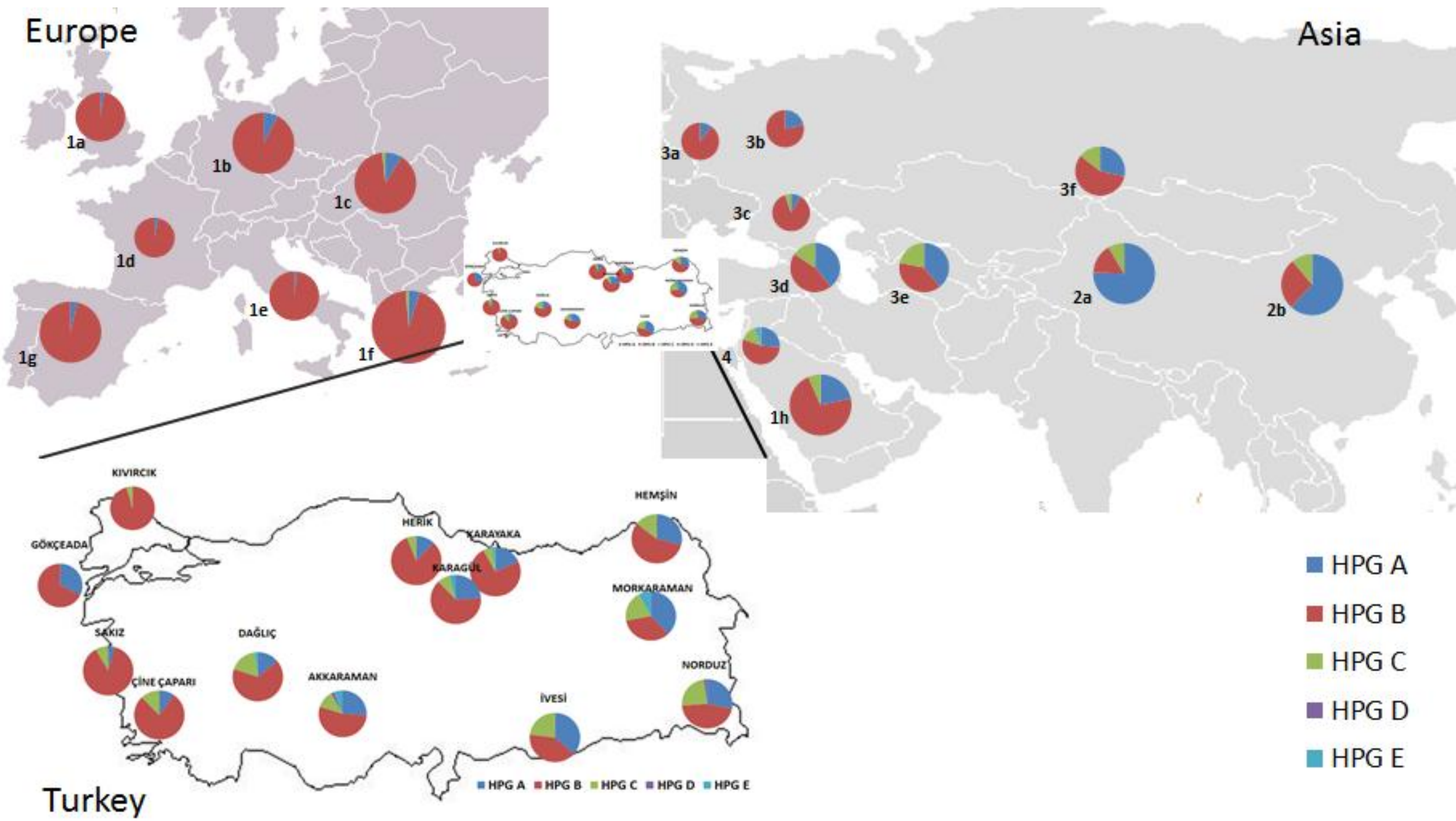


Figure 4.1 Haplogroup compositions of regions on Eurasia. 1a) United Kingdom (n=124); 1b) Poland and Germany (n=310); 1c) Romania and Hungary (n=248); 1d) France (n=31); 1e) Italy (n=155); 1f) Albania and Greece (n=403); 1g) Portugal and Spain (n=217); 1h) Saudi Arabia (n=93); 2a) northwest China (n=94); 2b) northeast China (n=137); 3a) west Russia and Russian Karelia (n=38); 3b) Middle Volga and Volga-Kama (n=64); 3c) Stravropol and Caspian depression (n=56); 3d) south and north Caucasus (n=102); 3e) east Caspian (n=23); 3f) Altay (n=21); 4) Bet Shean Valley, Israel (n= 57). Samples were taken from Econogene project (http://www.econogene.eu/) (1a-h), Guo *et al*.'s (2005) study (2a-b), Tapio *et al.*'s (2006) study (3a-f), and Meadows *et al.*'s (2007) study (4). Samples belong to Turkey were same as in Figure 3.11.

In Figure 4.1, HPG A and C were widely seen in Asia, whereas HPG B was widespread in Europe. On the other side, in the junction of these continents, Turkey was neither similar to Asia nor Europe in relation to domestic sheep mtDNA haplogroup composition. Yet, the gradual change of HPG A and C to HPG B from Asia to Europe can be observed over the Turkey. Although, eastern part of Turkey was rich or equal in composition of HPG A, B and C like Asia, western part of Turkey had breeds in high frequency of HPG B like Europe. This pattern supports the migration (first and/ or the second)ways of sheep through Europe from Middle East where the domestication first emerged and presumably the second expansion of sheep took place. The pattern observed (using the haplogroups observed in 106 flocks) in Spatial Correlation Analysis fits the intrusion model given in Sokal's (1979) study and long distance differentiation model in Barbujani *et al.*'s (1994) study. On the latter model, which is considered as ancient cline, observed pattern is believed to be generated by the effects of successive gene flow, drift and/or adaptation to local environmental factors have been superimposed (Bertorelle and Barbujani, 1995).

The genetic difference between Karayaka- Kıvırcık- Sakız and Morkaraman- Ivesi in terms of haplogroup diversity as visualized in median joining analysis (Figure 3.12) is another manifestation of the existence of two genetically as well as morphologically different sheep groups in Turkey.

In relation to domestic sheep evolution retrotypes indicated that "second sheep expansion" is governing the diversity pattern in sheep (Chessa *et al.,* 2009). If it involves mainly the migration of rams and limited migration of the ewes, in that case mtDNA pattern could still be reflecting the first migration of sheep in many parts of the world. In Turkey intrusion of HPG A and or C from east and southeast of Anatolia to northwest of Anatolia can be associated with a migration such as defined by Chessa *et al.* (2009) from Middle East to Europe.

The similarity pattern between breeds measured by haplogroup diversity was investigated by Kernel density estimation (KDE) on Synthetic maps. Since the distribution of data points over Turkey was not even and hence not suitable for interpolation KDE was the only option. The three patches seen on the synthetic map can be converted into two by combining the black patches seen on Figure 3.14 as there is no data between them (the area was shaded on Figure 3.14). So, the breeds can be separated by two relatively distinct zones: western-northwestern and central-eastern Turkey. If the microsatellite data is confirming the presence of these zones then priority setting for the conservation of sheep can be made within these zones seperately.

In Turkey breeds are not fully isolated (Açan, 2012). Therefore, thin, semi fat and fat tailed breeds are all admixed in various degrees. Yet, there is a global pattern for the tail types of the sheep in Turkey. In the KDE the white patch describes the fat tailed and to a certain degree fat rump tail (Hemsin) sheep breeds of Turkey. Two black patches correspond to thin tail or to a large degree to their hybrids where HPG B is relatively more frequent. Either the white or the black patches might be representing the intrusions. Some of them for instance, Sakız (Chios) might be the host of the Anatolia before the arrival of fat tail. With the private allele H12 in males (Meadows and Kijas, 2009) it seems to be quite isolated from European and Asian breeds. Since Asian haplotype H4 (Meadows and Kijas, 2008) is in low frequency but almost exclusively observed in white patch in Turkey (Öner *et al.*, 2011) and since it was never observed in Europe suggests that fat tail sheep migration is not competable with the second sheep migration. Alternatively it can be argued that emergence of H4 was more recent than the suggested second expansion of sheep or H4 with a low frequency might be lost by founder effects during the second migration. Yet quite distinct division line within the Anatolia in the absence of physical barrier at least between the breeds of western Anatolia suggests that perhaps arrival of fat tailed sheep together with HPG A and especially with HPG C (Tapio *et al*., 2006) must be pointing out another major migration, third migration of sheep brought along with the nomadic Turks arriving to Anatolia almost 1000 years ago.

The light brown color in Anatolia in Figure 4.2 is the borderline between Byzantine and Anatolian Seljuk Sultanate which was established by the nomadic Turks after 1071 (Toynbee, 1970; Lewis, 1995). This border of the sultanate stayed almost the same for another two centuries. May be the fat tailed sheep were kept behind the borders for 200 years. The similarity between the division of mtDNA haplogroups and the northern, western borders of Anatolian Seljuks is striking.

Figure 4.2 The borders of the Byzantine Empire in the 11th century. Picture was taken from http://crusadinghistory.wikispaces.com/Byzantine+Empire at 5/14/2012.

A study on shepherd dogs of Turkey confirms the division line between fat and thin tail breeds and suggests that the dog associated with the third migration to the fat tail seemed to accompany the sheep all the way from eastern Caspian Sea (Koban *et al.*, 2009).

Hence, results of the present study suggested that if there were two major migrations of domestic sheep in Anatolia there might be another migration of sheep associated with fat tailed sheep with an effect on genetic diversity distribution of sheep of Anatolia.

4.3 Sequences and some analyses of sheep mtDNA segments

At the start of the research it was thought that there may be a spatial distribution within the haplogroups (for instance haplotypes which are genetically more similar to wild sheep may exist in the east of Turkey). Therefore, equal number of representatives, 5 if possible, from each breed was selected for each halogroup in this study. However, within haplogroup spatial patterns were not searched yet, as the haplogroup composition of breeds was not represented in these sequences the comparisons in the breed level was not made.

After detecting that some individuals who exhibited HPG C band pattern by ND2-SSCP method had HPG E sequences, first of all HPG band pattern was called as HPG C/E and then all HPG C/E individuals were sequenced. This strategy gave the opportunity of obtaining large number of HPG C sequences (n=69) hence tests on HPG C sequences were reproduced with high confidence. Since all HPG C/E individuals were sequenced new 11 sequences of HPG E were obtained and number of available sequences, compatible with the presented ones (i.e. complete mtDNA CR), increased from 4 to 15. Similarly, all the individuals with haplogroup D (n=2) were sequenced again to increase the number of available sequences (n=3) (for partial/complete mtDNA) in the literature.

Nowadays, the whole mtDNA sequences, mitogenome are obtained for the 16 selected sheep where all 5 haplogroups observed in domestic sheep were represented (Meadows *et al.* 2011). In the present study, control region sequences were used mainly. Yet, in the detection of mtDNA based haplogroups and phylogenetic analyses, mtDNA control region (CR) was revealed as the most informative region on mtDNA by Meadows *et al.* (2011). They showed that the highest contribution to the topology constructed by mitogenome comes from variations in CR. Also the topologies drawn with CR and whole mitogenome on NJ tree were identical. So, at least for the domestic sheep, it is highly likely that the NJ tree constructed by mtDNA CR in the present study represents the true topology of the haplogroup relationship.

When phylogenetic tree using DNA sequences is searched, to find the correct topology of the tree, correct nucleotide substitution model must be used. Nucleotide substitution model of 240 mtDNA CR sequences was examined according to Bayesian Information Criterion (BIC) (Schwarz, 1978) and Akaike Information Criterion (AIC) (Akaike, 1974) with jModelTest 0.1.1 (Posada, 2008).

According to Akaike Information Criterion TPM1uf (Kimura, 1981) model with gamma-distributed rate heterogeneity (Γ=0.43) and proportion of invariable site (0.621), simply called TPM1uf+I+G model which has 5 free parameters, was found to be appropriate model for the studied region.

According to Bayesian Information Criterion HKY (Hasegawa, Kishino, and Yano 1985) model with gamma-distributed rate heterogeneity (Γ=0.429) and proportion of invariable site (0.621), simply called HKY+I+G model which has 4 free parameters, was found as the suitable model for the studied region.

Previous studies (Hiendleder *et al.*, 1998a; Tapio *et al.*, 2006; Meadows *et al.*, 2007), which have studied sheep mtDNA CR (partially/completely) sequences, indicated that HKY model was appropriate for the mentioned region. Although the results of AIC generally used for ecological data, to be compatible with previous studies, the result of BIC, which was HKY+I+G model, was used in the presented study.

NJ tree revealed that 11 individuals with 9 haplotypes detected as HPG E because they clustered with reference sequences from Meadows *et al.* (2011). These haplotypes observed in Akkaraman, Morkaraman, Dağlıç, Karagül and Karayaka breeds which were known as fat-tailed sheep breeds except Karayaka (however it must be mixed Açan, 2012). Previously identified (as HPG E) four individuals from Tuj and Awassi (which was a fat tail population from Israel and its counterpart is Ivesi in Turkey) breeds in Meadows *et al.*'s (2007) study and one individual from Mongolian sheep and one from Kazakh fat-rumped breed in Guo *et al.*'s (2005) study were known as fat- tailed (except Kazakh fat-rumped, which deposit the fat in their rump). If individuals from Karayaka sheep were ignored it may be considered as the presence of HPG E were associated with fat tail like in HPG A and C.

In the case of rare HPG D, 1 individual from Akkaraman breed and 1 individual from Norduz breed as different haplotypes were observed in NJ tree (Figure 3.7). Previously identified two individuals were from Morkaraman breed (Meadows *et al.*, 2007) and one individual from Karachai breed (Tapio *et al.*, 2006). All individuals (identified in both present and previous studies) belong to HPG D were known as fat-tailed. Although sample size of HPG D is small, the association of fat tail and HPG D may be considered as in HPG A, C and E.

The topology of NJ tree drawn by complete mtDNA CR sequences (886 bp long) in the presented study is similar to the one constructed by partial mtDNA CR sequences (432 bp long) in Tapio *et al.*'s (2006) study. In both HPG C and E were on one side of wild sheep split, HPG A, B and D were on the other side where HPG D placed between the HPG A and B. This topology were not seen in NJ tree drawn by *cytB* region in Meadows *et al.*'s (2007), as the HPG D were placed closer to wild sheep than other haplogroups. Yet, as the CR revealed as more informative region than *cytB* (Meadows *et al.*, 2011), the difference between the two NJ tree topologies were expected. Similar to NJ tree, HPG D is closer to HPG A in, MJ network in Meadows et al.'s (2007) and Olivieri *et al*.'s (2012) studies respectively where combined sequences of mtDNA control and *cytB* regions were used.

Nucleotide diversity (π) is a sample size independent measure, which defines the probability of randomly selected two sequences being different than each other. π was highest in HPG E (0.00662), secondly in HPG B (0.00550) the lowest in HPG C (0.00289). These observations also appeared in MJ network as dispersed haplotypes without central node in HPG E, and also dispersed haplotypes with a central node in HPG B where the others showing closer haplotypes with a central node except HPG D which is represented by very few individuals hence the rule does not apply. As high nucleotide diversity observed in a haplogroup was considered as a sign of early domestication (Jobling, 2004), HPG E may be thought as one of the early, relatively locally domesticated haplogroup. May be it did not dispersed under the pressure of dominant haplogroups found in the same area like it was in the case of HPG C in goats (Naderi *et al.*, 2008). Yet, it may have received introgression from wild sheep equipped with HPG E related sequences which were found to be present in the vicinity of the distribution area (Rezai *et al*., 2010) of domestic HPG E as presented in the present study.

4.4 About the history and origin of domestic haplogroups

While using the mtDNA CR sequences, observation of a bell-shaped mismatch distribution together with significant negative neutrality tests in HPG A, B and C confirms the population expansion in the past of these haplogroups which were first revealed by Meadows *et al.* (2007) for HPG A and B and Chen *et al.* (2006) for HPG C. In Meadows *et al.*'s (2007) study, the sudden expansion of HPG A and B were confirmed by samples from Turkey, but not in HPG C probably due to small sample size (11 haplotypes of 27 individuals). In the presented study, only with the samples from Turkey, the past population expansion of HPG C was confirmed.

Population size changes of HPG E was investigated eventually together with the new members found in Turkey in this study and observed ones from previous studies (Guo *et al.*, 2005; Pedrosa *et al.*, 2005; Meadows *et al.*, 2011; Koban *et al.* (unpublished)). The bell-shaped mismatch distributions were observed for HPG E in analyses containing both all of the available sequences of HPG E (n=20) where the length of the sequences were 665 bp and only the sequences covering complete mtDNA-CR (n=15).

4.5 About *Ovis gmelinii anatolica*

To understand the phylogenetic relationship between *Ovis gmelinii anatolica* and both with domestic sheep and other wild types three classes of comparative studies according to the availability of the data were carried out as follows: 1) CR sequences of *Ovis gmelinii anatolica* (n=30) were compared with the sequences of domestic individuals (240) used in the present study, 2) compatible mtDNA *cytB* sequences of selected *Ovis gmelinii anatolica* and domestic individuals of the present study were compared with those of largely wild sheep (*Ovis gmelinii anatolica*, *Ovis gmelinii orientalis, Ovis orientalis laristanica, Ovis orientalis isphahanica*, *Ovis orientalis X vignei and Ovis vignei*) given by Rezai *et al.* (2010) 3) Both CR and *cytB* sequences of selected *Ovis gmelinii anatolica* and domestic individuals of the present study were compared with the compatible region of those of mainly domestic sheep given by Meadows *et al*. (2005), Meadows *et al*. (2007), Pardeshi *et al.* (2007) and Olivieri *et al.* (2011).

First of all *Ovis gmelinii anatolica* on the contrary of early predictions (Hiendleder *et al.*, 2002) did not harbor haplogroup B related mtDNA CR (Figure 3.15- 3.19). It can be argued that *Ovis gmelinii anatolica* might have been harboring HPG B but during the bottlenecks in particular to the most recent bottleneck where the population size decreased to below 35 (Sezen, 2000) HPG B was lost. As effective population size is almost always considerably less than the census population size (Frankel and Soulé, 1981) and effective population size of mtDNA is ¼ th of that of nuclear chromosome mtDNA must have went through a very severe bottleneck, recently. Yet, observation of different haplotypes (3 haplotypes in Rezaei *et al*.'s (2010) study: 2 of them were identical to haplotypes in the present study, 1 of them was similar with 1 mutation difference to X) indicated that after the bottleneck mtDNA effective population size was at least 3.

Yet, this low effective mtDNA size would be under severe random genetic drift. Since, probability of losing a haplotype is related with its frequency in the population the least frequent haplotypes must have been lost during the bottleneck. If HPG B was lost in the last bottleneck it perhaps was not a very frequent haplotype.

Since, haplotype of HPG A in *Ovis gmelinii anatolica* was observed in *Ovis gmelinii orientalis and Ovis gmelinii isphahanica*, wild sheep, in Rezaei *et al.*'s (2010) study and in 1 of our domestic sheep from Herik breed, and it is in the central node both with respect to our samples (Figure 3.15) and in MJ network of Olivieri *et al.*'s (2012) study (Figure 3.16-3.19) which covers the domestic sheep one can suggest that HPG A of *Ovis gmelinii anatolica* is a product of either introgression of fully domestic sheep (not very likely because of the size differences between modern domestic sheep and wild sheep as was already noted by Pedrosa *et al.*, 2005) or with a high probability was a feral form of domestication process.

However, X haplotype is neither observed in domestic sheep nor in wild sheep other than *Ovis gmelinii anatolica*. In MJ network which is constructed with samples from Rezai *et al.*'s (2010) study based on *cytB* (Figure 3.19) as well as X, 1 haplotype

(H_25) closely related to X (H_12) (with one mutational difference) were observed. There is no measure to define a fully wild or feralized form of sheep along the domestication process. Only the European mouflons have Dxy=0.00749 with HPG B (Table 3.9). X is in between with HPG E and HPG C and being higher than C exhibits Dxy= 0.01486 with HPG C, based on the same segment of mtDNA which was used to obtain Dxy of European Mouflon. Hence, if it is a feral form of domestication process it has twice the difference that was observed between European mouflon and HPG B. Since the difference between E and C among the domestic sheep in mtDNA CR is less than the difference between C and X and X seems to be close to C, X might be the extent of the ancestral population given rise to domestic sheep with HPG C. However, in the network based on Rezai *et al.*'s 2010 study it can be considered as also just one of the wilds or domestic ferals of the wild which gave rise to HPG E and HPG C of the domestic sheep. Interestingly in this network (Figure 3.19) HPG C suggests itself as an extent of HPG E.

Similar to wild goats (Naderi *et al*., 2008) among the wild sheep there seems to be many wild haplotypes which were not observed in domestic sheep (Rezaei *et al*., 2010) as observed in Figure 3.19. Also in this figure X, E and C relationship together with the related wild haplotypes observed in *Ovis gmelinii orientalis*suggest that there are related haplotypes, where one of them is X and haplotype observed by Rezaei *et al.*, (2010) in *Ovis gmelinii anatolica* population, which can be called as a family of C and E. May be members of this family were subjected to more than one expansion given rise to E and C.

Another interesting observation was that among the examined wild sheep, there was no wild sheep carrying HPG B related sequence other than a single *Ovis gmelinii orientalis* which seems to carry identical sequence with the domestic HPG B sequence (Rezai *et al.* 2010) These observations supports that some wild sheep ancestors of haplogroups might become extinct (Meadows *et al.,* 2011) in particular HPG B. Perhaps, the wild ancestors of this haplogroup were extinct as was suspected by Meadows *et al*. (2011).

Since HPG B was predominat in Europe (Figure 4.1), since early migration presumably involving the primitive sheep of early domestication products (Chessa *et al.,* 2009) seems to have HPG B exclusively for instance in mouflons of Sardinia, Cyprus and primitive sheep of Soay (Bruford and Towsend, 2006) it can be suggested that HPG B ancestors of sheep was dominant in the western part of the domestication center.

Evolution of *Ovis gmelinii anatolica* is unknown. If in early days of sheep domestication, *Ovis gmelinii anatolica* was connected to a large continuous *Ovis gmelinii* population the population defined for that time would be covering the large segment of the sheep domestication center depicted in Zeder's (2008) study. It would also be covering the monumental archaeological site of Göbekli Tepe where early steps of domestication (keeping animals alive in captivity for some time) might have been taken. Just as was parallel to the plausible scenario for the goat domestication (Naderi *et al.,* 2008) it could have harbor females possessing many haplogroups. May be HPG B being dominant in the western part of the range. As was pointed out by Arıhan (http://www.metu.edu.tr/~cbilgin/gmelinii.htm) it was suggested *Ovis gmelinii anatolica* might be isolated from the main *Ovis gmelinii* population around 6000 B.P. (Adams, 1997). Since then it might have been further fragmented and subjected to drift. Along this process ancestors of B could have become extinct and the haplotype pool is reduced to A and X related haplotypes perhaps representing individuals from different stages of domestication.

CHAPTER 5

CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDIES

In the present study, history of domestic Turkish sheep breeds and *Ovis gmelinii anatolica* were investigated based on mtDNA control region and *cytochrome B* haplotypes. These molecular markers enabled us to merge the present data with the data from previous studies, some of which (sequences) could be retrieved from the GenBank (http://www.ncbi.nlm.nih.gov/). Hence, meta analyses contributing to the understanding of both evolutionary history of domestic sheep and sheep domestication process could be done. For the analyses, Neighbor joining (NJ), Median joining (MJ), mismatch distribution, nucleotide based diversity calculation, Bayesian skyline plot (BSP) and Kernel density estimation (KDE) methods through MEGA 5 (Tamura *et al.*, 2011), NETWORK 4.6.1.0 (http://www.fluxus-engineering.com/), Arlequin 3.11 (Excoffier, 2005), DnaSP v.5 (Librado and Rozas, 2009), BEAST v.1.6 (Drummond and Rambaut, 2007) and ArcGIS 10 (www.esri.com) softwares respectively were carried out.

The conclusions of the present study can be listed as follows:

1) The relationships of haplogroups (HPG A-E) with each other observed in this study were concordant with those given in previous studies.

2) With the newly identified and sequenced members of HPG E of the present study, the population expansion of HPG E as was previously observed for HPG A, B and C was confirmed.

3) For the first time in this study, the haplogroup identification of Anatolian wild sheep (*Ovis gmelinii anatolica*) was performed by mtDNA CR sequences. Observed HPG A and X were thought as the haplotypes remaining from the early existing wider population form of *Ovis gmelinii* which was containing repertoire of haplotypes related with HPG A, C, E and even possibly, B.

4) With the examination of HPG D in ND2-SSCP method, new banding pattern specific to HPG D were observed. This observation made ND2-SSCP the best method to identify mtDNA based haplogroups among ND4-SSCP and CR-RFLP methods.

5) As all members of HPG D and most of the members of HPG E observed in fat-tailed sheep breeds, these haplogroups may be associated with fat tail as HPG A and C which were frequent in Asia and Middle East.

6) According to the relationships between the haplotypes within the haplogroups revealed by the meta analyses (MJ network) based on mtCR-*cytB* sequences of the wild sheep and the domestic sheep, most probably, the wild ancestors of HPG B is extinct.

7) The spatial distribution of haplogroups analyzed by Spatial Autocorrelation Analyses revealed the intrusion and long distance differentiation model probably due to intrusion of HPG A and C together with fat tailed-sheep to eastern Turkey from Middle East during the third migration of sheep into already existing sheep of Anatolia perhaps best represented by Sakız. This third migration scenario is suggested for the first time in the present study as findings neither supported the second migration of sheep defined by Chessa *et al.* (2009) nor first migration of sheep suggested in Zeder's (2008) study. The similarity of the borders of Byzantine Empire-Anatolian Seljuk Sultanate (founded by the nomadic Turks whose life style depends on pastoral animal husbandry) and the division line between the two relatively homogenous area

of sheep in Synthetic map drawn by KDE was supporting the third migration scenario.

8) What ever the reason is the presence of mtDNA based two distinct sheep groups broadly associated with fat and thin tail sheep requires that conservation studies of sheep breeds must be done seperately within these groups seperately.

For further studies that can be done in Turkey suggestions are as follows:

1) To perform more reliable spatial analyses, more uniform data covering samples from more evenly located sampling sites is needed. Hence, new samples to fill the gaps would be useful.

2) Mitogenomes (may be 5-6) of *Ovis gmelinii anatolica* individuals must be obtained and can be presented on the tree given by Meadows *et al*. (2011). Phylogenetic relationship between the haplogroups of the domestic sheep and haplogroups of *Ovis gmelinii anatolica* can then be fully resolved.

3) Retrotypes defined by Chessa *et al*. (2009), both within the domestic sheep of Anatolia and *Ovis gmelinii anatolica* would provide to the understanding of the sheep migrations by providing an interpretation from the perspective of a neutral nuclear marker.

4) Data composed of Y-Chromosome related markers must be enlarged to better understand the migrations and their components.

5) In relation to ancient DNA (aDNA) mtDNA studies, maternal history of sheep

    i) aDNA mtDNA haplogroup identification study from the western part of ancient wild sheep distributions would be useful to check the hypotheses that "wild sheep harboring HPG B related haplotypes were in the western fringe of wild sheep distribution" and "It became extinct sometime after the first domestication".

ii) aDNA mtDNA haplogroup identification study, using the samples unearthed around the southeast of Anatolia in different time intervals, would contribute to the understanding of early domestication of sheep from the maternal side.

iii) aDNA mtDNA haplogroup identification study using the samples unearthed around all over Anatolia in different time intervals, would also contribute to the understanding of incoming domestic female sheep migrations to Turkey.

REFERENCES

Acar, H. (2010). Bioinformatic Analyses In Microsatellite-Based Genetic Diversity Of Turkish Sheep Breeds. MSc Thesis. The Graduate School Of Informatics, Middle East Technical University.

Açan, S.C. (2012). Genetic Diversity Of Turkish Sheep Breeds In The Focus Of Conservation Reasearch. PhD Thesis, Department of Biology, Middle East Technical University, Ankara, Turkey.

Adams, J.M. (1997). Global Land Environments Since the Last Interglacial. Retrieved March 30, 2012, from Oak Ridge National Laboratory, The Environmental Sciences Division Web site: http://www.esd.ornl.gov/ern/qen/nerc.html.

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, *19*:716-723.

ArcGIS 10 for Server Functionality Matrix (2011). Environmental Systems Research Institute. Redlands, CA.

Arıhan, O. (2000). Population Structure, Habitat Use, and Reintroduction of Anatolian Wild Sheep *Ovis gmeliniii anatolica*. MSc thesis. Biology Department, Middle East Technical University, Ankara, Turkey.

Arıhan, O. and Bilgin C.C. (2002) Turkish Mouflon (*Ovis gmelinii anatolica*). Retrieved March 30, 2012, from Middle East Technical University, Department of Biology, Bilgin Lab of Biodiversity and Conservation Web site: (http://www.metu.edu.tr/~cbilgin/gmelinii.htm)

Aris-Brosou, S., and Excoffier, L., (1996). The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Mol. Biol. Evol.* 13: 494-504.

Bandelt, H-J., Forster, P. and Röhl, A. (1999). Median-joining Networks for Inferring Intraspecific Phylogenies. *Mol. Biol. Evol., 16*:37-48.

Barbujani, G., Pilastro, A., De Domenico, S. and Renfrew, C. (1994). Genetic Variation in North Africa and Eurasia: Neolithic Demic Diffusion vs. Paleolithic Colonisation. *American Journal of Physical Anthropology, 95*:137-154.

Barker, G., (2002). In P. Bellwood and C. Renfrew (Eds.), *Examining the Farming/Language Dispersal Hypothesis* (pp. 151–162). McDonald Institute for Archaeological Research, Cambridge.

Bassam, B. J., Anollés, G. C. and Gresshoff, P. M. (1991). Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry, 196*:80-83.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Wheeler D. L. (2005). *GenBank* [database]. Maryland, USA: National Center for Biotechnology Information [Producer and Distributor].

Bertorelle, G. and Barbujani, G. (1995). Analysis of DNA diversity by spatial autocorrelation. *Genetics, 140:* 811-819.

Bruford M.W., Townsend S.J. (2006) Mitochondrial DNA diversity in modern sheep: Implications for domestication. In: Zeder M.A., Bradley D.G., Emshwiller E., Smith B.D. (Eds.) Documenting Domestication: New Genetic and Archaeological Paradigms. University of California Press, CA,USA, 307-317.

Bruford, M. W., Bradley, D. G. and Luikart G. (2003). *DNA markers reveal the complexity of livestock domestication.* Nature Review Genetics, 4: 900-910.

Bunch T. D., W. C. Foote, and Spillett, J.J. (1976). Translocations of acrocentric chromosomes and their implications in the evolution of sheep (*Ovis*). *Cytogenetics and Cell Genetics,* 17 (3):122–136.

Chen, S. Y., Duan, Z. Y., Sha, T., Xiangyu, J., Wu, S. F. and Zhang, Y. P. (2006). Origin, genetic diversity, and population structure of Chinese domestic sheep. *Gene, 376*: 216–223.

Chessa, B., Pereira, F., Arnaud, F., Amorim, A., Goyache, F., Mainland, I. *et.al.* (2009). Revealing the History of Sheep Domestication Using Retrovirus Integrations. *Science, 324*: 532-536.

Chromas Pro version 1.52. 2003-2008. Technelysium Pty Ltd, Tewantin QLD.

Clutton-Brock J. (1999). *A natural history of domesticated mammals*. Cambridge University Press, UK.

Dobney K. and Larson G. (2006). Genetics and animal domestication: new windows on an elusive process. *Journal of Zoology* 269 (2): 261–271.

Doğan A.Ş. (2009). Reassessment of Genetic Diversity in Native Turkish Sheep Breeds with Large Numbers of Microsatellite Markers and Mitochondrial DNA (mtDNA). MSc Thesis. Biology Department, Middle East Technical University, Ankara, Turkey.

Drummond, A.J. and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology,* **7**: 214.

Drummond, A.J., Rambaut, A., Shapiro, B. and Pybus, O.G. (2005). Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Mol Biol Evol,* 22:1185-1192.

Drummond, A.J., Ho, S.Y.W., Rawlence, N., and Rambaut, A., (2007, July 6). A Rough Guide to BEAST 1.4. Retrieved February 21, 2011, from http://beast.bio.ed.ac.uk/Main_Page

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

Farris J. S. (1970). Methods for Computing Wagner Trees. *Syst Biol* 19(1): 83-92.

Fernández, H., Hughes, S., Vigne, J.D., Helmer, D., Hodgins, G., Miquel, C., *et al*. (2006). Divergent mtDNA lineages of goats in an Early Neolithic site, far from the initial domestication areas. *Proc Natl Acad Sci USA,* 103:15375–15379.

Festa-Bianchet, M., (2000). A summary of discussion on the taxonomy of mountain ungulates and its conservation implications. Workshop on Caprinae taxonomy, Ankara, Turkey.

Forster, P., Torroni, A., Renfrew, C. and Röhl, A. (2001). Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Mol Biol Evol, 18*:1864-1881.

Frankel, O. H. &. Soule, M. E. (1981). *Conservation and Evolution*. Cambridge: Cambridge University Press.

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics, 147*: 915-925.

Gatrell, A.C., Bailey, T.C., Diggle P.J. and Rowlingson B. S. (1996). Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. *Trans Inst Br Geogr NS*, 21: 256-274.

Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika,* 53: 325–328.

Groeneveld, L. F., Lenstra, J. A., Eding, H., Toro, M. A., Scherf, B., Pilling, D. et al. (2010). Genetic diversity in farm animals – a review. *Animal Genetics, 41*: 6-31.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology,* 52:696-704.

Guo, J., Du, L. X., Ma, Y. H., Guan, W. J., Li, H. B., Zhao, Q. J. et al. (2005). A novel maternal lineage revealed in sheep. *Animal Genetics, 36*:331-336.

Hall, T. (1997-2011). BioEdit: Biological sequence alignment editor for Win95/98/NT/2K/XP/7. Ibis Biosciences, Carlsbad, CA 92008. Retrieved June 15, 2011, from http://www.mbio.ncsu.edu/bioedit/bioedit.html.

Haltenorth, T., (1963). Klassification der Saugethiere: Artiodactyla. *Handbuch der Zoologie,* 8(32). Walter de Gruyter, Berlin, Germany.

Harpending., H.C., Batzer, M.A., Gurven, M., Jorde, L.B., Rogers, A.R., and Sherry, S.T. (1998). Genetic traces of ancient demography. *Proc. Natl. Acad. Sci. USA* 95:1961-1967.

Harris, D.R., Gosden, C., (1996). in The Origins and Spread in Agriculture and Pastoralism in Eurasia, D. Harris, Ed. (UCL Press, London), pp. 370–389.

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

Hiendleder, S., Mainz, K., Plante, Y. and Lewalski, H. (1998a). Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different maternal sources: No evidence for contributions from Urial and Argali sheep. *Journal of Heredity, 89*:113-120.

Hiendleder, S., Lewalski, H., Wassmuth, R. and Janke, A. (1998b). The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major *Ovine* haplotypes. *Journal of Molecular Evolution, 47*:441-448.

Hiendleder S., Kaupe B.,Wassmuth R. and Janke, A. (2002). Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. Proceedings: *Biological Sciences, 269*:893-904.

Hudson, R.R. (1990). Gene genealogies and the coalescent process. In Futuyama, and J.D. Antonovics (Eds.), *Oxford Surveys in Evolutionary Biology* (pp. 1-44). Oxford University Press, New York.

IUCN/SSC-Caprine Specialist Group (2000). Workshop on Caprinae taxonomy, Ankara, Turkey, May 8–10.

Jobling, M. A., Hurles, M. E. and Tyler-Smith, C. (2004). *Human Evolutionary Genetics: origins, peoples and disease.* London/New York: Garland Science Publishing.

Jukes, T. and Cantor, C. (1969). Evolution of protein molecules. In H.N. Munro (Ed.), *Mammalian Protein Metabolism* (p. 21-132). Academic Press, New York, NY.

Kaya M. A., Bunch T. D. and Konuk M. (2004) On Konya wild sheep, *Ovis orientalis anatolica,* in the Bozdag protected area. *Mammalia 68* (2-3): 229-232.

Kimura, M. (1981). Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci. USA., 78*: 454-8.

Klug, W. S. and Cummings, M. R. (2000). Concepts of Genetics. (6^{th} Ed.), Prentice Hall, New Jersey.

Koban, E. (2004). Genetic Diversity of Native and Crossbreed Sheep Breeds in Anatolia. PhD Thesis. Middle East Technical University, Ankara, Turkey.

Koban, E., Perez, T., Bruford M.W., Demirci, S. And Togan I. (Unpublished). Genetic observations on sheep from the center of domestication with some implications on conservation studies in Turkey. Submitted Manuscript.

Koban, E., Gökçek Saraç, Ç., Açan, S.C., Savolainen, P. and Togan İ. (2009). Genetic relationship between Kangal, Akbash and other dog populations. *Discrete Applied Mathematics, 157*: 2335-2340.

Kruskal, J.B. (1956). On the shortest spanning subtree of a graph and a travelling salesman problem. *Proc. Amer. Math. Soc.*, 7: 48-50.

Kuijt, I. (2002). *Life in Neolithic Farming Communities - Social Organization, Identity, and Differentiation*. New York, Boston, Dordrecht, London, Moscow: Kluwer Academic Publishers.

Lawson Handley, L.J., Byrne K., Santucci F., Townsend S., Taylor M., Bruford M. W. et al. (2007). Genetic structure of European sheep breeds. *Heredity 99*(6): 620-631.

Lewis, B. (1995). *The Middle East: A Brief History of the Last 2000 Years*. New York: Touchstone.

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

Luikart G., Gielly L., Excoffier L., Vigne J-D., Bouvet J., Taberlet P. (2001). Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc. Natl. Acad. Sci. USA, 98*:5927-5932.

Meadows, J.R.S., Li, K., Kantanen, J., Tapio, M., Sipos, W., Pardeshi, V. *et al*. (2005). Mitochondrial sequence reveals high levels of gene flow between sheep breeds from Asia and Europe. *J Hered* 96: 494–501.

Meadows, J. R. S., Cemal, I., Karaca, O., Gootwine, E. and Kijas, J. W. (2007). Five ovine mitochondrial lineages identified from sheep breeds of the Near East. *Genetics, 175*:1371-1379.

Meadows J.R.S., Hanotte O., Droögemüller C., Calvo J., Godfrey R., Coltman D.*et al*. (2006) Globally dispersed Y chromosomal haplotypes in wild and domestic sheep. *Animal Genetics* 37, 444–53.

Meadows, J. R. S., Hiendleder, S. and Kijas, J. W. (2011). Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. *Heredity,106:* 700–706.

Meadows, J.R.S. and Kijas, J. W. (2008). Re-sequencing regions of the ovine Y chromosome in domestic and wild sheep reveals novel paternal haplotypes. *Animal Genetics, 40:* 119–123.

Naderi S., Rezaei H., Pompanon F., Blum, M. G. B, Negrini, R., Naghash, H. et al. (2008). The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proc. Natl. Acad. Sci. USA, 105*: 17659–17664.

Naderi, S., Rezaei, H-R., Taberlet, P., Zundel, S., Rafat, S-A., Naghash, H-R. et al. (2007). Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. Econogene Consortium. *PLoS One. 2*:e1012.

Nadler, C. F., Korobitsina, K. V., Hoffmann, R. S. &Vorontsov, N. N. (1973). Cytogenetic differentiation, geographic distribution and domestication of palaearctic sheep (*Ovis*). *Zeitschrift jiir Sarcgetierkrtrzde*, 38: 109-1 25.

Nei, M. (1987). *Molecular Evolutionary Genetics*. New York, NY: Columbia University Press.

NETWORK version 4.6.1.0 Free Phylogenetic Network Software. Downloaded January 24, 2012, from Web site: http://www.fluxus-engineering.com/sharenet.htm.

Olivieri, C., Ermini, L., Rizzi, E., Corti, G., Luciani, S., Marota, I. et al. (2012) Phylogenetic Position of a Copper Age Sheep (*Ovis aries*) Mitochondrial DNA. *PLoS ONE* 7(3):e33792.

Öner, Y., Calvo, J. H. and Elmaci, C. (2011). Y chromosomal characterization of Turkish native sheep breeds. *Livestock Science 136*: 277-280.

Pardeshi, V. C., Kadoo, N. Y., Sainani, M. N., Meadows, J. R. S., Kijas, J. W. and Gupta, V. S. (2007). Mitochondrial haplotypes reveal a strong genetic structure for three Indian sheep breeds. *Animal Genetics* 38 (5): 460–466.

Pariset, L., Mariotti, M., Gargani, M., Joost, S., Negrini, R., Perez, T. (2011). Genetic Diversity of Sheep Breeds from Albania, Greece, and Italy Assessed by Mitochondrial DNA and Nuclear Polymorphisms (SNPs). *The Scientific World Journal*, 11: 1641–1659.

Peakall, R. and Smouse, P. E. (2006). Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes, 6:* 288–295.

Pedrosa S., Uzun M., Arranz J., Gil B. G., Primitivo F. S. and Bayon, Y. (2005). Evidence of three maternal lineages in near eastern sheep supporting multiple domestication events. *Proc. R. Soc. B., 272:*2211-2217.

Pereira, F., Davis S.J.M., Pereira, L., McEvoy, B., Bradley, D.G., and Amorim, A. (2006). Genetic signatures of a Mediterranean influence in Iberian Peninsula sheep husbandry. *Mol. Biol. Evol.,* 23: 1420–1426.

Peter, C., Bruford M., Perez T., Dalamitra S., Hewitt G., Erhardt G. et al. (2007). Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. *Animal Genetics, 38*: 37-44.

Poplin, F. (1979). Origine du mouflon de Corse dans une nouvelle perspective paleontologique, par marronnage. [Origin of the Corsican mouflon in a newpaleontological perspective, for runaways.] *Annales de Ginitique et Silection Aninzale* 11, 133-143.

Posada, D. (2003). Using Modeltest and PAUP* to select a model of nucleotide substitution. In A.D. Baxevanis, D.B. Davison, R.D.M. Page, G.A. Petsko, L.D. Stein, and G.D. Stormo (Eds.), *Current Protocols in Bioinformatics* (pp. 6.5.1-6.5.14). John Wiley and Sons, Inc.

Posada, D. (2008). jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution, 25*: 1253-1256.

Price, T.D. (2000). *Europe's First farmers.* Cambridge Univ. Press, Cambridge

Rambaut, A. and Drummond A.J. (2007). Tracer version 1.4. Retrieved February 21, 2011, from http://beast.bio.ed.ac.uk/Tracer.

Rezaei, H.R, Naderi, S., Chintauan-Marquier, I.C., Taberlet, P., Virk, A.T., Naghash, H.R. et al. (2010). Evolution and taxonomy of the wild species of the genus *Ovis* (Mammalia, Artiodactyla, Bovidae). *Molecular Phylogenetics and Evolution* 54: 315–326.

Ryder, M.L. (1983). *Sheep and Man*. London: Duckworth.

Rzhetsky A. and Nei M. (1992). A simple method for estimating and testing minimum evolution trees. *Molecular Biology and Evolution* 9:945-967.

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

Sambrook, J., Foritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. Second Ed., Cold Spring Harbor Laboratory, New York.

Sanger, F., Donelson, J. E., Coulson, A. R., Kössel, H. and Fischer, D. (1973). Use of DNA Polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage f1 DNA, *Proc. Natl. Acad. Sci. USA, 70*: 1209–1213.

Schmidt, K. (2007). *Taş Çağı Avcılarının Gizemli Kutsal Alanı Göbekli Tepe En Eski Tapınağı Yapanlar*.[The Mysterious Sanctuary ofStone AgeHunters: Builders of The World's Oldest Temple]. Arkeoloji ve Sanat yayınları, İstanbul.

Schwarz, G. 1978. Estimating the dimension of a model. *The Annals of Statistics* 6:461-464.

Sezen (Yüksel), Z. (2000). Population viability analysis for reintroduction and harvesting of Turkish mouflon *Ovis gmelinii anatolica* valenciennes 1856**.** MSc thesis. Department of Biology, Middle East Technical University METU, Ankara, 2000.

Shackleton, D. M., ed. (1997). *Wild Sheep and Goats and their Relatives: Status Survey and Conservation Action Plan for Caprinae* (IUCN, Gland, Switzerland).

Shackleton, D.M. and Lovari, S., (1997). Classification Adopted for the Caprinae Survey. In D.M Shackleton (Ed.), *Wild Sheep and Goats and their Relatives. Status survey and conservation action plan for Caprinae* (pp. 9–14). IUCN, Gland, Switzerland.

Smouse, P.E. and Peakall, R. (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity. 82*, 561-573.

Sokal, R. R. (1979). Ecological parameters inferred from spatial correlograms. In G. P. Patil and M. L. Rosenzweig [eds.], Contemporary quantitative ecology and related ecometrics. Maryland: International Cooperative Publishing House.

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

Tajima, F. (1993). Measurement of DNA polymorphism. In N. Takahata and A.G. Clark (Eds.), *Mechanisms of Molecular Evolution. Introduction to Molecular Paleopopulation Biology* (pp. 37-59). Sinauer Associates, Inc., MA:Japan Scientific Societies Press, Sunderland, Tokyo.

Tajima, F., (1996). The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics* 143: 1457-1465.

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution, 10*:512-526.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution, 28*: 2731-2739.

Tapio, M., Marzanov, N., Ozerov, M., Cinkulov, M., Gonzarenko, G., Kiselvova, T. et al. (2006). Sheep Mitochondrial DNA Variation in European, Caucasian and Central Asian Areas. *Mol. Biol. Evol., 23*: 1776-1783.

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. In R.M. Miura (Ed.), *Some mathematical questions in biology - DNA sequence analysis*. (pp. 57-86) Amer. Math. Soc., Providence, RI.

Thompson, J.D., Higgins, D.G. and Gibson, T.C. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*., 22(22): 4673–4680.

Toro, M.A., Fernández, J., Caballero, A. (2009). Molecular characterization of breeds and its use in conservation. *Livestock Science,* 120: 174–195.

Townsend, S.J. (2000). Patterns of Genetic Diversity in European Sheep Breeds. PhD Thesis. University of East Anglia, United Kingdom.

Toynbee, A.J. (1970). *The Western Question in Greece and Turkey. A Study in the Contact of Civilizations.* 3rd ed. New York: Howard Fertig, Inc.

Tsalkin, V.I., (1951). European and Asian mountain sheep (Gornje barany evropy I asii). Moskovskoe Obshchest. Iss. Privody., Moscow Society of Naturalists Zool.

Sect.(Moskovskoe Obshchest. Isp. Prirody) Materials for the recognition of the fauna and flora of the USSR 27. Moscow.

Uzun, M., Gutierrez-Gil, B., Arranz, J., Primitivo, F., Saatci, M., Kaya, M. *et al*. (2006). Genetic relationships among Turkish sheep. *Genet. Sel. Evol.,* 38: 513–524.

Valdez, R. (1982). *The Wild Sheep of the World*. Wild Sheep and Goat International, Mesilla, New Mexico.

Vigne, J.-D. (1999). The large 'true' Mediterranean islands as a model for the Holocene human impact on the European vertebrate fauna? Recent data and new reflections. In *The Holoceze history, of the European vertebrate fauna. Modern aspects of research* (ed. N. Benecke), Rahdenmestf, Germany: Verlag Marie Leidorf GmbH.

Vigne, J.-D., Carrère, I. and Guilaine, J. (2003). Unstable status of early domestic ungulates in the Near East: the example of Shillourokambos (Cyprus, IX-VIIIth millennia cal. BC). In *Le Néolithique de Chypre, 43* (eds J. Guilaine and A. Le Brun), pp. 239–251. Greece: Bull. Corr. Héllenique.

Wilson, D.E. and Reeder, D.M., (1993). *Mammal Species of the World: A Taxonomic and Geographic Reference*. Smithsonian Institution Press, Washington, D.C.

Wood, N. J. and Phua, S. H. (1996). Variation in the control region sequence of the sheep mitochondrial genome. *Animal Genetics, 27*: 25-33.

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

Wright, S. (1965). The interpretation of population structure by F-Statistics with special regard to systems of mating. *Evolution, 19*(3):395-420.

Yüncü, E., Demirci, S., Koban Baştanlar E., Togan, I. (Unpublished) Comparison of three simple methods in relation to mtDNA haplogroup identification of Domestic Sheep. Submitted Manuscript.

Zeder, M. A., Dan Bradley, Eve Emshwiller, and Bruce Smith (editors). (2006). *Documenting Domestication: New Genetic and Archaeological Paradigms* University of California Press.

Zeder, M. A. (2008). Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proc. Natl. Acad. Sci. USA, 105*(33):11597-11604.

Ziaie, H. (1997). Iran, Middle East. In D.M. Shackleton (Ed.), *Wild Sheep and Goats and their Relatives: Status Survey and Conservation Action Plan for Caprinae* (pp. 53). IUCN, Gland, Switzerland.

Zilhão, J. (2001). Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe. *Proc. Natl. Acad. Sci. USA*, *98(*24):14180-14185.

APPENDICES

APPENDIX A

Haplogroup results of each sample according to the ND2-SSCP analysis and mtNDA CR sequences and the collection site for each sample.

Sample Name	Breed Name	ND2-Results	Latitude	Longitude	Sequence
KRY1	Karayaka	B	36.82	40.6391	B
KRY2	Karayaka	B	36.6566	40.5426	
KRY3	Karayaka	E	36.82	40.6391	E
KRY4	Karayaka	B	36.82	40.6391	
KRY5	Karayaka	B	36.82	40.6391	
KRY6	Karayaka	A	36.82	40.6391	A
KRY7	Karayaka	C/E	36.82	40.6391	C
KRY8	Karayaka	C/E	36.82	40.6391	C
KRY9	Karayaka	A	37.2129	40.5738	
KRY10	Karayaka	B	37.2129	40.5738	
KRY11	Karayaka	A	37.2129	40.5738	A
KRY12	Karayaka	B	37.2129	40.5738	
KRY13	Karayaka	B	37.2936	40.5417	
KRY14	Karayaka	A	36.6067	40.541	A
KRY15	Karayaka	B	35.9774	40.1455	
KRY16	Karayaka	B	35.7256	40.167	
KRY17	Karayaka	A	36.4705	40.3219	
KRY18	Karayaka	A	37.2129	40.5738	
KRY19	Karayaka	B	37.2129	40.5738	
KRY20	Karayaka	B	36.4705	40.3219	
KRY21	Karayaka	B	36.6566	40.5426	
KRY22	Karayaka	B	36.7333	40.5359	
KRY23	Karayaka	B	36.249	40.1685	

KRY24	Karayaka	B	36.249	40.1685	
KRY25	Karayaka	B	36.6067	40.541	B
KRY26	Karayaka	B	36.5748	40.5393	B
KRY27	Karayaka	A	36.6566	40.5426	A
KRY28	Karayaka	B	36.4705	40.3219	
KRY29	Karayaka	B	36.5955	40.7404	
KRY30	Karayaka	B	36.5955	40.7404	
KRY31	Karayaka	B	36.5955	40.7404	
KRY32	Karayaka	B	36.5955	40.7404	
KRY33	Karayaka	B	36.5955	40.7404	
KRY34	Karayaka	B	35.7256	40.167	
KRY35	Karayaka	B	36.5955	40.7404	B
KRY36	Karayaka	B	36.5955	40.7404	
KRY37	Karayaka	B	36.5955	40.7404	
KRY38	Karayaka	A	36.5955	40.7404	A
KRY39	Karayaka	B	36.5523	40.3351	
KRY40	Karayaka	B	36.7333	40.5359	B
KRY41	Karayaka	A	36.249	40.1685	
KRY42	Karayaka	B	37.4493	40.7503	
KRY43	Karayaka	B	37.4493	40.7503	
KRY44	Karayaka	C/E	37.4493	40.7503	C
KRY45	Karayaka	B	37.4493	40.7503	
KRY46	Karayaka	B	37.4493	40.7503	
KRY47	Karayaka	B	37.4493	40.7503	
KRY48	Karayaka	B	37.2936	40.5417	B
KRY49	Karayaka	B	37.2936	40.5417	
KRY50	Karayaka	B	37.2936	40.5417	
AKK1	Akkaraman	B	37.966667	32.9	B
AKK2	Akkaraman	A	37.966667	32.9	
AKK3	Akkaraman	A	37.966667	32.9	
AKK4	Akkaraman	B	37.966667	32.9	
AKK5	Akkaraman	B	37.966667	32.9	
AKK6	Akkaraman	B	37.966667	32.9	
AKK7	Akkaraman	B	37.9	32.883333	
AKK8	Akkaraman	B	37.9	32.883333	
AKK9	Akkaraman	B	37.9	32.883333	
AKK10	Akkaraman	C/E	37.9	32.883333	C
AKK11	Akkaraman	B	37.9	32.883333	
AKK12	Akkaraman	C/E	37.9	32.883333	C
AKK13	Akkaraman	E	37.9	32.883333	E
AKK14	Akkaraman	B	37.9	32.883333	B
AKK15	Akkaraman	B	37.9	32.883333	
AKK16	Akkaraman	C/E	37.9	32.883333	C
AKK17	Akkaraman	A	37.9	32.883333	

AKK18	Akkaraman	A	37.9	32.883333	
AKK19	Akkaraman	A	37.833333	32.783333	A
AKK20	Akkaraman	A	37.833333	32.783333	
AKK21	Akkaraman	B	37.833333	32.783333	
AKK22	Akkaraman	A	37.966667	32.9	
AKK23	Akkaraman	C/E	37.833333	32.783333	C
AKK24	Akkaraman	A	37.833333	32.783333	A
AKK25	Akkaraman	B	37.833333	32.783333	B
AKK26	Akkaraman	B	37.966667	32.9	B
AKK27	Akkaraman	B	37.966667	32.9	
AKK28	Akkaraman	A	37.966667	32.9	A
AKK29	Akkaraman	B	37.966667	32.9	
AKK30	Akkaraman	A	37.966667	32.9	A
AKK31	Akkaraman	B	37.966667	32.9	
AKK32	Akkaraman	C/E	37.9	32.883333	C
AKK33	Akkaraman	A	37.9	32.883333	
AKK34	Akkaraman	C/E	37.9	32.883333	C
AKK35	Akkaraman	C/E	37.9	32.883333	C
AKK36	Akkaraman	B	37.9	32.883333	
AKK37	Akkaraman	B	37.9	32.883333	
AKK38	Akkaraman	E	37.9	32.883333	E
AKK39	Akkaraman	D	37.9	32.883333	D
AKK40	Akkaraman	B	37.9	32.883333	
AKK41	Akkaraman	B	37.9	32.883333	
AKK42	Akkaraman	B	37.9	32.883333	
AKK43	Akkaraman	B	37.833333	32.783333	
AKK44	Akkaraman	B	37.833333	32.783333	
AKK45	Akkaraman	E	37.833333	32.783333	E
AKK46	Akkaraman	A	37.833333	32.783333	A
AKK47	Akkaraman	A	37.966667	32.9	A
AKK48	Akkaraman	B	37.833333	32.783333	
AKK49	Akkaraman	B	37.833333	32.783333	
AKK50	Akkaraman	B	37.833333	32.783333	
GOK1	Gökçeada	A	40.2321	25.9446	
GOK2	Gökçeada	A	40.2152	25.9128	A
GOK3	Gökçeada	A	40.2152	25.9128	
GOK4	Gökçeada	B	40.2152	25.9128	
GOK5	Gökçeada	B	40.2152	25.9128	
GOK6	Gökçeada	A	40.2152	25.9128	
GOK7	Gökçeada	B	40.2152	25.9128	
GOK8	Gökçeada	B	40.1297	25.9493	B
GOK9	Gökçeada	B	40.1297	25.9493	
GOK10	Gökçeada	B	40.2148	25.9373	
GOK11	Gökçeada	A	40.2148	25.9373	

GOK12	Gökçeada	B	40.2148	25.9373	
GOK13	Gökçeada	B	40.2148	25.9373	
GOK14	Gökçeada	A	40.2148	25.9373	
GOK15	Gökçeada	A	40.2148	25.9373	
GOK16	Gökçeada	B	40.2148	25.9373	
GOK17	Gökçeada	A	40.2148	25.9373	A
GOK18	Gökçeada	A	40.2148	25.9373	
GOK19	Gökçeada	A	40.2148	25.9373	
GOK20	Gökçeada	A	40.2148	25.9373	
GOK21	Gökçeada	A	40.2148	25.9373	
GOK22	Gökçeada	A	40.2148	25.9373	A
GOK23	Gökçeada	B	40.2148	25.9373	B
GOK24	Gökçeada	B	40.2148	25.9373	
GOK25	Gökçeada	B	40.2148	25.9373	
GOK26	Gökçeada	B	40.2305	25.9388	
GOK27	Gökçeada	A	40.2305	25.9388	A
GOK28	Gökçeada	B	40.2305	25.9388	
GOK29	Gökçeada	B	40.2305	25.9388	
GOK30	Gökçeada	B			
GOK31	Gökçeada	B	40.2152	25.9128	
GOK32	Gökçeada	B	40.2152	25.9128	
GOK33	Gökçeada	B	40.2152	25.9128	
GOK34	Gökçeada	B	40.2152	25.9128	B
GOK35	Gökçeada	B	40.2152	25.9128	
GOK36	Gökçeada	B	40.1297	25.9493	
GOK37	Gökçeada	B	40.1297	25.9493	
GOK38	Gökçeada	B	40.1297	25.9493	B
GOK39	Gökçeada	B	40.1297	25.9493	B
GOK40	Gökçeada	B	40.1297	25.9493	
GOK41	Gökçeada	A	40.1297	25.9493	A
GOK42	Gökçeada	B	40.1297	25.9493	
GOK43	Gökçeada	B	40.1297	25.9493	
GOK44	Gökçeada	B	40.1297	25.9493	
GOK45	Gökçeada	B	40.1297	25.9493	
GOK46	Gökçeada	B	40.1297	25.9493	
GOK47	Gökçeada	B	40.1297	25.9493	
GOK48	Gökçeada	B	40.1297	25.9493	B
GOK49	Gökçeada	B	40.1297	25.9493	
GOK50	Gökçeada	A	40.1297	25.9493	
DAG1	Dağlıç	C/E	38.47	31.05	C
DAG2	Dağlıç	B	38.47	31.05	
DAG3	Dağlıç	B	38.47	31.05	
DAG4	Dağlıç	B	38.46	31.06	B
DAG5	Dağlıç	B	38.46	31.06	

DAG6	Dağlıç	B	38.42	31.05	
DAG7	Dağlıç	B	38.42	31.05	
DAG8	Dağlıç	B	38.42	31.05	B
DAG9	Dağlıç	A	38.27	30.26	A
DAG10	Dağlıç	C/E	38.08	30.15	C
DAG11	Dağlıç	B	38.08	30.15	
DAG12	Dağlıç	A	38.08	30.15	A
DAG13	Dağlıç	B	38.08	30.15	
DAG14	Dağlıç	B	38.56	30.3	
DAG15	Dağlıç	B	38.59	30.3	
DAG16	Dağlıç	B	38.59	30.3	
DAG17	Dağlıç	B	38.59	30.3	
DAG18	Dağlıç	B	38.59	30.3	
DAG19	Dağlıç	B	38.51	30.31	
DAG20	Dağlıç	B	38.08	30.15	
DAG21	Dağlıç	A	38.37	29.55	A
DAG22	Dağlıç	B	38.37	29.55	B
DAG23	Dağlıç	B	38.37	29.55	
DAG24	Dağlıç	B	38.47	31.05	B
DAG25	Dağlıç	A	38.27	30.26	A
DAG31	Dağlıç	B	38.37	29.55	
DAG32	Dağlıç	B	38.37	29.55	B
DAG33	Dağlıç	B	38.37	29.55	B
DAG34	Dağlıç	B	38.47	31.05	
DAG35	Dağlıç	C/E	38.27	30.26	C
DAG36	Dağlıç	E	38.37	29.55	E
DAG37	Dağlıç	B	38.37	29.55	
DAG38	Dağlıç	B	38.37	29.55	
DAG39	Dağlıç	C/E	38.37	29.55	C
DAG40	Dağlıç	C/E	38.37	29.55	C
DAG41	Dağlıç	B	38.37	29.55	
DAG42	Dağlıç	B	38.37	29.55	
DAG43	Dağlıç	B	38.37	29.55	
DAG44	Dağlıç	C/E	38.37	29.55	C
DAG45	Dağlıç	B	38.37	29.55	
DAG46	Dağlıç	B	38.37	29.55	
DAG47	Dağlıç	B	38.08	30.15	B
DAG48	Dağlıç	B	38.08	30.15	
DAG49	Dağlıç	B	38.08	30.15	
DAG50	Dağlıç	A	38.59	30.3	A
DAG55	Dağlıç	C/E	38.56	30.3	C
DAG56	Dağlıç	A	38.56	30.3	A
DAG57	Dağlıç	A	38.27	30.26	A
DAG58	Dağlıç	B	38.47	31.05	

DAG59	Dağlıç	A	38.47	31.05	
MRK1	Morkaraman	A	41.09	39.79	A
MRK2	Morkaraman	A	41.76	39.77	
MRK3	Morkaraman	B	41.6	39.94	
MRK4	Morkaraman	A	41.8	40.07	
MRK5	Morkaraman	B	41.53	39.92	
MRK6	Morkaraman	A	41.76	39.77	A
MRK7	Morkaraman	A	41.82	40.01	
MRK8	Morkaraman	C/E	41.5	39.92	C
MRK9	Morkaraman	C/E	41.49	39.99	C
MRK10	Morkaraman	A	41.49	39.99	
MRK11	Morkaraman	B	42.39	39.7	
MRK12	Morkaraman	B	44.19	38	B
MRK13	Morkaraman	B	42.39	39.7	
MRK14	Morkaraman	E	41.67	39.89	E
MRK15	Morkaraman	B	41.8	40.07	
MRK16	Morkaraman	B	41.68	40.03	
MRK17	Morkaraman	B	41.09	40.04	B
MRK18	Morkaraman	C/E	41.67	39.89	C
MRK19	Morkaraman	A	41.68	40.03	A
MRK20	Morkaraman	B	41.68	40.03	
MRK21	Morkaraman	A	41.09	40.04	
MRK22	Morkaraman	A	41.68	40.03	
MRK23	Morkaraman	A	41.72	39.94	A
MRK24	Morkaraman	E	41.53	39.92	E
MRK25	Morkaraman	B	41.73	39.89	
MRK26	Morkaraman	A	41.8	40.07	
MRK27	Morkaraman	B	41.49	39.99	
MRK28	Morkaraman	C/E	41.9165	39.8642	C
MRK29	Morkaraman	E	41.68	40.03	E
MRK30	Morkaraman	C/E	41.82	40.01	C
MRK31	Morkaraman	C/E	41.8221	40.0145	C
MRK32	Morkaraman	A	41.68	40.03	A
MRK33	Morkaraman	B	41.4982	39.9284	B
MRK34	Morkaraman	B	41.68	40.03	
MRK35	Morkaraman	A	41.4982	39.9284	
MRK36	Morkaraman	A	41.68	40.03	
MRK37	Morkaraman	E	41.68	40.03	E
MRK38	Morkaraman	A	41.8	40.07	A
MRK39	Morkaraman	A	41.76	39.77	
MRK40	Morkaraman	C/E	42.39	39.7	
MRK41	Morkaraman	C/E	41.8221	40.0145	C
MRK42	Morkaraman	A	41.8221	40.0145	A
MRK43	Morkaraman	C/E	41.8	40.07	C

MRK44	Morkaraman	B	41.68	40.03	B
MRK45	Morkaraman	B	42.39	39.7	
MRK46	Morkaraman	C/E	41.9165	39.8642	C
MRK47	Morkaraman	A	42.39	39.7	A
MRK48	Morkaraman	C/E	41.68	40.03	C
MRK49	Morkaraman	B	41.68	40.03	
MRK50	Morkaraman	A	41.9165	39.8642	
KIV1	Kıvırcık	B	41.8377	27.4681	
KIV2	Kıvırcık	B	41.8377	27.4681	
KIV3	Kıvırcık	B	41.8377	27.4681	
KIV4	Kıvırcık	B	41.902	27.5458	B
KIV5	Kıvırcık	B	41.902	27.5458	
KIV6	Kıvırcık	B	41.8377	27.4681	
KIV7	Kıvırcık	B	41.8377	27.4681	
KIV9	Kıvırcık	B	41.8377	27.4681	B
KIV10	Kıvırcık	B	41.8377	27.4681	
KIV11	Kıvırcık	B	41.8377	27.4681	
KIV12	Kıvırcık	B	41.8377	27.4681	
KIV13	Kıvırcık	B	41.902	27.5458	
KIV14	Kıvırcık	B	41.8377	27.4681	
KIV15	Kıvırcık	C/E	41.8377	27.4681	C
KIV16	Kıvırcık	B	41.902	27.5458	
KIV17	Kıvırcık	B	41.902	27.5458	
KIV18	Kıvırcık	B	41.902	27.5458	B
KIV19	Kıvırcık	B	41.902	27.5458	
KIV20	Kıvırcık	B	41.8377	27.4681	
KIV21	Kıvırcık	B	41.8377	27.4681	
KIV22	Kıvırcık	B	41.8377	27.4681	
KIV23	Kıvırcık	B	41.8377	27.4681	
KIV24	Kıvırcık	B	41.8377	27.4681	
KIV26	Kıvırcık	B	41.8377	27.4681	
KIV27	Kıvırcık	B	41.8377	27.4681	
KIV28	Kıvırcık	B	41.8377	27.4681	
KIV29	Kıvırcık	B	41.8377	27.4681	B
KIV30	Kıvırcık	B	41.8377	27.4681	
KIV31	Kıvırcık	B	41.8377	27.4681	
KIV33	Kıvırcık	B	41.8377	27.4681	
KIV34	Kıvırcık	B	41.8377	27.4681	
KIV36	Kıvırcık	B	41.8377	27.4681	
KIV37	Kıvırcık	B	41.8377	27.4681	
KIV38	Kıvırcık	B	41.8377	27.4681	
KIV39	Kıvırcık	B	41.8377	27.4681	
KIV40	Kıvırcık	B	41.8377	27.4681	B
KIV41	Kıvırcık	B	41.8377	27.4681	

KIV42	Kıvırcık	B	41.8377	27.4681	
KIV43	Kıvırcık	C/E	41.8377	27.4681	C
KIV44	Kıvırcık	B	41.8377	27.4681	
KIV45	Kıvırcık	B	41.8377	27.4681	
KIV47	Kıvırcık	B	41.8377	27.4681	B
KIV48	Kıvırcık	B	41.8377	27.4681	
KIV50	Kıvırcık	B	41.8377	27.4681	
KIV51	Kıvırcık	B	41.8377	27.4681	X
IVE1	İvesi	A	39.14	36.99	
IVE2	İvesi	B	39.26	37.06	B
IVE3	İvesi	C/E	39.26	37.06	C
IVE4	İvesi	B	39.26	37.06	
IVE5	İvesi	B	39.26	37.06	
IVE6	İvesi	B	39.26	37.06	
IVE7	İvesi	A	39.26	37.06	A
IVE8	İvesi	B	39.26	37.06	B
IVE9	İvesi	B	39.26	37.06	B
IVE10	İvesi	B	39.26	37.06	B
IVE11	İvesi	B	38.79	37.14	
IVE12	İvesi	B	38.9	37.12	
IVE13	İvesi	A	39.14	36.99	
IVE14	İvesi	C/E	39.15	37.01	C
IVE15	İvesi	A	39.15	37.01	A
IVE16	İvesi	B	39.26	37.06	
IVE17	İvesi	A	38.93	37.1	
IVE18	İvesi	B	38.93	37.1	
IVE19	İvesi	C/E	38.79	37.14	C
IVE20	İvesi	A	38.79	37.14	
IVE21	İvesi	B	38.93	37.1	
IVE22	İvesi	A	38.79	37.14	
IVE23	İvesi	A	38.79	37.14	A
IVE24	İvesi	A	38.79	37.14	
IVE25	İvesi	C/E	39.14	36.99	C
IVE26	İvesi	B	39.14	36.99	B
IVE27	İvesi	B	39.14	36.99	
IVE28	İvesi	C/E	39.15	37.01	C
IVE29	İvesi	B	38.93	37.1	
IVE30	İvesi	A	39.26	37.06	
IVE31	İvesi	A	38.79	37.14	A
IVE32	İvesi	C/E	38.79	37.14	C
IVE33	İvesi	A	38.79	37.14	
IVE34	İvesi	A	38.79	37.14	A
IVE35	İvesi	A	38.79	37.14	
IVE36	İvesi	B	38.79	37.14	B

IVE37	İvesi	A	38.79	37.14	
IVE38	İvesi	C/E	38.79	37.14	C
IVE39	İvesi	A	38.79	37.14	
IVE41	İvesi	C/E	39.14	36.99	C
IVE42	İvesi	A	38.93	37.1	A
IVE43	İvesi	B	39.26	37.06	
IVE44	İvesi	A	39.26	37.06	A
IVE45	İvesi	C/E	39.26	37.06	C
IVE46	İvesi	B	39.26	37.06	
IVE47	İvesi	C/E	39.26	37.06	C
IVE48	İvesi	A	39.26	37.06	A
IVE49	İvesi	C/E	39.26	37.06	C
IVE50	İvesi	B	39.26	37.06	B
IVE51	İvesi	B	38.79	37.14	
IVE52	İvesi	B	39.26	37.06	
HER1	Herik	B	40.861944	35.469167	
HER2	Herik	B	40.890833	35.554722	
HER3	Herik	B	40.842778	35.613056	
HER4	Herik	A	40.861944	35.469167	A
HER5	Herik	B	40.861944	35.469167	
HER6	Herik	B	40.790833	35.465833	
HER7	Herik	B	40.8406	35.1761	B
HER8	Herik	B	40.868611	35.471944	
HER9	Herik	B	40.790833	35.465833	
HER10	Herik	B	40.704722	35.515556	
HER11	Herik	B	40.861944	35.469167	
HER12	Herik	B	40.790833	35.465833	
HER13	Herik	C/E	40.861944	35.469167	C
HER14	Herik	B	40.868611	35.471944	B
HER15	Herik	B	40.842778	35.613056	
HER16	Herik	B	40.842778	35.613056	B
HER17	Herik	B	40.842778	35.613056	
HER18	Herik	B	40.868611	35.471944	
HER19	Herik	B	40.868611	35.471944	
HER20	Herik	B	40.868611	35.471944	
HER21	Herik	B	40.790833	35.465833	
HER22	Herik	B	40.842778	35.613056	
HER23	Herik	A	40.842778	35.613056	A
HER24	Herik	A	40.790833	35.465833	X
HER25	Herik	B	40.8486	35.4812	
HER26	Herik	C/E	40.8486	35.4812	C
HER27	Herik	B	40.8749	35.4597	
HER28	Herik	B	40.8749	35.4597	
HER29	Herik	B	40.8749	35.4597	

HER30	Herik	B	40.8896	35.555	B
HER31	Herik	C/E	40.8896	35.555	C
HER32	Herik	B	40.9016	35.5201	
HER33	Herik	B	40.9016	35.5201	
HER34	Herik	B	40.8749	35.4597	
HER35	Herik	B	40.7562	35.5596	
HER36	Herik	B	40.7532	35.6702	B
HER37	Herik	B	40.7532	35.6702	
HER38	Herik	C/E	40.7532	35.6702	
HER39	Herik	C/E	40.7532	35.6702	C
HER40	Herik	B	40.7297	35.7045	B
HER41	Herik	B	40.7323	35.291	
HER42	Herik	A	40.8682	35.4691	
HER43	Herik	B	40.8682	35.4691	
HER44	Herik	B	40.8682	35.4691	
HER45	Herik	B	40.8682	35.4691	B
HER46	Herik	B	40.8682	35.4691	
HER47	Herik	B	40.7914	35.466	
HER48	Herik	B	40.7914	35.466	
HER49	Herik	A	40.7914	35.466	A
KRG1	Karagül	C/E	36.15	40.6	C
KRG2	Karagül	A	36.15	40.6	A
KRG3	Karagül	B	36.03	40.4	
KRG4	Karagül	B	36.03	40.4	B
KRG5	Karagül	B	36.39	40.29	
KRG6	Karagül	B	36.39	40.29	
KRG7	Karagül	B	36.39	40.29	
KRG8	Karagül	C/E	36.39	40.29	C
KRG9	Karagül	B	36.39	40.29	B
KRG10	Karagül	B	36.39	40.29	
KRG11	Karagül	A	36.39	40.29	
KRG12	Karagül	B	36.39	40.29	B
KRG13	Karagül	A	36.39	40.29	A
KRG14	Karagül	B	36.43	40.3	
KRG15	Karagül	A	36.43	40.3	
KRG16	Karagül	A	36.43	40.3	A
KRG17	Karagül	E	36.39	40.29	E
KRG18	Karagül	A	36.39	40.29	
KRG19	Karagül	B	36.39	40.29	
KRG20	Karagül	B	36.39	40.29	B
KRG21	Karagül	E	36.39	40.29	E
KRG22	Karagül	B	36.39	40.29	
KRG23	Karagül	A	36.39	40.29	A
KRG24	Karagül	A	36.39	40.29	A

KRG25	Karagül	B	36.39	40.29	
KRG31	Karagül	A	36.75	40.46	
KRG32	Karagül	B	36.75	40.46	B
KRG33	Karagül	B	36.39	40.29	
KRG34	Karagül	B	36.216667	40.335	
KRG35	Karagül	A	36.216667	40.335	
KRG36	Karagül	B	36.216667	40.335	
KRG37	Karagül	B	36.216667	40.335	
KRG38	Karagül	B	36.216667	40.335	B
KRG39	Karagül	B	36.393611	40.463889	
KRG40	Karagül	B	36.393611	40.463889	
KRG41	Karagül	B	36.393611	40.463889	B
KRG42	Karagül	B	36.393611	40.463889	
KRG43	Karagül	B	36.393611	40.463889	
KRG44	Karagül	B	36.393611	40.463889	B
KRG45	Karagül	B	36.433056	40.303333	
KRG46	Karagül	B	36.433056	40.303333	
KRG47	Karagül	A	36.433056	40.303333	
KRG48	Karagül	B	36.433056	40.303333	
KRG49	Karagül	B	36.39	40.29	
KRG50	Karagül	B	36.39	40.29	
KRG51	Karagül	C/E	36.39	40.29	A
KRG52	Karagül	B	36.43	40.3	
KRG53	Karagül	A	36.43	40.3	A
KRG54	Karagül	A	36.43	40.3	A
KRG55	Karagül	B	36.43	40.3	
HEM1	Hemşin	B	41.1103	42.0721	
HEM2	Hemşin	B	42.2116	41.0406	B
HEM3	Hemşin	A	41.1103	42.0721	A
HEM4	Hemşin	B			
HEM5	Hemşin	B	41.9804	41.0741	B
HEM6	Hemşin	A	42.0617	41.1431	
HEM7	Hemşin	C/E	42.1512	41.1189	C
HEM8	Hemşin	B	41.6168	41.4407	
HEM9	Hemşin	B	42.2116	41.0406	
HEM10	Hemşin	A	41.1103	42.0721	A
HEM11	Hemşin	B	42.1512	41.0374	
HEM12	Hemşin	B	41.1103	42.0721	
HEM13	Hemşin	A	41.1135	42.1531	
HEM14	Hemşin	A	42.1603	41.0444	A
HEM15	Hemşin	A	42.2135	41.0392	
HEM16	Hemşin	A	42.0652	41.1605	
HEM17	Hemşin	B	42.2116	41.0406	B
HEM18	Hemşin	B	41.9804	41.0741	

HEM19	Hemşin	C/E	42.0652	41.1605	
HEM20	Hemşin	B	42.3169	41.2994	
HEM21	Hemşin	C/E	42.2116	41.0406	C
HEM22	Hemşin	A	42.0652	41.1605	
HEM23	Hemşin	A	42.3169	41.2994	A
HEM24	Hemşin	A	42.2116	41.0406	A
HEM25	Hemşin	A	41.1103	42.0721	A
HEM26	Hemşin	B	42.3169	41.2994	B
HEM27	Hemşin	B	42.2116	41.0406	
HEM28	Hemşin	B	42.0944	41.0445	
HEM29	Hemşin	B	42.2135	41.0392	
HEM30	Hemşin	B	42.3169	41.2994	
HEM31	Hemşin	B	42.2135	41.0392	
HEM32	Hemşin	C/E	41.9804	41.0741	C
HEM33	Hemşin	B	42.0652	41.1605	
HEM34	Hemşin	A	42.1603	41.0444	
HEM35	Hemşin	B	42.0652	41.1605	
HEM36	Hemşin	A	42.2116	41.0406	
HEM37	Hemşin	B	42.1512	41.1189	
HEM38	Hemşin	A	41.1103	42.0721	
HEM39	Hemşin	A	42.1513	41.0367	A
HEM40	Hemşin	B	42.2116	41.0406	
HEM41	Hemşin	A	41.1135	42.1531	A
HEM42	Hemşin	B	42.0617	41.1431	B
HEM43	Hemşin	C/E	41.1135	42.1531	C
HEM44	Hemşin	C/E	42.0652	41.1605	C
HEM45	Hemşin	B	41.6168	41.4407	
HEM46	Hemşin	B	41.6168	41.4407	
HEM47	Hemşin	C/E	42.0827	41.1114	
HEM48	Hemşin	B	41.9804	41.0741	
CIC1	Çine Çaparı	A	27.69	37.75	A
CIC2	Çine Çaparı	B	27.69	37.75	
CIC3	Çine Çaparı	B	27.69	37.75	
CIC4	Çine Çaparı	B	27.69	37.75	
CIC5	Çine Çaparı	B	27.69	37.75	
CIC6	Çine Çaparı	A	27.72	37.59	A
CIC7	Çine Çaparı	B	27.72	37.59	B
CIC8	Çine Çaparı	B	27.72	37.59	
CIC9	Çine Çaparı	B	27.72	37.59	
CIC10	Çine Çaparı	A	27.72	37.59	A
CIC11	Çine Çaparı	B	27.69	37.75	
CIC12	Çine Çaparı	B	28.09	37.54	
CIC13	Çine Çaparı	B	28.09	37.54	
CIC14	Çine Çaparı	B	28.09	37.54	

CIC15	Çine Çaparı	B	28.09	37.54	
CIC16	Çine Çaparı	C/E	28.09	37.54	C
CIC17	Çine Çaparı	B	28.09	37.54	
CIC18	Çine Çaparı	B	28.09	37.54	B
CIC19	Çine Çaparı	B	28.09	37.54	
CIC20	Çine Çaparı	B	28.09	37.54	
CIC21	Çine Çaparı	B	28.09	37.54	
CIC22	Çine Çaparı	B	28.09	37.54	B
CIC23	Çine Çaparı	C/E	28.09	37.54	C
CIC24	Çine Çaparı	C/E	28.09	37.54	C
CIC25	Çine Çaparı	B	27.69	37.75	
CIC27	Çine Çaparı	A	27.69	37.75	A
CIC28	Çine Çaparı	B	27.69	37.75	
CIC29	Çine Çaparı	A			A
CIC30	Çine Çaparı	B	27.69	37.75	B
CIC31	Çine Çaparı	B	27.69	37.75	B
CIC32	Çine Çaparı	B	27.69	37.75	
CIC33	Çine Çaparı	B	27.69	37.75	
CIC34	Çine Çaparı	C/E	27.69	37.75	C
CIC35	Çine Çaparı	B	27.69	37.75	B
CIC36	Çine Çaparı	C/E			C
CIC37	Çine Çaparı	B	27.69	37.75	
CIC38	Çine Çaparı	B	27.69	37.75	
CIC39	Çine Çaparı	B	27.69	37.75	
CIC40	Çine Çaparı	B	27.69	37.75	
CIC41	Çine Çaparı	B	27.69	37.75	
SAK1	Sakız	B	38.340833	26.290833	
SAK2	Sakız	B	38.340833	26.290833	
SAK3	Sakız	B	38.340833	26.290833	
SAK4	Sakız	B	38.340833	26.290833	
SAK5	Sakız	B	38.340833	26.290833	B
SAK6	Sakız	B	38.291389	26.333056	
SAK7	Sakız	B	38.291389	26.333056	B
SAK8	Sakız	B	38.291389	26.333056	B
SAK9	Sakız	B	38.291389	26.333056	
SAK10	Sakız	B	38.291389	26.333056	
SAK11	Sakız	B	38.3	26.45	
SAK12	Sakız	B	38.3	26.45	
SAK13	Sakız	B	38.3	26.45	
SAK14	Sakız	B	38.3	26.45	
SAK15	Sakız	B	38.3	26.45	
SAK16	Sakız	B	38.3	26.45	B
SAK17	Sakız	C/E	38.3	26.416667	C
SAK18	Sakız	C/E	38.3	26.416667	C

SAK19	Sakız	B	38.25	26.3	
SAK20	Sakız	B	38.283333	26.283333	
SAK21	Sakız	B	38.283333	26.283333	B
SAK22	Sakız	B	38.283333	26.283333	
SAK23	Sakız	B	38.283333	26.283333	
SAK24	Sakız	A	38.283333	26.283333	Aq
SAK25	Sakız	B	38.283333	26.283333	
SAK26	Sakız	B	38.340833	26.290833	
SAK27	Sakız	B	38.340833	26.290833	
SAK28	Sakız	B	38.340833	26.290833	
SAK29	Sakız	B	38.340833	26.290833	B
SAK30	Sakız	B	38.340833	26.290833	
SAK32	Sakız	B	38.291389	26.333056	B
SAK33	Sakız	B	38.291389	26.333056	
SAK34	Sakız	B	38.291389	26.333056	
SAK35	Sakız	B	38.291389	26.333056	B
SAK36	Sakız	B	38.3	26.45	
SAK37	Sakız	A	38.3	26.45	A
SAK38	Sakız	B	38.3	26.45	
SAK39	Sakız	B	38.3	26.45	B
SAK40	Sakız	B	38.3	26.45	
SAK41	Sakız	B	38.3	26.45	
SAK42	Sakız	B	38.3	26.45	B
SAK43	Sakız	B	38.3	26.45	B
SAK44	Sakız	B	38.3	26.45	
SAK45	Sakız	B	38.3	26.45	B
SAK46	Sakız	B	38.3	26.45	
SAK47	Sakız	C/E	38.3	26.416667	C
SAK48	Sakız	C/E	38.316667	26.4	C
SAK49	Sakız	B	38.25	26.3	B
SAK50	Sakız	B	38.25	26.3	
NOR1	Norduz	A	43.37	38.03	
NOR2	Norduz	B	43.37	38.03	B
NOR3	Norduz	A	43.44	37.87	
NOR4	Norduz	A	43.44	37.87	A
NOR5	Norduz	A	43.44	37.87	
NOR6	Norduz	B	43.44	37.87	
NOR7	Norduz	B	43.44	37.87	
NOR8	Norduz	C/E	43.44	37.87	C
NOR9	Norduz	B	43.44	37.87	
NOR10	Norduz	B	43.44	37.87	
NOR11	Norduz	B	43.44	37.87	
NOR12	Norduz	A	43.37	38.03	A
NOR13	Norduz	B	43.37	38.03	B

NOR14	Norduz	B	43.44	37.87	
NOR15	Norduz	B	43.44	37.87	
NOR16	Norduz	C/E	43.44	37.87	C
NOR17	Norduz	C/E	43.44	37.87	C
NOR18	Norduz	C/E	43.37	38.03	C
NOR19	Norduz	B	44.01	38.04	
NOR20	Norduz	A	44.01	38.04	
NOR21	Norduz	A	44.01	38.04	A
NOR22	Norduz	C/E	44.01	38.04	C
NOR23	Norduz	B	44.01	38.04	B
NOR24	Norduz	C/E	44.01	38.04	C
NOR25	Norduz	B	44.01	38.04	
NOR26	Norduz	C/E	43.41	38.32	C
NOR27	Norduz	C/E	43.41	38.32	C
NOR28	Norduz	A	43.41	38.32	A
NOR29	Norduz	B	43.41	38.32	B
NOR30	Norduz	B	43.41	38.32	
NOR31	Norduz	B	43.41	38.32	
NOR32	Norduz	A	43.41	38.32	
NOR33	Norduz	B	43.41	38.32	
NOR34	Norduz	B	43.41	38.32	
NOR35	Norduz	C/E	43.41	38.32	C
NOR36	Norduz	B	43.41	38.32	
NOR37	Norduz	B	43.41	38.32	
NOR38	Norduz	A	43.41	38.32	
NOR39	Norduz	A	43.41	38.32	
NOR40	Norduz	A	43.41	38.32	A
NOR41	Norduz	B	43.41	38.32	
NOR42	Norduz	B	44.01	38.04	B
NOR43	Norduz	C/E	43.37	38.03	C
NOR44	Norduz	A	43.41	38.32	
NOR45	Norduz	D	43.41	38.32	D
NOR46	Norduz	C/E	43.41	38.32	C

APPENDIX B

The accession numbers of mtDNA CR and *cytB* sequences with their respective haplotypes used in MJ network in Figure 3.16-3.18.

ctyB	**CR**	**Sample Name**	**Haplotype**
FJ218148	EF056465	Sn_3648	H_1
FJ218146	EF056462	Sn_3628	H_2
FJ218143	EF056459	Sn_3621	H_3
FJ218139	EF056457	Sn_3612	H_4
FJ218128	EF056453	Sn_2932	H_5
FJ218126	EF056450	Sn_2870	H_6
FJ218091	EF056447	Lo_696	H_7
FJ218089	EF056446	Lo_656	H_8
FJ218086	EF056443	Lo_354	H_9
FJ218085	EF056442	Lo_349	H_10
FJ218084	EF056440	Lo_326	H_11
FJ218083	EF056437	Lo_308	H_12
FJ218094	EF056434	Lo_27811	H_13
FJ218092	EF056433	Lo_27805	H_14
FJ218058	EF056431	Ga_G9	H_5
FJ218049	EF056428	Ga_689	H_15
FJ218041	EF056427	Ga_54	H_16
FJ218040	EF056425	Ga_40	H_17
FJ218057	EF056424	Ga_34743	H_18
FJ218054	EF056423	Ga_16290	H_19
FJ218053	EF056421	Ga_16283	H_20
FJ218051	EF056419	Ga_16251	H_21
FJ218050	EF056418	Ga_16222	H_22
FJ218039	EF056417	Ga_16	H_23
FJ218043	EF056415	Ga_142	H_24
FJ218038	EF056414	Ga_1	H_25
FJ218036	EF056412	Bn_D98036	H_26
FJ218034	EF056410	Bn_566	H_27
FJ218033	EF056408	Bn_552	H_28
FJ218032	EF056407	Bn_544	H_27
FJ218029	EF056405	Bn_527	H_5
FJ218028	EF056404	Bn_512	H_29
FJ218027	EF056403	Bn_509	H_27

FJ218026	EF056402	Bn_502	H_30
FJ218035	EF056400	Bn_34009	H_31
FJ218023	EF056397	Bn_260	H_32
DQ852079	DQ852278	AW24	H_33
DQ852078	DQ852277	AW12	H_33
DQ852076	DQ852273	AW62	H_34
DQ852075	DQ852272	AW58	H_35
DQ852074	DQ852271	AW51	H_35
DQ852073	DQ852270	AW41	H_35
DQ852072	DQ852269	AW31	H_34
DQ852070	DQ852267	AW4	H_35
DQ852069	DQ852266	AW3	H_35
DQ852067	DQ852264	cc226	H_36
DQ852066	DQ852263	cc202	H_36
DQ852065	DQ852262	cc51	H_35
DQ852064	DQ852261	cc21	H_35
DQ852063	DQ852260	nz14	H_37
DQ852062	DQ852259	nz13	H_38
DQ852061	DQ852258	nz8	H_38
DQ852060	DQ852257	nz5	H_39
DQ852059	DQ852256	tj16	H_40
DQ852057	DQ852254	tj10	H_40
DQ852056	DQ852253	tj2	H_41
DQ852055	DQ852252	tj1	H_41
DQ852053	DQ852250	kk19	H_35
DQ852052	DQ852249	kk13	H_35
DQ852050	DQ852247	kk3	H_40
DQ852049	DQ852246	AW63	H_42
DQ852048	DQ852245	AW55	H_43
DQ852047	DQ852244	AW53	H_43
DQ852046	DQ852243	AW52	H_44
DQ852045	DQ852242	AW50	H_42
DQ852044	DQ852241	AW49	H_45
DQ852043	DQ852240	AW47	H_43
DQ852042	DQ852239	AW46	H_45
DQ852041	DQ852238	AW45	H_44
DQ852040	DQ852237	AW42	H_42
DQ852039	DQ852236	AW40	H_43
DQ852038	DQ852235	AW39	H_46
DQ852037	DQ852234	AW38	H_47
DQ852036	DQ852233	AW37	H_48
DQ852035	DQ852232	AW32	H_42
DQ852034	DQ852231	AW29	H_49
DQ852033	DQ852230	AW28	H_50

DQ852031	DQ852228	AW22	H_51
DQ852030	DQ852227	AW20	H_45
DQ852029	DQ852226	AW19	H_52
DQ852027	DQ852224	AW17	H_53
DQ852026	DQ852223	AW15	H_54
DQ852025	DQ852222	AW13	H_48
DQ852024	DQ852221	AW11	H_49
DQ852023	DQ852220	AW10	H_54
DQ852022	DQ852219	AW9	H_42
DQ852021	DQ852218	AW7	H_46
DQ852020	DQ852217	AW5	H_42
DQ852019	DQ852216	AW1	H_42
DQ852018	DQ852215	sz36	H_55
DQ852017	DQ852214	sz35	H_56
DQ852016	DQ852213	sz34	H_56
DQ852015	DQ852212	sz31	H_56
DQ852014	DQ852211	sz27	H_56
DQ852013	DQ852210	sz24	H_56
DQ852012	DQ852209	sz22	H_56
DQ852011	DQ852208	sz21	H_56
DQ852010	DQ852207	sz20	H_56
DQ852009	DQ852206	sz17	H_55
DQ852008	DQ852205	sz16	H_56
DQ852007	DQ852204	sz13	H_56
DQ852006	DQ852203	sz9	H_56
DQ852005	DQ852202	sz8	H_56
DQ852004	DQ852201	sz6	H_56
DQ852003	DQ852200	sz5	H_56
DQ852002	DQ852199	sz4	H_56
DQ852001	DQ852198	ky16	H_57
DQ852000	DQ852197	ky13	H_58
DQ851999	DQ852196	ky12	H_42
DQ851998	DQ852195	ky8	H_59
DQ851997	DQ852194	ky7	H_59
DQ851996	DQ852193	ky6	H_58
DQ851995	DQ852192	ky5	H_42
DQ851994	DQ852191	ky3	H_60
DQ851993	DQ852190	ky2	H_56
DQ851992	DQ852189	cc301	H_61
DQ851991	DQ852188	cc300	H_42
DQ851990	DQ852187	cc224	H_62
DQ851989	DQ852186	cc64	H_42
DQ851988	DQ852185	cc58	H_63
DQ851987	DQ852184	cc50	H_64

DQ851986	DQ852183	cc31	H_65
DQ851985	DQ852182	nz19	H_42
DQ851984	DQ852181	nz16	H_42
DQ851983	DQ852180	nz7	H_42
DQ851982	DQ852179	nz3	H_66
DQ851981	DQ852178	kr23	H_67
DQ851980	DQ852177	kr22	H_68
DQ851979	DQ852176	kr21	H_69
DQ851978	DQ852175	kr19	H_42
DQ851977	DQ852174	kr18	H_68
DQ851976	DQ852173	kr17	H_68
DQ851975	DQ852172	kr16	H_69
DQ851974	DQ852171	kr15	H_42
DQ851973	DQ852170	kr14	H_67
DQ851972	DQ852169	kr12	H_67
DQ851971	DQ852168	kr11	H_70
DQ851970	DQ852167	kr10	H_68
DQ851969	DQ852166	kr9	H_71
DQ851968	DQ852165	kr8	H_68
DQ851967	DQ852164	kr7	H_72
DQ851966	DQ852163	kr6	H_73
DQ851965	DQ852162	kr5	H_74
DQ851964	DQ852161	kr4	H_42
DQ851963	DQ852160	kr3	H_75
DQ851962	DQ852159	kr2	H_42
DQ851961	DQ852158	kr1	H_56
DQ851960	DQ852157	tj14	H_76
DQ851958	DQ852155	tj11	H_77
DQ851957	DQ852154	tj8	H_78
DQ851956	DQ852153	tj7	H_49
DQ851955	DQ852152	tj4	H_49
DQ851954	DQ852151	tj3	H_79
DQ851953	DQ852150	mk20	H_80
DQ851952	DQ852149	mk17	H_81
DQ851951	DQ852148	mk16	H_42
DQ851950	DQ852147	mk15	H_81
DQ851949	DQ852146	mk12	H_81
DQ851948	DQ852145	mk10	H_80
DQ851947	DQ852144	mk8	H_82
DQ851946	DQ852143	mk7	H_81
DQ851945	DQ852142	mk6	H_83
DQ851943	DQ852140	mk1	H_81
DQ851942	DQ852139	kk20	H_42
DQ851941	DQ852138	kk17	H_84

DQ851940	DQ852137	kk15	H_85
DQ851939	DQ852136	kk14	H_84
DQ851938	DQ852135	kk11	H_42
DQ851937	DQ852134	kk10	h_86
DQ851936	DQ852133	kk9	H_84
DQ851935	DQ852132	kk8	H_84
DQ851934	DQ852131	kk5	H_42
DQ851933	DQ852130	kk4	H_42
DQ851930	DQ852127	AW60	H_5
DQ851929	DQ852126	AW59	H_87
DQ851926	DQ852123	AW44	H_88
DQ851925	DQ852122	AW43	H_87
DQ851923	DQ852120	AW35	H_89
DQ851922	DQ852119	AW33	H_90
DQ851921	DQ852118	AW27	H_91
DQ851920	DQ852117	AW26	H_5
DQ851919	DQ852116	AW21	H_87
DQ851918	DQ852115	AW8	H_5
DQ851917	DQ852114	AW6	H_5
DQ851916	DQ852113	AW2	H_87
DQ851915	DQ852112	ky14	H_92
DQ851914	DQ852111	ky11	H_93
DQ851913	DQ852110	ky10	H_92
DQ851912	DQ852109	ky9	H_92
DQ851911	DQ852108	ky4	H_94
DQ851910	DQ852107	ky1	H_95
DQ851909	DQ852106	cc304	H_96
DQ851908	DQ852105	cc53	H_97
DQ851907	DQ852104	nz20	H_5
DQ851904	DQ852101	nz9	H_5
DQ851903	DQ852100	nz4	H_98
DQ851902	DQ852099	nz2	H_98
DQ851901	DQ852098	nz1	H_5
DQ851900	DQ852097	kr24	H_95
DQ851899	DQ852096	kr20	H_95
DQ851898	DQ852095	kr13	H_95
DQ851897	DQ852094	tj13	H_99
DQ851896	DQ852093	tj9	H_5
DQ851895	DQ852092	tj5	H_100
DQ851894	DQ852091	mk19	H_100
DQ851893	DQ852090	mk14	H_23
DQ851892	DQ852089	mk13	H_100
DQ851891	DQ852088	mk11	H_100
DQ851890	DQ852087	mk5	H_23

DQ851886	DQ852083	kk6	H_5
AY879584	AY879463	mongolian1	H_101
AY879583	AY879462	zd19	H_5
AY879582	AY879461	zd15	H_102
AY879581	AY879460	zd11	H_92
AY879580	AY879459	zd04	H_103
AY879579	AY879458	ZB09	H_104
AY879578	AY879457	TSS36	H_105
AY879577	AY879456	TMS9	H_106
AY879576	AY879455	TMS8	H_106
AY879575	AY879454	TMS7	H_107
AY879574	AY879453	TMS3	H_106
AY879573	AY879452	TMS18	H_108
AY879572	AY879451	TMS14	H_106
AY879570	AY879449	L8991	H_5
AY879569	AY879448	L125	H_5
AY879568	AY879447	JTT92090	H_109
AY879567	AY879446	JTT90079	H_5
AY879566	AY879445	JTT90034	H_5
AY879565	AY879444	JTT89171	H_5
AY879564	AY879443	FS68	H_5
AY879563	AY879442	FS62	H_5
AY879561	AY879440	AH25	H_110
AY879556	AY879435	LX187	H_42
AY879555	AY879434	LX106	H_111
AY879554	AY879433	zb13	H_112
AY879553	AY879432	TSS40	H_113
AY879552	AY879431	TSS38	H_114
AY879551	AY879430	TSS37	H_115
AY879550	AY879429	TSS32	H_116
AY879549	AY879428	TSS31	H_42
AY879548	AY879427	TSS29	H_113
AY879547	AY879426	TSS28	H_117
AY879546	AY879425	TSS27	H_113
AY879545	AY879424	TSS24	H_113
AY879544	AY879423	TSS23	H_114
AY879543	AY879422	TSS22	H_114
AY879542	AY879421	TSS21	H_68
AY879541	AY879420	TMS6	H_118
AY879539	AY879418	TMS4	H_119
AY879538	AY879417	TMS20	H_120
AY879536	AY879415	TMS19	H_42
AY879535	AY879414	TMS17	H_121
AY879534	AY879413	TMS15	H_119

AY879529	AY879408	O1	H_122
AY879528	AY879407	O15	H_56
AY879527	AY879406	LAK1	H_123
AY879526	AY879405	LAF2	H_124
AY879525	AY879404	L98	H_125
AY879524	AY879403	L95	H_126
AY879523	AY879402	L93	H_127
AY879522	AY879401	L44	H_128
AY879521	AY879400	L449	H_129
AY879520	AY879399	L35	H_130
AY879519	AY879398	L33	H_131
AY879518	AY879397	L2691	H_132
AY879517	AY879396	L199	H_133
AY879516	AY879395	L128	H_134
AY879515	AY879394	L127	H_42
AY879514	AY879393	L112	H_42
AY879513	AY879392	L110	H_135
AY879512	AY879391	L109	H_131
AY879511	AY879390	L108	H_131
AY879510	AY879389	L107	H_131
AY879509	AY879388	L101	H_131
AY879508	AY879387	JTT92127	H_136
AY879507	AY879386	JTT92123	H_136
AY879506	AY879385	JTT92068	H_136
AY879505	AY879384	JTT92041	H_137
AY879504	AY879383	JTT91113	H_42
AY879503	AY879382	JTT91021	H_136
AY879502	AY879381	JTT91020	H_136
AY879501	AY879380	JTT91014	H_136
AY879500	AY879379	JTT90125	H_136
AY879499	AY879378	JTT90057	H_136
AY879498	AY879377	JTT89424	H_136
AY879497	AY879376	JTT89344	H_136
AY879496	AY879375	JTT89205	H_136
AY879495	AY879374	JTT87346	H_136
AY879494	AY879373	FS73	H_138
AY879493	AY879372	FS72	H_139
AY879492	AY879371	FS71	H_140
AY879491	AY879370	FS70	H_140
AY879490	AY879369	FS69	H_140
AY879489	AY879368	FS67	H_141
AY879488	AY879367	FS66	H_140
AY879487	AY879366	FS64	H_139
AY879486	AY879365	FS63	H_142

AY879485	AY879364	FS61	H_140
AY879484	AY879363	CS60	H_143
AY879483	AY879362	CS59	H_143
AY879482	AY879361	CS58	H_143
AY879481	AY879360	CS57	H_144
AY879480	AY879359	CS56	H_145
AY879479	AY879358	CS54	H_145
AY879478	AY879357	CS53	H_145
AY879477	AY879356	CS51	H_143
AY879476	AY879355	CS50	H_143
AY879475	AY879354	CS49	H_143
AY879474	AY879353	CS48	H_143
AY879473	AY879352	CS47	H_143
AY879472	AY879351	CS46	H_143
AY879471	AY879350	CS44	H_143
AY879470	AY879349	CS43	H_146
AY879469	AY879348	CS42	H_147
AY879468	AY879347	CS41	H_145
AY879467	AY879346	AH8	H_133
AY879466	AY879345	AH18	H_148
AY879465	AY879344	AH11	H_148
AY879464	AY879343	CS52	H_149
HM236174	HM236174	cl122	H_5
HM236175	HM236175	r359	H_5
HM236176	HM236176	kk1	H_42
HM236177	HM236177	kk2	H_42
HM236178	HM236178	kk12	H_35
HM236179	HM236179	mk4	H_35
HM236180	HM236180	mk3	H_150
HM236181	HM236181	mk9	H_150
HM236182	HM236182	aw25	H_33
HM236183	HM236183	tj6	H_151
HM236184	HM236184	h1	H_56
HM236185	HM236185	h2	H_56
-	-	DAG_21	H_152
-	-	HER_24	H_5
-	-	KRG_17	H_153
-	-	CIC_23	H_35
-	-	IVE_14	H_40
-	-	AKK_39	H_154
-	-	NOR_45	H_155
-	-	OGA_14	H_156
-	-	OGA_21	H_156
-	-	DAG_36	H_33

-	-	HEM_17	H_157
-	-	OGA_9	H_5
-	-	OGA_18	H_5
-	-	NOR_18	H_40
HM236187	HM236187	h76	H_158
NC_001941	NC_001941	NC_001941	H_159
		copper age sheep	H_160

APPENDIX C

The accession numbers of mtDNA *cytB* sequences with their respective haplotypes used in MJ network in Figure 3.19.

Accession Number	Sample Name	Haplotype	Species
FJ936211	OOS1	H_1	Ovis orientalis
FJ936210	OOL2	H_2	Ovis orientalis laristanica
FJ936209	OOL1	H_3	Ovis orientalis laristanica
FJ936208	OOKo2	H_4	Ovis orientalis
FJ936207	OOKo1	H_5	Ovis orientalis
FJ936206	OOI6	H_6	Ovis orientalis isphahanica
FJ936205	OOI5	H_6	Ovis orientalis isphahanica
FJ936204	OOI3	H_7	Ovis orientalis isphahanica
FJ936203	OOG9	H_8	Ovis orientalis gmelini
FJ936202	OOG8	H_2	Ovis orientalis gmelini
FJ936201	OOG5	H_9	Ovis orientalis gmelini
FJ936200	OOG4	H_10	Ovis orientalis gmelini
FJ936199	OOG33	H_7	Ovis orientalis gmelini
FJ936198	OOG32	H_7	Ovis orientalis gmelini
FJ936195	OOG27	H_11	Ovis orientalis gmelini
FJ936194	OOG26	H_7	Ovis orientalis gmelini
FJ936193	OOG25	H_3	Ovis orientalis gmelini
FJ936192	OOG23	H_2	Ovis orientalis gmelini
FJ936191	OOG22	H_2	Ovis orientalis gmelini
FJ936190	OOG2	H_2	Ovis orientalis gmelini
FJ936189	OOG15	H_3	Ovis orientalis gmelini
FJ936188	OOG13	H_7	Ovis orientalis gmelini
FJ936187	OOG12	H_11	Ovis orientalis gmelini
FJ936186	OOG11	H_3	Ovis orientalis gmelini
FJ936185	OOA3	H_12	Ovis orientalis anatolica
EU366068	OogKh3	H_13	Ovis orientalis gmelini
FJ936197	OOG31	H_3	Ovis orientalis gmelini
EU366040	OogAr1	H_14	Ovis orientalis gmelini
EU366055	OoiAz3	H_15	Ovis orientalis isphahanica
EU366053	OogGa2	H_16	Ovis orientalis gmelini
EU366016	OgiAz1	H_17	Ovis orientalis isphahanica
EU366015	OogMa2	H_18	Ovis orientalis gmelini

EU366010	Ooils1	H_6	Ovis orientalis isphahanica
EU366009	OogKh2	H_19	Ovis orientalis gmelini
EU366004	OogMa1	H_3	Ovis orientalis gmelini
EU366003	OogGa1	H_20	Ovis orientalis gmelini
EU366001	OogMk7	H_2	Ovis orientalis gmelini
EU366000	OogMk6	H_21	Ovis orientalis gmelini
EU365998	OogMk4	H_21	Ovis orientalis gmelini
EU365997	OogBi2	H_11	Ovis orientalis gmelini
EU365996	OogBi1	H_8	Ovis orientalis gmelini
EU365991	OogKh1	H_22	Ovis orientalis gmelini
EU365990	OomFr2	H_23	Ovis aries musimon
EU365989	OogMk3	H_24	Ovis orientalis gmelini
EU365988	OogMk2	H_2	Ovis orientalis gmelini
EU365987	OoaTk4	H_25	Ovis orientalis anatolica
EU365986	OoaTk2	H_12	Ovis orientalis anatolica
EU365982	OolBa3	H_26	Ovis orientalis laristanica
EU365980	OogSn2	H_27	Ovis orientalis gmelini
EU365979	OogSn1	H_3	Ovis orientalis gmelini
EU365978	OolBa1	H_28	Ovis orientalis laristanica
EU365976	OoiAz2	H_7	Ovis orientalis isphahanica
EU365975	OogMk1	H_29	Ovis orientalis gmelini
EU365974	OoaTk3	H_7	Ovis orientalis anatolica
EU365973	OoaTk1	H_7	Ovis orientalis anatolica
EU365999	OogMk5	H_2	Ovis orientalis gmelini
EU366002	OogZa1	H_30	Ovis orientalis gmelini
HM236174	cl122	H_7	Ovis aries
HM236175	r359	H_7	Ovis aries
HM236176	kk1	H_29	Ovis aries
HM236177	kk2	H_29	Ovis aries
HM236178	kk12	H_31	Ovis aries
HM236179	mk4	H_31	Ovis aries
HM236180	mk3	H_32	Ovis aries
HM236181	mk9	H_32	Ovis aries
HM236182	aw25	H_2	Ovis aries
HM236183	tj6	H_33	Ovis aries
HM236184	h1	H_29	European musimon
HM236185	h2	H_29	European musimon
-	Dag 21	H_34	Ovis aries
-	Her 24	H_7	Ovis aries
-	Krg 17	H_29	Ovis aries
-	Cic 23	H_31	Ovis aries
-	Ive 14	H_4	Ovis aries
-	Akk 39	H_32	Ovis aries
-	Nor 45	H_35	Ovis aries

-	Oga 14	H_12	Ovis gmelinii anatolica
-	Oga 21	H_12	Ovis gmelinii anatolica
-	Dag 36	H_2	Ovis aries
-	Hem 17	H_29	Ovis aries
-	Oga 9	H_7	Ovis gmelinii anatolica
-	Oga 18	H_7	Ovis gmelinii anatolica
-	Nor 18	H_4	Ovis aries
EU365977	OomFr1	H_36	Ovis aries musimon
FJ936235	OxV8	H_37	Ovis orientalis x vignei
FJ936234	OxV7	H_37	Ovis orientalis x vignei
FJ936233	OXV6	H_28	Ovis orientalis x vignei
FJ936232	OxV37	H_38	Ovis orientalis x vignei
FJ936231	OxV36	H_38	Ovis orientalis x vignei
FJ936230	OxV35	H_39	Ovis orientalis x vignei
FJ936229	OxV34	H_40	Ovis orientalis x vignei
FJ936228	OxV33	H_19	Ovis orientalis x vignei
FJ936227	OxV32	H_41	Ovis orientalis x vignei
FJ936226	OxV25	H_42	Ovis orientalis x vignei
FJ936225	OxV24	H_41	Ovis orientalis x vignei
FJ936224	OxV14	H_40	Ovis orientalis x vignei
FJ936223	OxV13	H_19	Ovis orientalis x vignei
FJ936222	OxV12	H_39	Ovis orientalis x vignei
FJ936212	OVA12	H_43	Ovis vignei arkal
FJ936220	OVV3	H_44	Ovis vignei vignei

METU LIBRARY

TEZ FOTOKOPİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü ☐

Sosyal Bilimler Enstitüsü ☐

Uygulamalı Matematik Enstitüsü ☐

Enformatik Enstitüsü ■

Deniz Bilimleri Enstitüsü ☐

YAZARIN

Soyadı :Demirci...
Adı : Sevgin...
Bölümü : Bioinformatics...

TEZİN ADI (İngilizce) : MTDNA BASED GENETIC DIVERSITY OF NATIVE SHEEP BREEDS AND ANATOLIAN MOUFLON (*OVIS GMELINII ANATOLICA*) IN TURKEY...

TEZİN TÜRÜ : Yüksek Lisans ■ Doktora ☐

1. Tezimin tamamı dünya çapında erişime açılsın ve kaynak gösterilmek şartıyla tezimin bir kısmı veya tamamının fotokopisi alınsın. ☐
2. Tezimin tamamı yalnızca Orta Doğu Teknik Üniversitesi kullancılarının erişimine açılsın. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.) ☐
3. Tezim bir (1) yıl süreyle erişime kapalı olsun. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.) ■

Yazarın imzası .... Tarih ....