MITOCHONDRIAL DNA (MTDNA) HAPLOGROUP COMPOSITION IN TURKISH SHEEP BREEDS

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

BY

EREN YÜNCÜ

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

JANUARY 2009

Approval of the thesis:

MITOCHONDRIAL DNA (MTDNA) HAPLOGROUP COMPOSITION IN TURKISH SHEEP BREEDS

Submitted by **EREN YÜNCÜ** in partial fulfillment of the requirements for the degree of **Master of Science in Department of Biological Sciences, Middle East Technical University** by,

Prof. Dr. Canan Özgen ________________
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Zeki Kaya ________________
Head of Department, **Department of Biological Sciences**

Prof. Dr. İnci Togan ________________
Supervisor, **Department of Biological Sciences, METU**

Examining Committee Members

Prof. Dr. Zeki Kaya ________________
Department of Biological Sciences, METU

Prof. Dr. İnci Togan ________________
Department of Biological Sciences, METU

Prof. Dr. Mesude İşcan ________________
Department of Biological Sciences, METU

Prof. Dr. Okan Ertuğrul ________________
Veterinary Faculty, Ankara University

Assoc. Prof. Dr. İrfan Kandemir ________________
Biology Department, Ankara University

Date: 13.01.2009

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: EREN YÜNCÜ
Signature:

ABSTRACT

MITOCHONDRIAL DNA (MTDNA) HAPLOGROUP COMPOSITION IN TURKISH SHEEP BREEDS

Yüncü, Eren
M.Sc. Department of Biology
Supervisor: Prof. Dr. İnci Togan

January 2009, 88 pages

In the present study, haplogroup composition of five native Turkish sheep breeds, (Karayaka, Akkaraman, Gökçeada, Dağlıç, Morkaraman) and two sheep breeds from neighboring countries (Herik from Iran, samples from Azerbaijan) were determined by single strand length polymorphism (SSCP) analysis of mitochondrial DNA (mtDNA) NADH dehydrogenase subunit 4 (ND4) region and restriction fragment length polymorphism (RFLP) analysis of mtDNA control (CR) region.

Results of the SSCP and RFLP approaches were found to be 96,82% consistent. Most of the 3,18% inconsistency was due to unidentified band patterns of 9 individuals. SSCP method could identify haplogroups A, B and C, but not D and E. Similarly RFLP method could identify haplogroup A, B and possibly D, but not E and C. Data of the present study were compared with those of the previous studies to

test the robustness of results under different samplings and were found to be homogenous with a previous study with similar sampling strategy.

Neighbor joining tree, principal component analysis (PCA), Delaunay network analysis and analysis of molecular variance (AMOVA) were employed to analyze the haplogroup frequencies and breeds were separated in four groups according to the genetic barriers between breeds from different geographical locations. Strongest differentiation was present between two groups which were eastern breeds (Morkaraman, Herik-Iran and Azerbaijan) and western breeds (Gökçeada, Akkaraman, Karayaka and Dağlıç). Additionally, Azerbaijan was proposed as the entrance point of the haplogroup A and the Iran was proposed as the entrance point of haplogroup C to Anatolia with the Spearman rank correlation test.

Keywords: Turkish native sheep breeds, Mitochondrial DNA (mtDNA), Control (CR) Region, ND4 Region, SSCP, RFLP

ÖZ

TÜRK KOYUN IRKLARINDA MİTOKONDRİYAL DNA (MTDNA) HAPLOGRUP KOMPOZİSYONU

Yüncü, EREN
Yüksek Lisans, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. İnci Togan

Ocak 2009, 88 sayfa

Bu çalışmada beş yerli koyun ırkı (Karayaka, Akkaraman, Gökçeada, Dağlıç ve Morkaraman) ile komşu ülkelerden toplanan iki ırkın (İran'dan Herik ırkı ve Azerbaycan örnekleri2) haplogrup kompozisyonu, mitokondriyal DNA ND4 bölgesinin tek zincir konformasyonel polimorfizmi (SSCP) ve mitokondriyal DNA kontrol (CR) bölgesi restriksiyon parça uzunluk polimorfizmi (RFLP) analizleri kullanılarak saptanmıştır.

SSCP ve RFLP yöntemlerinin sonuçları birbirleri ile %96,82 uyumlu bulunmuştur. Sonuçlardaki %3,18 uyumsuzluğun büyük bir kısmı ise 9 örnekteki tanımlanamayan bant örgüleri nedeniyle ortaya çıkmıştır. SSCP yöntemi haplogrup A, B ve C'yi tanımlayabilmektedir ancak D ve E bu yöntemle ayrılamamıştır. RFLP yöntemi Haplogrup A, B ve büyük ihtimalle D'yi tanımlayabilmektedir ancak E ve C bu yöntemle ayrılamamıştır. Bu çalışmanın sonuçları, farklı örnekleme ile bulunan

haplogrup frekans sonuçlarının güvenilirliğini test etmek için önce yapılan çalışmaların sonuçları ile karşılaştırılmıştır ve sonuçlar daha önce benzer örnekleme stratejisi ile yapılan bir sonuç ile homojen bulunmuştur.

Komşu birleştirme ağacı, temel öğeler analizi, Delaunay örgüsü ve moleküler varyans analizi (AMOVA), haplogrup frekanslarına uygulanmıştır ve ırklar bulundukları coğrafi bölgelerin arasındaki genetik bariyerlere göre 4 gruba ayrılmıştır. En güçlü ayrım batı ırkları (Gökçeada, Akkaraman, Karayaka ve Dağlıç) ve doğu ırkları (Morkaraman, Herik-İran ve Azerbaycan) olarak ayrılan iki grup arasında bulunmuştur. Ayrıca Spearman'ın sıralama korelasyon testi sonuçlarına göre Azerbaycan haplogrup A'nın Anadolu'ya giriş noktası olarak, İran ise haplogrup C'nin Anadolu'ya giriş noktası olarak önerilmiştir.

Anahtar Kelimeler: Türk yerli koyun ırkları, Mitokondriyal DNA (mtDNA), Kontrol (CR) bölgesi, ND4 bölgesi, SSCP, RFLP

ACKNOWLEDGEMENTS

First I would like to express my deepest gratitude and thanks to my supervisor Prof. Dr. İnci Togan for her guidance, attention, recommendations and support throughout my MSc studies and this research.

I am indebted to my lab mates and my friends Havva Dinç, Dr. Ceren Berkman, Anıl Doğan, Assist Prof. Dr. Emel Özkan, Dr. Evren Koban, Eda Şeyma Kepenek, Hande Acar, Dilek Çevik who were always with me and endlessly helped me in every step of my experiments, and preparation of this thesis.

I would like to thank to Bekir Ankaralı for the collection of native sheep individuals and to TURHAYGEN-I project team members for taking the blood samples. Also I would like to thank Refik Gurbanov and Rodin Mazaheri for collecting samples from Iran and Azerbaijan.

My special thanks go to to all my dear friends for their invaluable moral support, motivation and companionship.

Finally, I declare my best appreciation to my parents Hafit and Selma Yüncü, my sister Bilgen Yüncü and all my family for their endless help, support and confidence.

This study was supported by Scientific and Technical Research Council of Turkey (TUBITAK) as a part of the project In Vitro Conservation and Preliminary Molecular Identification of Some Turkish Domestic Animal Genetic Resources-I (TURKHAYGEN-I) under the grant number 106G115 and METU-BAP under the grant number 2006-07-02-00-01.

TABLE OF CONTENTS

ABSTRACT iv
ÖZ..... vi
ACKNOWLEDGEMENTS..... viii
TABLE OF CONTENTS..... ix
LIST OF FIGURES..... xii
LIST OF TABLES xiv
LIST OF ABBREVIATIONS xvi
CHAPTERS
1.INTRODUCTION..... 1
1.1 Mitochondrial DNA (mtDNA) as Marker and Sheep mtDNA 4
1.2 Domestication of Sheep 5
1.3 Importance of Turkish Sheep Breeds..... 6
1.4 Mitochondrial DNA Control (CR) Region 8
1.5 Mitochondrial DNA NADH Dehydrogenase Subunit 4 (ND4) Region 8
1.6 Purposes and Outcomes of the Study..... 9
2.MATERIALS AND METHODS 10
2.1 Materials 10
2.2 Methods 12
2.2.1 DNA Isolation..... 12
2.2.1.1 DNA Isolation from Blood Samples 12
2.2.1.2 DNA Isolation from Buccal Samples..... 12
2.2.1.3 Checking the Presence of DNA..... 13

2.2.2 Polymerase Chain Reaction (PCR) 14
2.2.2.1 Amplification of mtDNA CR Region 14
2.2.2.2 Amplification of mtDNA ND4 Region... 15
2.2.2.3 Checking the Presence of PCR Products 16
2.2.3 RFLP Analysis of CR Region 17
2.2.4 SSCP Analysis of ND4 Region 17
2.2.5 Statistical Analysis 18
2.2.5.1 Chi Square (X^2) Test of Homogeneity 18
2.2.5.2 Frequency of Haplogroups.. 20
2.2.5.3 Pairwise F_{ST} Values 21
2.2.5.4 Nei's Genetic Distance (D_A) 22
2.2.5.5 Neighbor Joining (NJ) Tree 23
2.2.5.6 Spearman Rank Correlation Test.. 23
2.2.5.7 Delaunay Network Analysis 24
2.2.5.8 Principal Component Analysis (PCA) 25
2.2.5.9 Analysis of Molecular Variance (AMOVA) 25
3.RESULTS 28
3.1 Results of the Laboratory Experiments.. 28
3.1.1 DNA Extraction and PCR 28
3.1.2 Genotyping... 31
3.1.2.1 Genotyping of mtDNA CR Region by RFLP 31
3.1.2.2 Genotyping of mtDNA ND4 Region by SSCP 32
3.2 Results of Statistical Analyses 35
3.2.1 Comparison of Results Obtained by Two Methods 35
3.2.2 Frequency of Haplogroups... 36
3.2.3 Pairwise F_{ST} Values... 38
3.2.4 Nei's Genetic Distance (D_A) Between The Breeds 40

3.2.5 Neighbor Joining (NJ) Tree of the Breeds 42
3.2.6 Analysis of Geographical Distribution of the Breeds 43
3.2.6.1 Spearman Rank Correlation Test 43
3.2.6.2 Delaunay Network Analysis 45
3.2.6.3 Principal Component Analysis (PCA) 46
3.2.6.4 Analysis of Molecular Variance (AMOVA) 48
4.DISCUSSION 54
4.1 Consistency of Results Obtained by Two Methods 55
4.2 Distribution of HPG Frequencies in Different Studies 56
4.3 Interpretations of Statistical Analyses 59
5.CONCLUSIONS 66
6.RECOMMENDATIONS FOR FUTURE STUDIES 68
REFERENCES 69
APPENDIX A 76
APPENDIX B 78
APPENDIX C 87
APPENDIX D 88

LIST OF FIGURES

Figure 1.1 Three main livestock domestication centers on a world map, taken from Bruford *et al.* (2003) 2

Figure 1.2 Domestication sites of the sheep, pig, cattle and goats in Fertile Crescent taken from Zeder (2008) 3

Figure 2.1 Location of the sample sites 11

Figure 3.1 Gel image of total DNA extracts after isolation 29

Figure 3.2 Gel image of mtDNA CR region PCR amplification products 30

Figure 3.3 Gel image of partial mtDNA ND4 region PCR amplification products ... 30

Figure 3.4 Restriction enzyme profiles of mtDNA CR analysis 31

Figure 3.5 Restriction enzyme profiles of unidentified band patterns 32

Figure 3.6 Conformational polymorphism profiles of mtDNA ND4 analysis. 33

Figure 3.7 Conformational polymorphism profiles of unidentified band patterns 34

Figure 3.8 Conformational polymorphism profiles of HPG-D and HPG-E together with unidentified band patterns of SSCP and RFLP method 35

Figure 3.9 Distribution of mtDNA haplogroups in 7 sheep breeds from Gökçeada to Iran region 37

Figure 3.10 NJ tree that was constructed according to the Nei's D_A genetic distances between seven populations. 42

Figure 3.11 NJ tree that was constructed according to the pairwise F_{ST} values between seven populations 43

Figure 3.12 Delaunay Network Approach performed by using Nei's D_A values. Dots represent the locations of the sampling sites of the breeds... 45

Figure 3.13 Delaunay Network Approach performed by using pairwise F_{ST} values. Dots represent the locations of the sampling sites of the breeds. 46

Figure 3.14 Principal component analysis of seven breeds in 2 dimensional model based on mtDNA haplogroup frequencies 47

Figure 4.1 Comparison of the data in the present study and other studies 57

LIST OF TABLES

Table 2.1 PCR mixture for the CR region. 15

Table 2.2 PCR conditions for the CR region. 15

Table 2.3 PCR mixture for the ND4 region. 16

Table 2.4 PCR conditions for the ND4 region. 16

Table 2.5 The enzyme digestion mixture for CR region. 17

Table 2.6 Representative contingency table for the X^2 test of homogeneity. 19

Table 2.7 General AMOVA table taken from Arlequin package program (Excoffier *et al.*, 2006). 27

Table 3.1 Ambiguous and different haplogroups based on RFLP and SSCP screening. 36

Table 3.2 Haplogroup frequencies of seven studied breeds. 37

Table 3.3 Pairwise F_{ST} values of the seven studied breeds based on mtDNA haplogroups. 39

Table 3.4 Pairwise Nei's D_A genetic distance values of the seven studied breeds based on mtDNA haplogroups. 41

Table 3.5 Conducted r_s values for Spearman Rank Correlation tests. 44

Table 3.6 Spearman Rank Correlation r_s coefficient when Gökçeada population was not included from the analysis. 44

Table 3.7 AMOVA table for the two major groups determined based on PCA-C1 and the neighbor joining trees. 49

Table 3.8 AMOVA table for the three groups determined based on PCA-C1 and PCA-C2 dimensions. 50

Table 3.9 AMOVA table for the three groups determined based on first and third genetic barriers in Delaunay Network Analysis that formed with Nei's D_A genetic distances. 51

Table 3.10 AMOVA table for the three groups determined based on first and second genetic barriers in Delaunay Network Analysis that formed with Nei's D_A genetic distances. 52

Table 3.11 AMOVA table for the four groups determined based on Delaunay Network Analysis that formed with Nei's D_A genetic distances. 53

Table 4.1 Computed X^2 values of the X^2 test of homogeneity for the data of present study and Pedrosa *et al.*'s (2005); Meadows *et al.*'s (2007); Koban *et al.'s* (2008) study. 58

LIST OF ABBREVIATIONS

°C : Degrees Celsius

µl : Microliter

AMOVA: Analysis of Molecular Variance

APS: Ammonium Per Sulfate

Arlequin: An Integrated Software Package for Population Genetics Data Analysis

bp : Base Pair

BP: Before Present

BSA: Bovine Serum Albumine

CR: Control Region

dNTP: Deoxynucleotide Triphosphate

dH_2O : Distilled Water

DNA : Deoxyribonucleic Acid

EDTA : Ethylene Diamine Tetra Acetic Acid

e.g: For example

HPG: Haplogroup

K_3EDTA: Potassium EDTA

M: Molar

MEGA : Molecular Evolutionary Genetics Analysis

mg: Miligram

$MgCl_2$: Magnesium Chloride

MARA: Ministry of Agriculture and Rural Affairs

ml: Milliliter

mM: Millimolar

mtDNA : Mitochondrial DNA

NaAc: Sodium Acetate

ND4: NADH Dehydrogenase Subunit 4

ng: Nanogram

NJ: Neighbor Joining

NTSYS: Numerical Taxonomy and Multivariate Analysis System

PAGE: Polyacrylamide Gel Electrophoresis

PCR: Polymerase Chain Reaction

pH : Potential of Hydrogen

PHYLIP: Phylogeny Inference Package Software

pmol: Picomoles

RFLP: Restriction Fragment Length Polymorphism

rpm : Rotations per Minute

RT: Room Temperature

SDS: Sodium Dodecyl Sulfate

SSCP: Single Strand Conformational Polymorphism

Taq : Thermus aquaticus

TBE: Tris Borate EDTA

TEMED: Tetramethylethylenediamine

UV: Ultra Violet

V: Volt

CHAPTER 1

INTRODUCTION

Earliest domestication is believed to be started during the short dry period that aroused at the end of the Pleistocene. Climatic and environmental changes in this period caused a ‘dry climate and the steppe-desert dry environment’ and mammals, which were the preys of humans, moved to the ‘well forested and well watered mountains’. Humans became less successful in hunting, so they started to domesticate first plants and then animals to meet the increased food demand (Bökönyi, 1976). Domestication did not only provide a steady food supply like meat, milk and eggs, it also rendered companionship, protection (as by the dogs), clothing material like wool and hides and certain jobs like carrying heavy loads.

Domestication is believed to occur in three main areas. Two of these areas: Southwest Asia (place named as The Fertile Crescent and east of the Fertile Crescent to the Indus valley) and East Asia (China and south of China) is the place where cattle, sheep, goats, pigs and buffalos were domesticated. The third one, Adnean chain of South America is known as the domestication center of llamas and alphacas (Bruford *et al.*, 2003). Figure 1.1 represents three main domestication centers on a world map taken from Bruford *et al.* (2003).

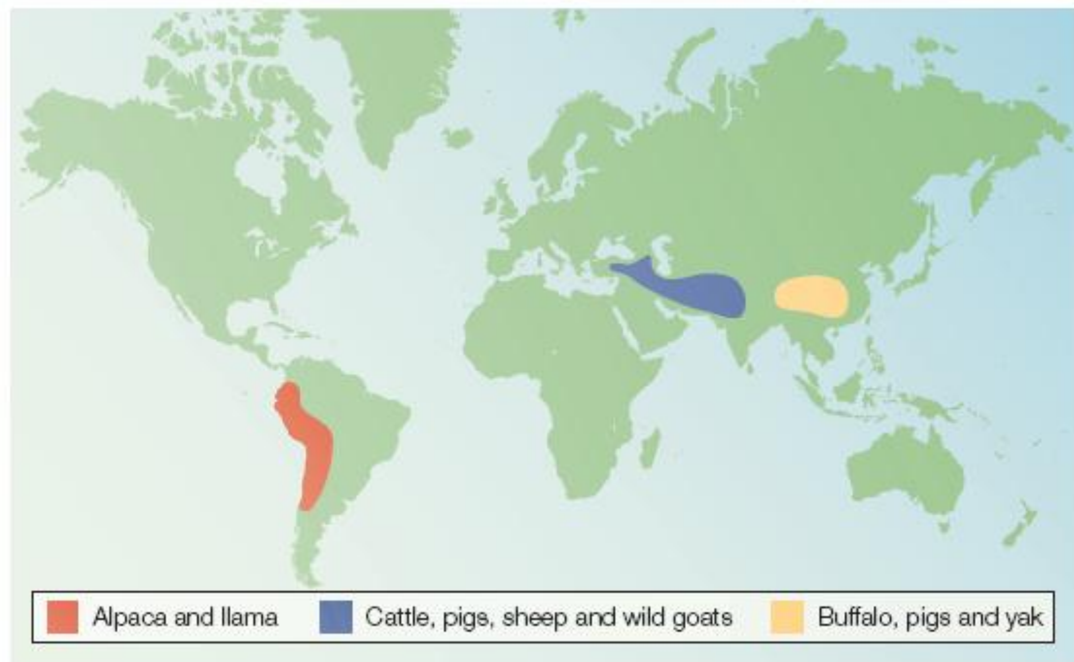


Figure 1.1 Three main livestock domestication centers on a world map, taken from Bruford *et al*. (2003).

The Fertile Crescent is the region that extends among today's Israel, Jordan, Lebanon, west of Syria, southeast of Turkey, along the Tigris and Euphrates rivers, Iraq and the western of Iran. (Bruford *et al*., 2003) It compromises the areas of Northern Levant, Western Zagros east Anatolia and central Anatolia. As revealed by archeological remains, this area is also the site of important domestication centers like Çatalhöyük and Zagros (Renfrew, 1991). Furthermore, new advancements in archeological methodology refined the locations of the earliest domestication sites of the sheep, cattle, goats and pigs (Zeder, 2008). Figure 1.2 represents the domestication sites of these animals on a map constructed very recently (Zeder, 2008).

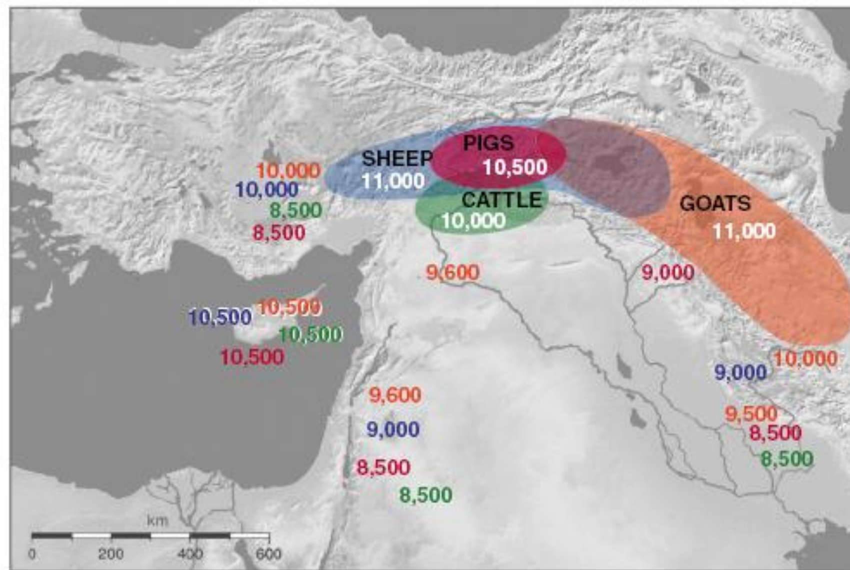


Figure 1.2 Domestication sites of the sheep, pig, cattle and goats in Fertile Crescent taken from Zeder (2008). Numbers inside the shaded regions represents the date of the initial domestication in years scaled before present (BP) and number outside of the shaded region are the dates of appearance of the first domestics in the region in years scaled BP (Zeder, 2008).

The new map indicates that Anatolia is the center of domestication for many important livestock species and it is highly possible that Anatolian native breeds might be the extents of earliest domesticated animals (Zeder, 2008).

Wild looking sheep and goats in the Europe are shown to be the only feral descendants of the Near Eastern caprines by the means of genetic analyses (Bruford and Townsend, 2006; Luikart *et al.*, 2006). So these domestication products must be diffused to Europe from Near East. It is known that Neolithic farmers travel from Near East to Cyprus first, then they diffused to Mediterranean Basin and Europe and it is proposed that they carried domesticated livestock species with them (Zeder, 2008). Also it is suggested that many colonists migrated from Near East to Iberian Peninsula through Mediterranean with these livestock species which leads to high gene flow between Iberian Peninsula and Near East (Pereira *et al.*, 2006).

It can be anticipated that during the colonization, subsets of the gene pool was taken to new areas and most of the genetic diversity is lost, as colonists went from the center of domestication to the periphery. Hence it is well accepted that Anatolian native breeds are very special because they may still harbor very valuable genetic diversity to be explored. Therefore they must have higher priority in conservation (Bruford *et al*., 2003; Zeder, 2008).

1.1 Mitochondrial DNA (mtDNA) as Marker and Sheep mtDNA

Mitochondrial DNA (mtDNA) is a circular DNA, found in the mitochondrion of the cell. It is an important molecular marker, widely used to trace domestication history. It has high rate of evolution so recently diverged species can be studied comparatively; since it does not go under recombination, differences between mitochondrial sequences occur due to mutations only. Therefore, dates of certain polymorphisms can be estimated and hence evolutionary history of the species based on mtDNA can be obtained. Yet, it must be emphasized that mtDNA reflects only the history of the maternal lineage. Data obtained from mtDNA can scarcely be used for interpreting the evolution of whole genomic diversity because it acts as a one gene and it has its own specific 'evolutionary dynamics' (Bruford *et al.,* 2003; Bruford 2004).

Complete sheep mitochondrial DNA is found to be approximately 16 616 base pair (bp) long. Length of the strands can vary because of different number of 75bp long tandem repeats which contain two octamer sequences of mirror symmetry (TTAATGTA, TACATTAA). (Hiendleder *et al*., 1998[a])

Firstly two common distinct mtDNA haplogroups (HPG's) were identified and called as HPG-A and HPG-B (Wood and Phua, 1996; Hiendleder *et al.,* 1998[b]). In addition to these two common haplogroups, recently three more haplogroups were identified. These haplogroups are called as HPG-C, HPG-D and HPG-E. (Guo *et al.*, 2005;

Pedrosa *et al.*, 2005; Tapio *et al.*, 2006; Bruford and Townsend, 2006; Meadows *et al.*, 2007, Koban *et al.*, 2008). Additionally in the study of Guo *et al.*'s (2005) a new candidate haplogroup, which is named as HY030, is found. It is shown that this haplogroup is not clustered with any of these five haplogroups. (Koban *et al.*, 2008) Three of the haplogroups are also named according to their distribution patterns. HPG-A, which is dominantly found in Asia, is called as Asian Lineage. HPG-B is called as European Lineage because it is the only haplogroup that can be found predominantly in Europe and HPG-C is called as South Asian Lineage because it is frequently found in South Asia (Hiendleder *et al.*, 1998[b]; Guo *et al.*, 2005; Bruford and Townsend, 2006). HPG-D and HPG-E are rarely indentified among sheep in Eurasia so far but these HPG's were found in Turkish domestic sheep breeds (Pedrosa *et al.*, 2005; Meadows *et al.*, 2007; Koban *et al.*, 2008).

1.2 Domestication of Sheep

Domestic sheep (*Ovis aries*), member of genus *Ovis* were among the first domesticated animals. Archeological findings show they were domesticated approximately 11 000 BP in the Southwest Asia. (Bruford *et al.*, 2003; Zeder, 2008)

Morphological changes used to be the indicator for detecting domestication. Reduction in the overall body size that is observed in archaeological sheep findings (e.g.bones) were interpreted by domestication of these sheep populations. But it is observed gender differences and environmental factors were affecting this reduction. Recently, indication of the start of domestication is expected as the deviation of female to male ratio in favor of females and therefore sites and dates of the earliest domestication of sheep have changed. (Zeder, 2008) Recently, the accepted sites of the earliest sheep domestication and locations are as it was represented by Figure 1.2 (Zeder, 2008).

Presence of distinct mtDNA haplogroups usually interpreted as the existence of different domestication events, which may be taking place in different domestication centers (Luikart *et al.*, 2001). For sheep, at least three different domestication events, which took place in separate domestication centers or at different times in the same region is assumed. The oldest sheep domestication center is believed to harbor archaeological sites Nevali Çori and Çayönü in Turkey (Peters *et al.*, 1999). Second sheep domestication center is believed to be in the Zagros Mountains which placed in the border of Iran, Iraq and Turkey (Zeder, 1999).

Ancestors of domestic sheep are still unclear. Some candidate wild sheep populations were studied for this purpose. Wild sheep species *Ovis vignei* (Urial sheep) and *Ovis ammon* (Argali sheep) were shown to be genetically distant so less likely to be related with domestic sheep (Hiendleder *et al.*, 1998[b]). Wild looking sheep used to be named as *Ovis musimon* (European muflon) is shown to be a feral descendant of first domestic sheep entering Europe because they are genetically very close to the domestic sheep mtDNA European lineage B, and fossil records of the prior of Neolithic is not harboring *Ovis* genus (Clutton-Brock, 1981; Hermans, 1996; Hiendleder *et al.*, 2002; Pedrosa *et al.*, 2005; Bruford and Townsend, 2006; Meadows *et al.*, 2007). Currently O*vis gmelini* is the strongest candidate (Hiendleder *et al.*, 2002; Pedrosa *et al.*, 2005; Bruford and Townsend, 2006; Meadows *et al.*, 2007, Koban *et al.*, 2008). Question of ancestors of domestic sheep is not answered yet.

1.3 Importance of Turkish Sheep Breeds

Turkey is harboring the earliest sheep domestication center. So, Turkish sheep can be closest living relatives of first domesticated ancestors. Due to the proximity to the domestication centers. Since the highest diversity is expected to be maintained at the earliest domestication events (Bradley *et al.*, 1996; Loftus *et al.*, 1999), Turkish sheep are expected to have high genetic diversity. Indeed there are some studies that

show Turkish sheep are exhibiting high genetic diversity (Bruford *et al.*, 2003; Pedrosa *et al.*, 2005; Bruford and Townsend, 2006; Uzun *et al.*, 2006; Lawson-Handley *et al.*, 2007; Meadows *et al.*, 2007; Peter *et al.*, 2007; Koban *et al,*. 2008).

It is very well known that Turkish sheep have prime importance in conservation studies (Bruford *et al.*, 2003; Zeder, 2008). Unfortunately, native Turkish sheep breeds are shown to be under a severe threat. According to the Longworth, (2005)'s study number of sheep in Turkey is decreased by 47% in the last two decades due to indiscriminate hybridization and hybridization according to the demands of market economy. If only the native breeds are considered there are 13 sheep breeds in Turkey. These are Karayaka, Akkaraman, Kıvırcık, Dağlıç, Herik, Gökçeada, Morkaraman, Sakız, İvesi, Çine Çapari, Hemşin, Tuj and Norduz. All these native Turkish breeds have distinct geographic distributions. However in Turkey there is no herd-book based mating practice, hence they are not genetically isolated (Akçapınar, 2000). It is quite expensive to conserve them all. To prioritize the breeds, conservation strategies can be developed. Koban *et al.*'s (2008) study revealed that, Dağlıç, and perhaps Karayaka were the closest relatives of the earliest sheep breeds of Anatolia. However confidence level of the previous data must be evaluated. If the previous conclusion was made based on a data which cannot be repeated, then the conservation studies based on the previous data can be highly inaccurate.

In the present study, mtDNA haplogroup compositions of five native domestic breeds; Karayaka, Akkaraman, Gökçeada, Dağlıç and Morkaraman, with two foreign breeds; Herik from Iran and samples from Azerbaijan were examined. Since breeds in Azerbaijan are mixed with each other, different breeds cannot be identified (personal communication). For this reason samples collected from this location is called as Azerbaijan breed. While analyzing the haplogroup frequency of the breeds two different regions, Control (CR) region (Bruford and Townsend, 2006; Koban *et al.*, 2008) and NADH dehydrogenase subunit 4 (ND4) region (Guo *et al.*, 2005) of mtDNA were analyzed separately and results were compared.

1.4 Mitochondrial DNA Control (CR) Region

Control region is the most rapidly evolving region of the mtDNA. It is nearly 1260bp long and contains the displacement loop (d-loop/origin of replacement) (Bruford, 2004). Tandem repeats (as explained above) are in this region so the length of the CR region can change due to this vary number of tandem repeats. Number of these repeats is usually four, in some samples these numbers are changing between 3 and 5 (Hiendleder *et al.*, 1998[a]). Since it is rapidly evolving it is frequently used to track geographic patterns of diversity and evolution studies (Bruford *et al.*, 2003). In the studies of Bruford and Townsend's (2006); Koban *et al.*'s (2008) it is shown that Restriction Fragment Length Polymorphism (RFLP) method can be used to analyze CR region to determine the haplogroup of the sheep. It is shown that, by this method, HPG-A can be differentiated by one specific band pattern. HPG-B is shown to have two band patterns due to the G-A transition in the enzyme restriction site. One is specific for HPG-B and the others have the same band pattern with HPG-C. So haplogroups can be differentiated as HPG-A, HPG-B and HPG-B/HPG-C (Bruford and Townsend, 2006; Koban *et al.*, 2008).

1.5 Mitochondrial DNA NADH Dehydrogenase Subunit 4 (ND4) Region

ND4 region is 1378 bp long gene (Hiendleder *et al.*, 1998[a]) which codes NADH dehydrogenase subunit 4. In the study of Guo *et al.*'s (2005) it is shown that Single Strand Conformational Polymorphism (SSCP) can be used to analyze partial ND4 region to determine the haplogroup of the sheep. It is shown that same three major haplogroups HPG-A, HPG-B and HPG-C can be determined distinctly by SSCP analysis of partial ND4 (Guo *et al.,* 2005).

1.6 Purposes and Outcomes of the Study

Turkish sheep breeds have prior importance in conservation studies and to make appropriate breed priorization, genetic characterization of the breeds must be made accurately. Purposes of the study are to compare the efficiency two methods, SSCP analyses of ND4 region and RFLP analyses of CR region, which were separately proposed for determination of sheep mtDNA haplogroups and to check the consistency in their results.

Outcomes the study are as follows:

- To test the sufficiency of the employed sampling strategy in characterizing the breeds of the present study.

- To make contributions to the studies on evolutionary history of domestic sheep by obtaining and analyzing genetic data of some Turkish breeds.

- To study and use widely employed statistical analyses on the obtained data to understand evolutionary history of domestic sheep.

- To contribute conservation of domestic sheep studies in Turkey.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

In this study, 314 individuals from 5 different Turkish sheep breeds (Karayaka, Akkaraman, Gökçeada, Dağlıç and Morkaraman) and 2 different sheep breeds from the neighboring countries (Herik-Iran and Azerbaijan) were analyzed. Turkish sheep were represented with 50 individuals for each breed and collected by Ministry of Agriculture and Rural Affairs (MARA) within project TURKHAYGEN-I (Project No: 106G115) Samples from Karayaka breed were collected from Ordu and Tokat provinces in Black Sea Region, Akkaraman breed was collected from Konya province in Central Anatolia Region, Gökçeada breed was collected from an Aegean Island, Gökçeada, near to Çanakkale in Marmara Region, Dağlıç breeds was collected from Afyon province in Aegean Region and Morkaraman breeds were collected from Erzurum province in East Anatolia Region. For the foreign breeds, Herik samples were collected from local Iran villages and Azerbaijan samples were collected from local Azerbaijan villages. Breeds were collected from different local farms and only 2-3 individuals were collected from one herd. Also samples were collected so that individuals representing the breeds were not relatives. Figure 2.1 represents the location of the samples sites on a Turkey map.

Figure 2.1 Location of the sample sites.

Samples were collected by two types of tissues. Blood samples were collected from Turkish breeds and epithelial cell (as buccal swap samples from inside the cheek) samples were collected from neighboring breeds. Blood samples were collected from each individual by taking 10ml blood into a vacuum blood tube. Vacuum tube contains K_3EDTA to prevent coagulation. Samples were stored in 4°C until DNA isolation. Buccal samples were collected from both of the cheeks of each individual, by strongly scraping with a sterile brush to collect as many cells as possible. The scraping should be continued at least one minute.

Ingredients of chemical solutions used in experiments are presented in Appendix A.

2.2 Methods

2.2.1 DNA Isolation

2.2.1.1 DNA Isolation from Blood Samples

Each blood sample (10ml) was poured to a falcon tube that was already containing 0.5ml EDTA and completed to 50ml by 2X lysis buffer, which was used to lyse red blood cells. All tubes were mixed by inversion for 10 minutes then waited in ice for 30 minutes. After this step tubes were centrifuged for 15 minutes, at 3000rpm, in 4°C. Supernatants were discarded and 3ml of Salt-EDTA buffer (pH:7.5), 0.3ml SDS solution (%10) and 0.15ml proteinase K solution (10mg/ml) were added to each tube. Tubes were incubated for 3 hours in 55°C. After incubation, 3ml phenol was added to each tube; tubes were mixed by inversion for 10 minutes and centrifuged for 15 minutes at 3000rpm in 4°C. After centrifugation supernatants were taken to new falcon tubes and phenol chloroform isoamylalcohol (in ratio of 25:24:1) was added to each tube. Tubes were centrifuged for 15 minutes at 3000rpm in 4°C. After centrifugation supernatants were taken to glass tubes and DNA was precipitated by shaking each sample with 2 volume of pure cold ethanol. Each precipitated DNA sample was taken to 1.5ml eppendorf tube and 0.5ml Tris-HCl (pH:8.0) was added to each tube. DNA samples were stored in -20°C. (Sambrook *et al.*, 1989)

2.2.1.2 DNA Isolation from Buccal Samples

The head of the brushes for each sample were placed into an eppendorf tube; 700µl 2X lysis buffer and 10ml proteinase-K solution (20mg/ml) were added to each tube. Tubes were digested overnight at 37°C. After digestion 500µl phenol chloroform isoamylalcohol (in ratio of 25:24:1) was added to each tube and tubes were centrifuged for 2 minutes at 10000rpm, in room temperature (RT). The upper aqueous phases were transferred to new eppendorf tubes and phenol chloroform

isoamlyalcohol extraction was repeated. After transferring the aqueous phases into new eppendorf tubes, 500µl chloroform isoamlyalcohol (in ratio of 24:1) was added to each tube and tubes were centrifuged for 2 minutes, at 10000rpm in RT. After centrifugation, aqueous phases were transferred to a new tube. 0.1 volume of sodium acetate (NaAc,3M) and 0.6 volume of isopropanol were added to each tube and tubes were waited overnight in -20°C. After this step, tubes were centrifuged for 10 minutes, at 10000rpm, in RT and each pellet was rinsed by adding 500µl cold ethanol (70%) to each tube. Each DNA was precipitated again by centrifuging tubes for 5 minutes, at 10000rpm in RT. Supernatants were discarded, DNA's were dried at room temperature and melted in 100µl of Tris-HCl (pH:8.0). Samples were stored in -20°C (Dinç, 2003).

2.2.1.3 Checking the Presence of DNA

After the isolation process presence, concentration and the amount of DNA were checked by agarose gel electrophoresis. 0.8% agarose gel was prepared by boiling agarose in 0.5X TBE buffer. Then agarose solution was poured into an electrophoresis plate and waited for approximately 20 minutes for polymerization. Agarose plate, which carries polymerized agarose gel, was placed into electrophoresis tank and the tank was filled with 0.5X TBE buffer until the all gel was covered with buffer. Samples were prepared for loading by mixing 2µl DNA with 3µl, 6X loading buffer (bromophenol blue dye) and 3µl dH_2O and loaded into the wells of the gel. The gel was run at 100V, for 20 minutes. After this step, gel was stained by etidium bromide (Et-Br) solution (0.5µl/ml) and visualized and photographed under UV light by a gel image system. The presence, concentration and the quality of the DNA were decided by observing the thickness, presence of smears and the migration patterns of the corresponding bands on the gel.

DNA samples with high concentration were diluted to use in PCR. First, these samples were incubated in 55 C° for 20 minutes or in 37°C for overnight, to dissolve

DNA properly. After incubation certain amount of DNA was taken to a new tube and diluted with Tris-HCl (pH:8.0) buffer. Diluted samples were checked with 1% agarose gel again. Amount of samples that must be put to PCR was decided by comparing DNA bands with the bands of a standard lambda DNA marker.

2.2.2 Polymerase Chain Reaction (PCR)

Complete mtDNA CR region and partial mtDNA ND4 region were amplified by using Polymerase Chain Reaction (PCR). This technique is based on three steps which are called denaturation, annealing and extension. In denaturation DNA's single strands are separated from each other, in annealing oligonucleotide primers are bound to their complements on single stranded template DNA and in extension primers are extended in the 5'- to- 3' direction. As a result, specific DNA regions were enzymatically amplified in microgram quantities (Klug and Cummings, 2000).

2.2.2.1 Amplification of mtDNA CR Region

Complete sheep mtDNA CR region, which is 1260 bp long (vary because of tandem repeats, as explained in introduction) was amplified by PCR (Hiendleder *et al.*, 2002). Primers (Hiendleder *et al.*, 2002) that were used for amplification is given below.

Forward-Proline (Pro) : 5'-CTCACCATCAACCCCCAAAGC-3'
Reverse-Phenylalanine (Phe) : 5'-TCATCTAGGCATTTTCAGTG-3'

Table 2.1 represents the concentrations of the PCR mixture contents and Table 2.2 represents the PCR conditions.

Table 2.1 PCR mixture for the CR region.

PCR Buffer	1X
$MgCl_2$	1.5mM
dNTP	0.2mM
Primer	5pmol
BSA	0.5mg/ml
Taq Polymerase	1u
DNA	50ng-100ng
Total Volume	**25µl**

Table 2.2 PCR conditions for the CR region.

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

2.2.2.2 Amplification of mtDNA ND4 Region

Partial sheep mtDNA ND4 region, which is 280 bp long, was amplified by PCR (Guo *et al*., 2005). OV11 primers (Guo *et al*., 2005) that were used for amplification is given below.

Forward -OV11: 5-GACTCCACCTCTGACTTCC-3

Reverse -OV11: 5-TGAATGAGAATGGCAACA-3

Table 2.3 represents the concentrations of the PCR mixture contents and Table 2.4 represents the PCR conditions.

Table 2.3 PCR mixture for the ND4 region.

PCR Buffer	1X
$MgCl_2$	2.5mM
dNTP	0.2mM
Primer	15pmol
BSA	0.4mg/ml
Taq Polymerase	0.5u
DNA	50ng-100ng
Total Volume	**15µl**

Table 2.4 PCR conditions for the ND4 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	1.15 seconds	
Final Extension	72°C	15 minutes	1

2.2.2.3 Checking the Presence of PCR Products

The results of the PCR amplification was checked by visualizing the presence of the PCR products with agarose gel electrophoresis. For ND4 region analysis %2 agarose gel and for CR region analysis %1 agarose gel were prepared by boiling agarose in 0.5X TBE buffer. Then agarose solution was poured into an electrophoresis plate and waited for approximately 20 minutes for polymerization. Agarose plate, which carries frozen agarose gel, was placed into electrophoresis tank and the tank was filled with 0.5X TBE buffer until the all gel was covered with buffer. Samples were prepared for loading by mixing 2µl DNA with 3µl, 6X loading buffer (bromophenol blue dye) and 3µl dH_2O and loaded into the wells of the gel. The gel was run at 100V, for 20 minutes. After this step, gel was stained by etidium bromide (Et-Br) solution (0.5µl/ml) and visualized and photographed under UV light by a gel image

system. Bands that are expected base pair long presented near the corresponded bands of the marker after successful amplification.

2.2.3 RFLP Analysis of CR Region

RFLP method is based on using an enzyme for detecting the differences between sequences of haplogroups. Restriction sites, where the different sequences present, are recognized and digested with the enzyme or not. As a result, different lengths of fragments are obtained according to the haplogroups. (Lewis, 2005)

After amplification of CR region samples were digested with *NSiI* enzyme (Bruford and Townsend, 2006; Koban *et al*., 2008) in 37°C for overnight. Table 2.5 represents the enzyme digestion mixture for CR region.

Table 2.5 The enzyme digestion mixture for CR region.

Enzyme Buffer	1X
PCR Product	20µl
Enzyme (*NSiI*)	2u
Total Volume	**25µl**

After digestion products were mixed with 5µl loading dye (bromophenol blue) loaded to %3 agarose gel to check the results. Gel was run at 120V until the dye disappears. After this step, gel was stained by etidium bromide solution (0.5µl/ml) and visualized and photographed under UV light by a gel image system.

2.2.4 SSCP Analysis of ND4 Region

SCCP method is based on separating double strands of the DNA to identify the haplogroups. DNA strands take new conformations after separation of the strands

and these conformations will be different according to the different sequences of the haplogroups. (Klug and Cummings, 2000)

Amplification products of the ND4 region, 12µl from each PCR product, were mixed with 12µl formamide and incubated in 98ºC to separate double strands. After incubation, samples were placed into ice to stop the reaction. Results were visualized with Polyacrylamide Gel Electrophoresis (PAGE). Polyacrylamide gel was prepared by dissolving 12% acrylamide with bisacrylamide (39:1) in 0.5X 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 30-40 minutes for polymerization. After polymerization of the polyacrylamide gel samples were loaded to the wells. Gel was run with 100V for 19-22 hours in 0.5X TBE and bands were visualized by silver staining. (Guo *et al*., 2005) Silver staining procedure is as follows. First gel was incubated in 5% acetic acid solution for 10 minutes. After this step, acetic acid solution is poured and gel was imposed to 10% ethanol solution and incubated for 5 minutes. Then ethanol solution is poured and gel was incubated in 1X silver nitrate solution ($AgNO_3$) for 25 minutes. After 25 minutes gel was washed with deionized water for 90 seconds and incubated in developer solution (0.75M NaOH and 37% formaldehyde) until the bands are visualized. Photograph was taken under white light, by a gel image system (Bassam *et al*., 1991).

2.2.5 Statistical Analysis

2.2.5.1 Chi Square (X^2) Test of Homogeneity

Chi Square (X^2) test of homogeneity is used to check the assumption if samples that are drawn from the same population are similar (homogeneous) with respects to a frequencies of a variable (e.g. alleles, haplotypes or haplogroups of certain gene). Value of X^2 is calculated by obtaining expected and observed frequency values with

the help of a contingency table. Observed frequencies are the actual frequencies in the data and expected frequencies are the frequencies that are expected to be observed when the assumption is true. Table 2.6 represents the contingency table that is used to calculate the expected values for X^2 test of homogeneity (Daniel, 1999). Abbreviations used in table are as follows;

n_{A1} = Frequency of variable A in population 1.
n_{B1} = Frequency of variable B in population 1.
n_{C1} = Frequency of variable C in population 1.
n_{A2} = Frequency of variable A in population 2.
n_{B2} = Frequency of variable B in population 2.
n_{C2} = Frequency of variable C in population 2.
$n_A = n_{A1} + n_{A2}$ (Total number of individuals that carry variable A.)
$n_B = n_{B1} + n_{B2}$ (Total number of individuals that carry variable B.)
$n_C = n_{C1} + n_{C2}$ (Total number of individuals that carry variable C.)
$n_1 = n_{A1} + n_{B1} + n_{C1}$ (Total number of individuals in Population 1.)
$n_2 = n_{A2} + n_{B2} + n_{C2}$ (Total number of individuals in Population 2.)
$n = n_1 + n_2$ or $n_A + n_B + n_C$ (Total number of analyzed individuals.)

Table 2.6 Representative contingency table for the X^2 test of homogeneity.

	Variable Category			
	A	B	C	**Total**
Population 1	n_{A1}	n_{B1}	n_{C1}	$\mathbf{n_1}$
Population 2	n_{A2}	n_{B2}	n_{C2}	$\mathbf{n_2}$
Total	$\mathbf{n_A}$	$\mathbf{n_B}$	$\mathbf{n_C}$	**n**

Expected values of the frequencies are calculated according to the table as the example given below;

$$Expected\ n_{A1} = \frac{n_A \times n_1}{n}$$

where,

n_{A1} = Frequency of variable A in population 1.

$n_A = n_{A1} + n_{A2}$ (Total number of individuals that carry variable A.)

$n_1 = n_{A1} + n_{B1} + n_{C1}$ (Total number of individuals in Population 1.)

X^2 value is calculated by the formula given below;

$$X^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

Haplogroup frequencies of the four sheep breeds that were obtained from this study is compared with the data from Pedrosa *et al.*'s (2005), Meadows *et al.*'s (2007) and Koban *et al.*'s (2008) study by X^2 test of homogeneity. Significance of results was decided by comparing obtained p value with α values in given degrees of freedom. Values of X^2 and p were calculated by Minitab package program (Minitab Inc., State College, PA, USA).

2.2.5.2 Frequency of Haplogroups

Haplogroup frequency is a measure of the relative frequency of a haplogroup in a population and presents the genetic diversity of a species population. In the present study, there were 3 haplogroups for sheep mtDNA. Haplogroup frequencies of each breed were calculated in the studied seven populations (Karayaka, Akkaraman,

Gökçeada, Dağlıç, Morkaraman, Azerbaijan and Herik-Iran). Haplogroup frequencies for a population were calculated according to the formula given below;

$$Haplogroup\ Frequency = \frac{Total\ Number\ of\ Individuals\ Carrying\ the\ Haplogroup}{Total\ Number\ of\ Individuals}$$

2.2.5.3 Pairwise F_{ST} Values

F_{ST} is the type of distance measure that computes the degree of genetic differentiation between pairs of populations of species (Nei, 1977). It can be calculated by the formula given below:

$$F_{ST} = 1 - \frac{H_S}{H_T}$$

where,

H_S = Mean expected heterozygosity within populations.

H_T = Expected heterozygosity when two populations are considered as one big population. (Nei and Kumar, 2000)

F_{ST} values change between 0 and ± 1 whereas value "0" is for identical populations and "± 1" is for populations that share no alleles. F_{ST} values that are close to ± 1 refers to high genetic differentiation between populations usually due to low gene flow between these populations and F_{ST} values close to "0" refer to two genetically similar populations usually due to the high gene flow between these populations.

F_{ST} values by pairwise comparisons of seven breeds in this study, with their significance levels, were calculated by Arlequin package program (Excoffier *et al.*, 2006) Significance of calculated F_{ST} values were tested by applying 1000 random permutations.

2.2.5.4 Nei's Genetic Distance (D_A)

Nei's Genetic Distance (D_A) is commonly used measure calculate average genetic distances between populations (Takezaki and Nei, 1996). It can be calculated by the formula given below:

$$D_A = 1 - \frac{1}{r}\sum_{j}^{r}\sum_{i}^{m_j}\sqrt{x_{ij}y_{ij}}$$

where,

x_{ij} = Frequencies of the ith allele at the jth locus in samples X.

y_{ij} = Frequencies of the ith allele at the jth locus in samples Y.

m_j =Number of alleles at the jth locus.

r =Number of loci examined.

Nei's Genetic Distance (D_A) also change between 0 and 1 whereas value "0" is for identical populations and "± 1" is for populations that share no alleles.

Pairwise genetic distances of seven breeds in this study were calculated by GENDIST program in PHYLIP: Phylogeny Inference Package Software (Felsenstein, 1993).

2.2.5.5 Neighbor Joining (NJ) Tree

Neighbor Joining (NJ) tree is a distance based phylogenetic tree which is used to visualize genetic relationship of the populations, especially clusters. This tree is constructed by first picking two populations with least genetic distances and forming first two branches. Then, next least distant population is chosen and added as third branch and so on. NJ tree does not assume equal rate of evolution of the breeds after the divergence. Therefore is a realistic method for representing the evolution of breeds. (Allendorf and Luikart, 2007)

In this study, two NJ trees were constructed to observe the relationship of the seven populations. Pairwise F_{ST} values were used to obtain one NJ tree and Nei's D_A distances were used to obtain the other NJ tree. All analyses were performed with the help of the software called MEGA: Molecular Evolutionary Genetic Analysis (Tamura *et al.*, 2008).

2.2.5.6 Spearman Rank Correlation Test

Spearman rank correlation test is used to analyze if two sets of data ranks are correlated or independent. Two data sets, which are called as X and Y, are prepared by ranking them from 1 to n. Possible correlation is tested by calculating the Spearman rank correlation coefficient (r_s) with the formula given next page;

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

n= Number of values ($n<30$)

d_i=Difference between i[th] value of X and Y.

Values of r_s change between 0 and ± 1. Results which are “0” or close to “0”, shows that two data are independent. Results which are -1 or close to -1 mean negative correlation between data while, +1 or close to +1 mean positive correlation between data (Daniel, 1999).

In this study, Spearman rank correlation test was used to check the possible correlation between the HPG-A and HPG-C frequency and distance between sampling sites of the breeds and assumed point of entrance of these haplogroups, Azerbaijan and Iran respectively. Significance of the r_s value was determined by comparing calculated r_s with the critical value of r_s, in given degrees of freedom and α values.

2.2.5.7 Delaunay Network Analysis

Delaunay Network Analysis is performed to view possible genetic barriers between populations according to their geographical positions, by using Thiessian polygons. (Monmonier, 1973) A study area is constructed by symbolizing populations as points which placed on map relative geographical positions of populations. The closest geographical locations of the samples on the map are connected to form a network by the Delaunay triangulation approach (Brassel and Reif, 1979). In this approach first, points in the outer region of the studied are linked to construct the outer boundary of the study area. Then interior points are connected based on “short criteria distances” and triangles are formed. The name of the resulting network is "Delaunay Network". Genetic barriers are formed by using genetic distance measurements. All genetic distances are written to the matching sides of the triangles and a perpendicular is drawn to the triangle edges with highest pairwise distance in the outer boundary. Drawing of the perpendicular line is continued by choosing the edges with highest distance until the outside of the polygon is reached. The resulted lines represent the first barrier. Second barrier is started from second highest genetic distance and so on. Perpendiculars are drawn until all the outer edges are checked. When the perpendicular line drive circles inside the polygon, it shows there is no genetic

barrier. Priorities of the genetic barriers are determined by the order with which the barriers are drawn.

In this study, two different Delaunay Network Analyses were performed by using pairwise F_{ST} values and Nei's D_A genetic distances on the map of Turkey and neighbor regions as the study area.

2.2.5.8 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is used to determine the relationship of the populations and this relationship is defined based on relative positions of the populations in multi dimensional space. It is performed by composing two or three independent compounds from certain number of variables. Relationship of the populations is explained visualizing the relative positions of populations on a space constructed by these independent compounds. The highest variation of the data is explained by the first compound axis, the next highest variation is explained by second compound axis and so on. Weightings of the axes, which are displayed as equations, show the amount of contribution of the variable to the variation (Dytham, 2003).

For this study, three haplogroups were used as variables and two independent compounds were composed to execute two dimensional PCA. The computer program, named as NTSYS: Numerical Taxonomy and Multivariate Analysis System, is used to perform the PCA (Rohlf, 2000).

2.2.5.9 Analysis of Molecular Variance (AMOVA)

Analysis of Molecular Variance (AMOVA) is an approach that is used to assess the differentiation between hierarchical groups (Excoffier *et al.*, 1992). It exhibits the correlations among genotype distances at various hierarchical levels by using Φ

statistics, which is an analogue of the F-statistics. This approach is parallel to the Analysis of Variance (ANOVA). Three components are analyzed by ϕ statistics. The correlation of haplogroups within a breed relative to those from the whole population is called ϕ_{ST}. It measures the proportion of genetic variation when all breeds are considered as one population. The correlation of random haplogroups within a group of breeds relative to those drawn from the entire population is called ϕ_{CT}. It measures the proportion of genetic variation among groupings of breeds. Lastly, the correlation of random haplogroups within breeds relative to those within a grouping of breeds is called ϕ_{SC} and it measures the proportion of variation among breeds within a group (Hartl and Clark, 1997). Table 2.7 represents the design of AMOVA and formulas of the calculations taken from Arlequin package program (Excoffier *et al.*, 2006). Abbreviations used in table are as follows;

SSD (T): Total sum of squared deviations.
SSD (AG): Sum of squared deviations Among Groups of breeds.
SSD (WP): Sum of squared deviations Within breeds.
SSD (AP\WG): Sum of squared deviations Among Breeds, Within Groups.
G: Number of groups in the structure.
P: Total number of breeds.
N: Total number of gene copies for haplogroup data.

Table 2.7 General AMOVA table taken from Arlequin package program (Excoffier *et al.*, 2006).

Source of Variation	Degrees of Freedom	Sum of Squares (SSD)	Expected Mean Squares
Among Groups	G-1	SSD (AG)	$n''\sigma_a^2 + n'\sigma_b^2 + \sigma_c^2$
Among Populations / Within Groups	P-G	SSD (AP/WG)	$n\sigma_b^2 + \sigma_c^2$
Within Populations	N-P	SSD (WP)	σ_c^2
Total	**N-1**	**SSD (T)**	σ_T^2

It this study Analysis of Molecular Variance (AMOVA) is used to observe how the genetic variation is partitioned within and among the constructed two groups and four groups from the breeds based on NJ tree and PCA. The significance results were evaluated by 1000 random permutations. Calculations were performed using Arlequin version 3.1 (Excoffier *et al.*, 2006).

CHAPTER 3

RESULTS

3.1 Results of the Laboratory Experiments

3.1.1 DNA Extraction and PCR

After isolation of the DNA by the methods described in materials and methods, DNA samples were checked by the agarose gel electrophoresis, to control the presence, quality and concentration of the isolated DNA. Figure 3.1 represents an example for the agarose gel image of the DNA check, which was scanned under UV light. Concentrations of the DNA samples were observed approximately by comparing sample DNA bands with the bands in the standard size markers (HindIII size marker), visually. Samples which have no DNA or very low amount of DNA were not used for PCR (DNA samples 14, 34, 35, 39, 41, and 43 in Figure 3.1), new DNA samples were isolated from other blood samples which belong to the same individual. DNA samples which have high concentrations were diluted as described in materials and methods (DNA samples 2, 31, 38, 40, 44, 46 and 56 in Figure 3.1), to adjust their concentrations appropriate for PCR and checked by agarose gel electrophoresis again.

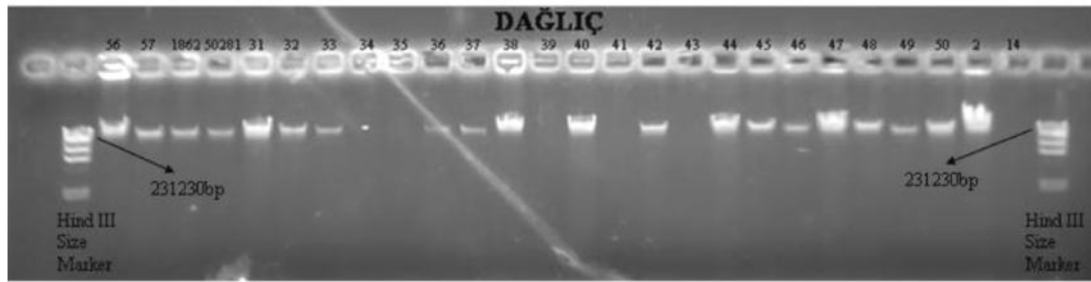


Figure 3.1 Gel image of total DNA extracts after isolation. Numbers shown above the wells are the DNA sample numbers and name on the top of the image is the name of the breed.

After checking and adjusting the concentrations, DNA samples were used for amplification of the partial ND4 region and complete CR region of mtDNA by procedures previously described in materials and methods. Presence and the amount of PCR products were checked by agarose gel electrophoresis too. Figure 3.2 represents an example for the agarose gel image of the amplification results of approximately (varies because of tandem repeats) 1260 bp long mtDNA CR region and Figure 3.3 represents an example for the agarose gel image of the amplification results of 280 bp long mtDNA ND4 region which were scanned under UV light. Samples which were not amplified or gave a weak amplification products were not used for further analysis, these samples were amplified again. Also during amplification, a negative control (as shown in Figure 3.2) was used that contained PCR mixture without DNA, to be able to detect if there is a contamination in solutions that is used for PCR reaction. Negative control is expected to show no bands in agarose gel electrophoresis.

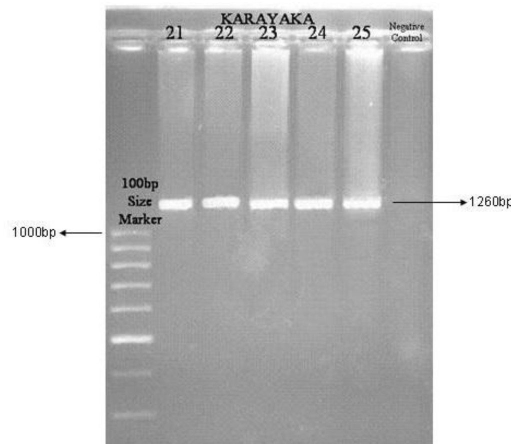


Figure 3.2 Gel image of mtDNA CR region PCR amplification products. Numbers shown above the wells are the DNA sample numbers and name on the top of the image is the name of the breed.

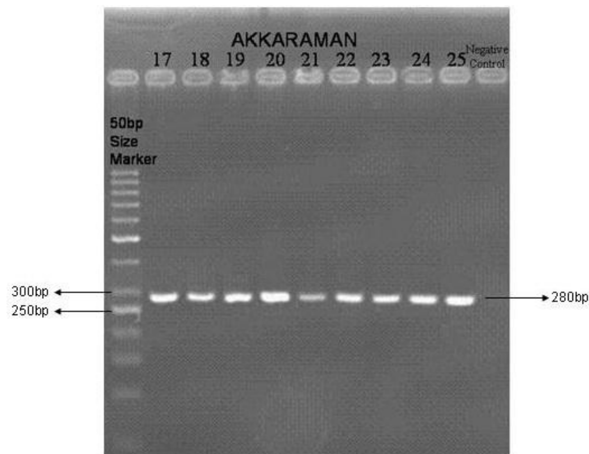
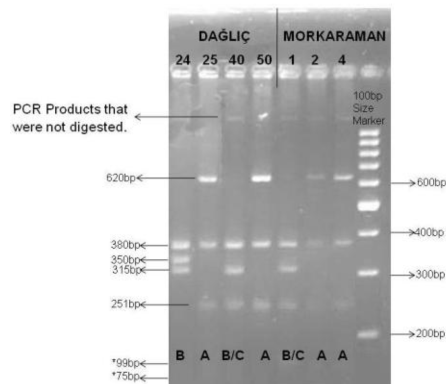


Figure 3.3 Gel image of partial mtDNA ND4 region PCR amplification products. Numbers shown above the wells are the DNA sample numbers and name on the top of the image is the name of the breed.

3.1.2 Genotyping

3.1.2.1 Genotyping of mtDNA CR Region by RFLP

Haplogroup analysis of mtDNA CR region was carried out by RFLP method. After amplification, products were digested with *NsiI* restriction enzyme as described in previous chapter and digestion products were checked by agarose gel electrophoresis. Digestion products revealed three distinct profiles. Figure 3.4 represents an example for the agarose gel image of the restriction profiles of samples which were scanned under UV light. Profiles which composed of 620 bp, 380 bp and 251 bp long bands were genotyped as HPG-A, profiles which composed of 380bp, 350bp and 315bp long bands were genotyped as HPG-B and profiles which composed of 380bp, 315bp and 251bp long bands were genotyped as HPG-B/HPG-C. Because of a G-A transition in the restriction site, this digestion profile cannot be separated in a conclusive manner (Bruford and Townsend, 2006; Koban *et al*., 2008).



*Indicated fragments are not seen.

Figure 3.4 Restriction enzyme profiles of mtDNA CR analysis. Numbers shown above the wells are the DNA sample numbers, names on the top of the image are the name of the breeds and letters at the bottom of the patterns are haplogroups that are represented by digestion profiles above them. Samples that belong to two different breeds are separated by a vertical line. 99bp and 75bp fragments are not seen because they can not be visualized by agarose gel electrophoresis and also, they are not diagnostic.

Akkaraman individual number 46 and Dağlıç individual number 56 showed different patterns than expected. These patterns were called unidentified band patterns and symbolized as “?”. Figure 3.5 represents these unidentified band patterns.

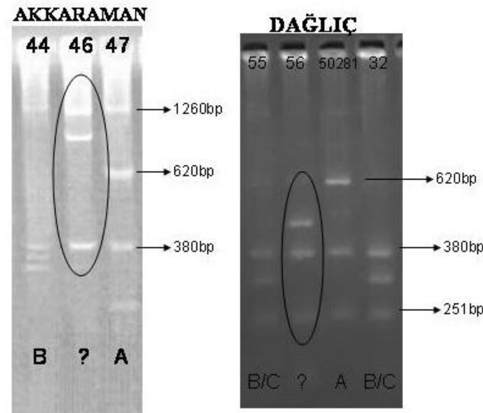


Figure 3.5 Restriction enzyme profiles of unidentified band patterns. Numbers shown above the wells are the DNA sample numbers, names on the top of the image are the name of the breeds and letters at the bottom of the patterns are haplogroups that are represented by digestion profiles above them. Unidentified band patterns were symbolized by ‘‘?’’ and shown in an ellipse.

The haplogroup of each individual which was determined by mtDNA CR region analysis is given in Appendix B.

3.1.2.2 Genotyping of mtDNA ND4 Region by SSCP

Haplogroup analysis of mtDNA ND4 region was carried out by SSCP method. (Guo *et al*., 2005) After amplification, products were denaturated to separate the double strands of the DNA, by procedure previously described in materials and methods. New conformations of the strands were checked by PAGE (Polyacrylamide Gel Electrophoresis) for genotyping. Figure 3.6 represents an example for the PAGE

image of conformational polymorphism profiles which were visualized by silver staining. Patterns were decided by using samples with known haplogrups as standards and comparing their patterns with unknown sample patterns.

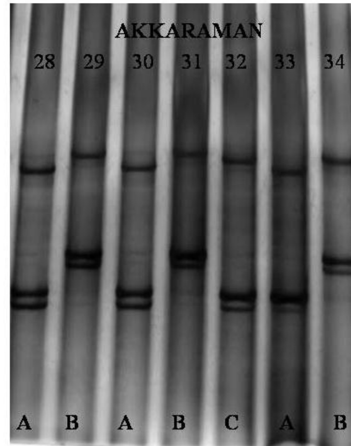


Figure 3.6 Conformational polymorphism profiles of mtDNA ND4 analysis. Numbers shown above the patters are the DNA sample numbers, names on the top of the image are the name of the breeds and letters at the bottom of the patterns are haplogroups that are represented by conformational polymorphism profiles above them.

Akkaraman individual number 13, 38 and Dağlıç individual number 9, 32, 55, 24 showed different patterns than expected. These patterns were called unidentified band patterns and symbolized by “?”. Figure 3.7 represents these unidentified band patterns.

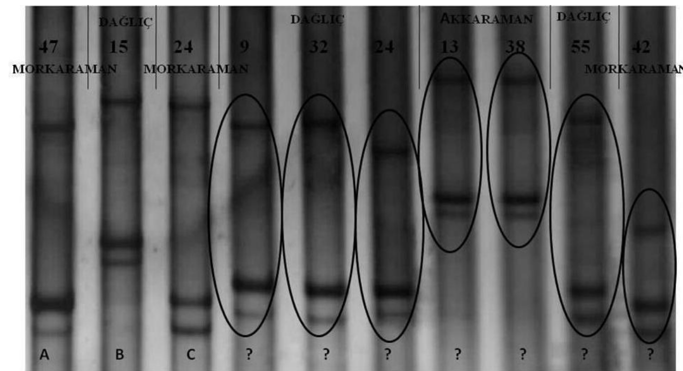


Figure 3.7 Conformational polymorphism profiles of unidentified band patterns. Numbers shown above the patterns are the DNA sample numbers, names on the top of the image are the name of the breeds and letters at the bottom of the patterns are haplogroups that are represented by conformational polymorphism profiles above them. Samples that belong to two different breeds are separated by a vertical line. Unidentified band patterns were symbolized by "?" and shown in an ellipse

Individuals, which are found to be carrying HPG-D (econogene sample) and HPG-E by sequencing in Koban *et al.*'s (2008) study, were analyzed with SSCP and RFLP methods to determine if they can be differentiated by these methods. Also their band patterns are compared with unidentified band patterns. Figure 3.8 represents the band patterns of these individuals and unidentified individuals. By SSCP, HPG-D gave similar band pattern with HPG-A and HPG-E gave similar band pattern with HPG-B. By RFLP, HPG-D gave a unique band pattern and HPG-E gave similar band pattern with HPG-B/HPG-C. Band pattern of the HPG-D is not similar with any of the unidentified band patterns of the RFLP. As a result, unidentified band patterns can not be identified yet and it is observed that RFLP analysis may identify HPG-D but SSCP analysis can not identify HPG-D and HPG-E.

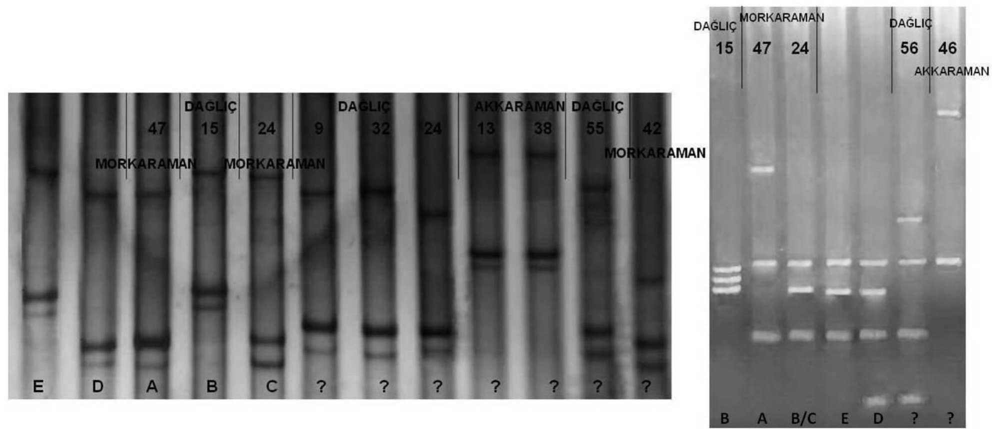


Figure 3.8 Conformational and restriction enzyme polymorphism profiles of HPG-D and HPG-E together with unidentified band patterns of SSCP and RFLP method. Numbers shown above the wells and the patters are the DNA sample numbers, names on the top of the images are the name of the breeds and letters at the bottom of the patterns are haplogroups that are represented by digestion and conformational polymorphism profiles above them. HPG-D and HPG-E samples are from Koban *et al.*'s (2008) study. Samples that belong to two different breeds are separated by a vertical line.

The haplogroup of each individual which was determined by mtDNA ND4 region analysis is given in Appendix B.

3.2 Results of Statistical Analyses

3.2.1 Comparison of Results Obtained by Two Methods

All breeds, total of 314 individuals, were analyzed with both of the methods (RFLP and SSCP) for determining the haplogroup compositions. Results that were obtained from two different methods were compared to decide the goodness of fit of results obtained from the two methods. As an outcome, 50 out of 50 individuals gave same results for Karayaka (100%), 47 out of 50 individuals gave same results for Akkaraman (94%), 49 out of 50 individuals gave same results for Gökçeada (98%), 45 out of 50 individuals gave same results for Dağlıç (90%), 49 out of 50 individuals

gave same results for Morkaraman (98%), 20 out of 20 individuals gave same results for Herik-Iran (100%) and 44 out of 44 individuals gave same results for Azerbaijan (100%). For the total 304 out of 314 individuals were gave same haplogroup with two methods with 96.82% total goodness of fit. Table 3.1 represents the samples which showed different or unidentified band patters with two different methods.

Table 3.1 Ambiguous and different haplogroups based on RFLP and SSCP screening.

BREEDS	Individual Number	SSCP Results	RFLP Results
AKKARAMAN	13	?	HPG-B/HPG-C
AKKARAMAN	38	?	HPG-B/HPG-C
AKKARAMAN	46	HPG-A	?
GÖKÇEADA	38	HPG-C	HPG-B
DAĞLIÇ	9	?	HPG-B/HPG-C
DAĞLIÇ	24	?	HPG-B
DAĞLIÇ	32	?	HPG-B/HPG-C
DAĞLIÇ	55	?	HPG-B/HPG-C
DAĞLIÇ	56	HPG-A	?
MORKARAMAN	42	?	HPG-A

Akkaraman individual number 13-38 and Dağlıç individual number 9-32-55 have same unidentified SSCP band patterns. (Figure 3.7)

3.2.2 Frequency of Haplogroups

Haplogroup frequencies of all studied breeds were calculated to observe genetic structure of the populations in relation to mtDNA haplogroup composition. While calculating the frequency of the haplogroups for the breeds, 10 individuals which gave different results with two analyses were excluded from the calculation and individuals that gave HPG-B/HPG-C with mtDNA CR region analysis were included

in accordance with the results based on mtDNA ND4 analysis. Table 3.2 represents frequency of the haplogroups with the haplogroup diversities for the breeds and Figure 3.9 represents the pie charts of the breeds constructed by haplogroup frequencies, overlaid on map of Turkey.

Table 3.2 Haplogroup frequencies of seven studied breeds.

BREEDS	HAPLOGROUP FREQUENCIES			SAMPLE SIZE (N)
	HPG-A	HPG-B	HPG-C	
KARAYAKA	0.2000	0.7200	0.0800	50
AKKARAMAN	0.2553	0.5745	0.1702	47
GÖKÇEADA	0.3265	0.6735	0.0000	49
DAĞLIÇ	0.1111	0.7111	0.1778	45
MORKARAMAN	0.3469	0.3469	0.3962	49
HERIK-IRAN	0.3000	0.3500	0.3500	20
AZERBAIJAN	0.5455	0.3864	0.0681	44

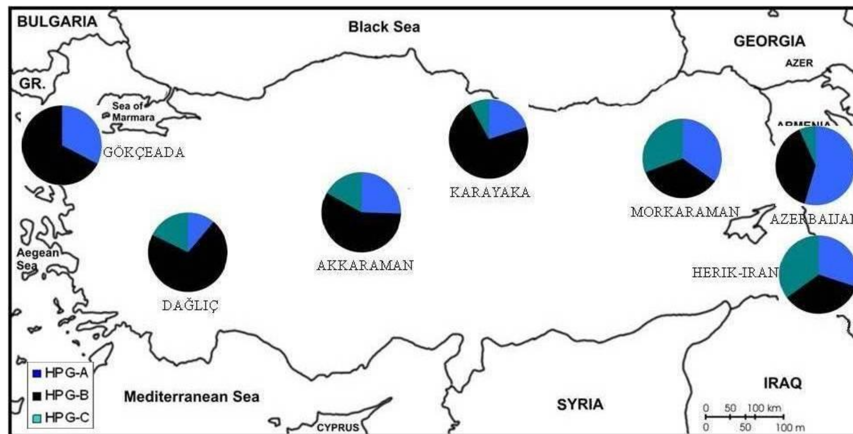


Figure 3.9 Distribution of mtDNA haplogroups in 7 sheep breeds from Gökçeada to Iran region.

All haplogroups are present in all of the breeds except in Gökçeada, the highest HPG-A frequency is observed in Azerbaijan. Karayaka and Dağlıç exhibited the maximum frequencies of HPGB. HPG-C frequency peaked in Herik-Iran.

3.2.3 Pairwise F_{ST} Values

Degree of genetic differentiation based on mtDNA haplogroups, among the studied breeds are determined with pairwise F_{ST} index, by Arlequin package program (Excoffier *et al*., 2006). Table 3.3 represents the pairwise F_{ST} values among breeds with their significance levels

.

.

Table 3.3 Pairwise F_{ST} values of the seven studied breeds based on mtDNA haplogroups.

	KARAYAKA	AKKARAMAN	GÖKÇEADA	DAĞLIÇ	MORKARAMAN	HERIK-IRAN	AZERBAIJAN
KARAYAKA	0.00000						
AKKARAMAN	0.01075	0.00000					
GÖKÇEADA	0.00726	0.02109	0.00000				
DAĞLIÇ	-0.00171	0.01558	0.06196*	0.00000			
MORKARAMAN	0.14443***	0.03947	0.13607**	0.13552***	0.00000		
HERIK-IRAN	0.15121**	0.03154	0.15526**	0.12860*	-0.03326	0.00000	
AZERBAİJAN	0.17352***	0.08338*	0.10143*	0.21718***	0.05363*	0.07441*	0.00000

$*p<0.05$ $**p<0.01$ $***P<0.001$

According to the pairwise F_{ST} values, Morkaraman, Herik-Iran and Azerbaijan populations are significantly differentiated from other populations in most of the pariwise comparisons. Especially Azerbaijan population is significantly different from other Turkish populations in varying α level.

3.2.4 Nei's Genetic Distance (D_A) Between The Breeds

Nei's D_A genetic distances for the seven populations were calculated with the GENDIST program in PHYLIP package software (Felsenstein, 1993), based on the mtDNA haplogroup frequencies, to estimate genetic relationship between populations. Table 3.4 represents the Nei's D genetic distance matrix of the studied breeds.

Table 3.4 Pairwise Nei's D_A genetic distance values of the seven studied breeds based on mtDNA haplogroups.

	KARAYAKA	AKKARAMAN	GÖKÇEADA	DAĞLIÇ	MORKARAMAN	HERIK-IRAN	AZERBAIJAN
KARAYAKA	0.000000						
AKKARAMAN	0.023062	0.000000					
GÖKÇEADA	0.022069	0.035918	0.000000				
DAĞLIÇ	0.015778	0.033056	0.074209	0.000000			
MORKARAMAN	0.234855	0.102458	0.221340	0.232883	0.000000		
HERIK-IRAN	0.246355	0.111384	0.261049	0.219774	0.005185	0.000000	
AZERBAİJAN	0.251349	0.160241	0.137444	0.360201	0.121512	0.186567	0.00000

According to these genetic distances most diverged breeds are Dağlıç and Azerbaijan breeds while most genetically close breeds are Iran and Morkaraman.

3.2.5 Neighbor Joining (NJ) Tree of the Breeds

NJ trees were constructed using Nei's D_A genetic distances and pairwise F_{ST} values, to visualize the relationship of the populations. All trees were drawn with MEGA package program (Tamura, 2008). Figure 3.10 represents the NJ tree constructed by Nei's D_A values and Figure 3.11 represents NJ tree constructed by F_{ST} values,

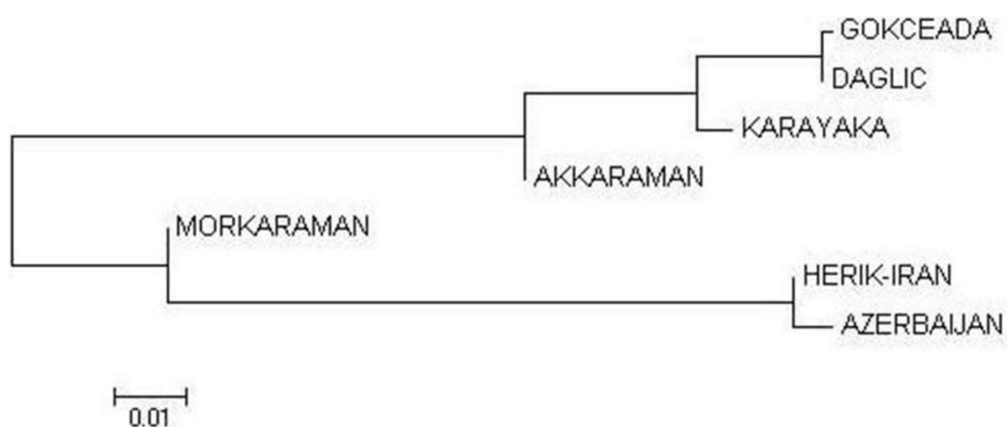


Figure 3.10 NJ tree that was constructed according to the Nei's D_A genetic distances between seven populations.

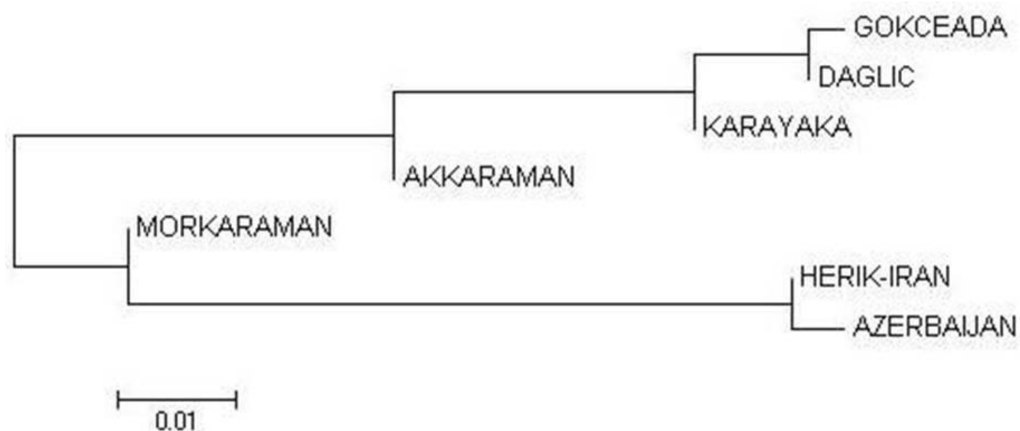


Figure 3.11 NJ tree that was constructed according to the pairwise F_{ST} values between seven populations.

In both neighbor joining trees there are two distinct branches. First branch is composed of population sampled from the west and the middle Anatolia (Karayaka, Akkaraman, Gökçeada and Dağlıç) and second group is sampled from east of Anatolia (Morkaraman, Herik-Iran and Azerbaijan).

3.2.6 Analysis of Geographical Distribution of the Breeds

3.2.6.1 Spearman Rank Correlation Test

Two Spearman rank correlation tests were conducted to test if HPG-A and HPG-C frequencies were decreasing gradually with the increase in distance. The point of origin for HPG-A correlation test was chosen as Azerbaijan breed which have highest HPG-A frequency and the point of origin for the HPG-C correlation test was chosen as Herik-Iran bred which have highest HPG-C. Table 3.5 represents r_s values of those correlation tests with their significance levels.

Table 3.5 Conducted r_s values for Spearman Rank Correlation tests.

	r_s Values
HPG-A Spearman Rank Correlation Test	-0.600000
HPG-C Spearman Rank Correlation Test	-0.771429*

*p<0.05 **p<0.01 ***p<0.001

There is a significant negative correlation (p<0.05) between HPG-C frequency and the distances of the populations from point of origin (Iran), suggesting frequency of HPG-C is gradually decreasing from the point of origin, Iran. Negative correlation between HPG-A frequency and the distances of the populations from the point of origin (Azerbaijan) is not significant (p>0.05) because of the Gökçeada population. Although this population has the highest distance from the Azerbaijan, it has a higher HPC-A frequency compared to other breeds. Spearman rank correlation test for HPG-A is repeated by excluding the Gökçeada population. Table 3.6 represents r_s value of this correlation test with its significance level.

Table 3.6 Spearman Rank Correlation r_s coefficient when Gökçeada population was excluded from the analysis.

	r_s Value
HPG-A Spearman Rank Correlation Test (Without Gökçeada)	-0.900000**

*p<0.05 **p<0.01 ***p<0.001

After excluding the island population (Gökçeada), negative correlation between HPG-A frequency and the distances of the populations from point of origin (Azerbaijan), became significant (p<0.01) suggesting frequency of HPG-A is a gradually decreasing from the point of origin, Azerbaijan.

Ranks of the X and Y components of three tests with their di values are given in Appendix C.

3.2.6.2 Delaunay Network Analysis

Delaunay Network Analysis is conducted with the data obtained from the studied breeds by using Nei's D_A genetic distances and pairwise F_{ST} values to construct the genetic barriers on map. Genetic barriers were formed with thiessian triangles by shortest distance criteria. Figure 3.12 represents the Delaunay Network Analysis formed with Nei's D_A values and Figure 3.13 represents Delaunay Network Analysis formed with pairwise F_{ST} values.

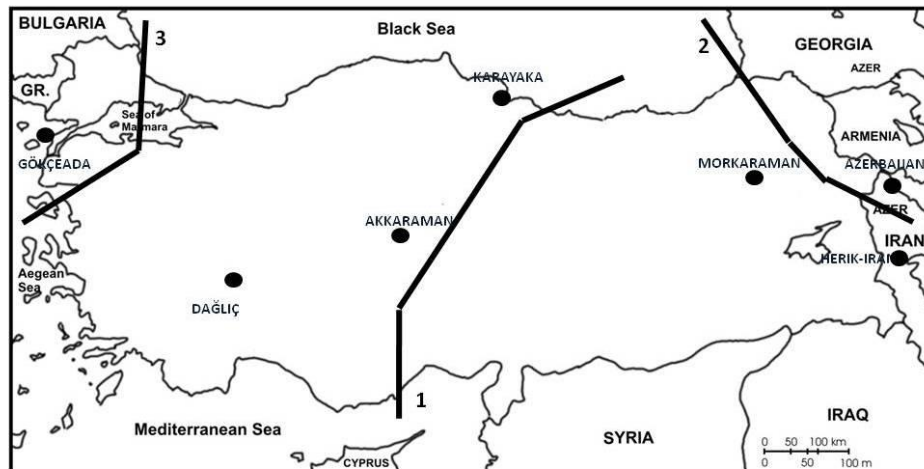


Figure 3.12 Delaunay Network Approach performed by using Nei's D_A values. Dots represent the locations of the sampling sites of the breeds. Black lines are the constructed genetic barriers and numbers are the priority of the barriers near them.

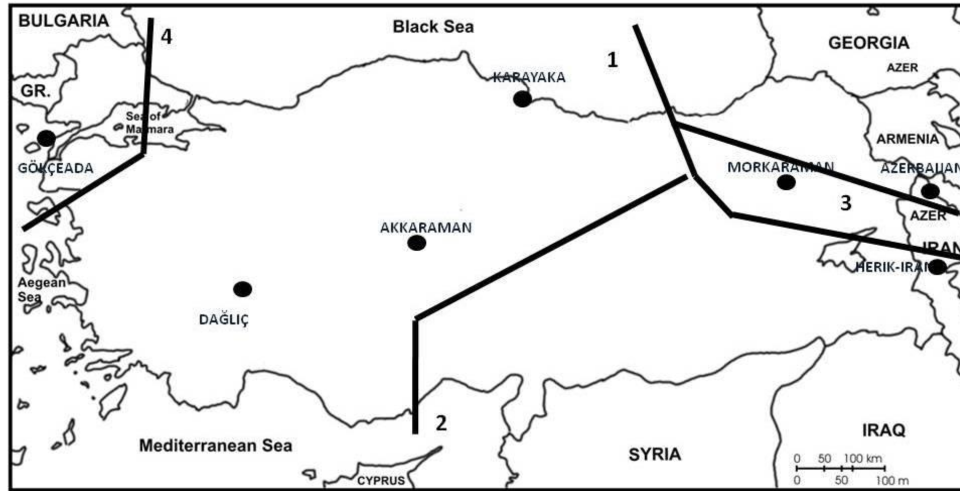


Figure 3.13 Delaunay Network Approach performed by using pairwise F_{ST} values. Dots represent the locations of the sampling sites of the breeds. Black lines are the constructed genetic barriers and numbers are the priority of the barriers near them.

Priorities of the genetic barriers are as follows; In the first map, the first barrier separates breeds as western breeds (Karayaka, Akkaraman, Dağlıç, Gökçeada) and eastern breeds (Morkaraman, Azerbaijan and Herik-Iran). Second barrier separates Azerbaijan from rest. Finally third barrier separates Gökçeada from rest of the breeds. In the second map, first barrier separates Azerbaijan and Morkaraman from the rest, second barrier separates breeds as western breeds (Karayaka, Akkaraman, Dağlıç, Gökçeada) and eastern breeds (Morkaraman, Azerbaijan and Herik-Iran). Third barrier separates Azerbaijan from Morkaraman and the fourth barrier separates Gökçeada from rest.

3.2.6.3 Principal Component Analysis (PCA)

Relative relatedness of populations was observed based on haplogroup frequencies by performing Principle Component Analysis (PCA) in 2 dimensional spaces. The PCA was constructed with NTSYS package program (Rohlf, 2000). Figure 3.14 represents two principal components of the PCA constructed based on haplogroup

frequencies. The equations of the two components (C1 and C2) of the PCA were as follows;

PC1=0.689 HGPA - 0.999 HPGB + 0.613 HPGC

PC2=0.724 HPGA + 0.015 HPGB - 0.790 HPGC

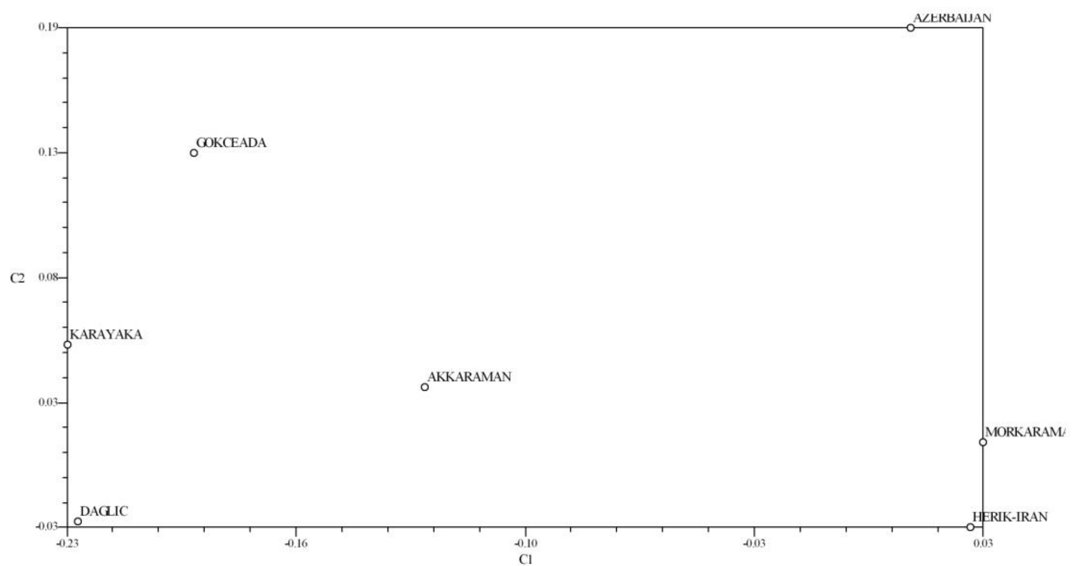


Figure 3.14 Principal component analysis of seven breeds in 2 dimensional model based on mtDNA haplogroup frequencies.

In this PCA, first component (C1) explains 61.6742% and the second component explains 38.3258% of the genetic variation. C1 and C2 components together can explain 100% of the genetic variation. C1 axis, heavily weighted by HPG-B, discriminates and separates populations into two groups as in NJ tree. First group is composed of breeds from the west Anatolia (Karayaka, Akkaraman, Gökçeada and Dağlıç) and second group is composed of breeds from east Anatolia (Morkaraman Herik-Iran and Azerbaijan). C2 is further resolving the populations based on their

HPG-A and HPG-C frequencies. In this component Azerbaijan and Gökçeada breeds were separated from other populations.

3.2.6.4 Analysis of Molecular Variance (AMOVA)

Five different grouping were made according to the results of neighbor joining trees, PCA and Delaunay Network Analysis that formed with Nei's D_A genetic distances. These groups were analyzed with AMOVA to observe, how the amount of genetic variation within and among groups is partitioned, with Arlequin package program (Excoffier *et al.*, 2006). Table 3.7 represents AMOVA results constructed by using two groups according to the PCA-C1 dimension and the neighbor joining trees. These two groups used in AMOVA analysis were as follows;

Group 1 (West): Karayaka
Akkaraman
Gökçeada
Dağlıç

Group 2 (East): Morkaraman
Herik-Iran
Azerbaijan

Table 3.7 AMOVA table for the two major groups determined based on PCA-C1 and the neighbor joining trees.

Source of Variation	**df**	**Variance Components**	**Percentage of Variation**	**Fixation Index**
Among Groups	1	0.03196	10.32	0.10319* ϕ_{CT}
Among Breeds Within Groups	5	0.00722	2.33	0.02600* ϕ_{SC}
Within Breeds	297	0.27052	87.35	0.12651*** ϕ_{ST}
Total	**303**	**0.30970**		

*p<0.05 **p<0.01 ***p<0.001

Table 3.8 represents AMOVA results constructed by using three groups according to the PCA-C1 and PCA-C2 dimensions. These three groups used in AMOVA analysis were as follows;

Group 1: Gökçeada
Azerbaijan

Group 2: Dağlıç
Karayaka
Akkaraman

Group 3: Morkaraman
Herik-Iran

Table 3.8 AMOVA table for the three groups determined based on PCA-C1 and PCA-C2 dimensions.

Source of Variation	df	Variance Components	Percentage of Variation	Fixation Index
Among Groups	2	0.02352	7.81	0.07808* ϕ_{CT}
Among Breeds Within Groups	4	0.00717	2.38	0.02580*** ϕ_{SC}
Within Breeds	297	0.27052	89.81	0.10187*** ϕ_{ST}
Total	**303**	**0,30121**		

*p<0.05 **p<0.01 ***p<0.001

Table 3.9 AMOVA results constructed by using three groups according to the first and third genetic barriers in Delaunay Network Analysis that formed with Nei's D_A genetic distances. These three groups used in AMOVA analysis were as follows;

Group 1: Gökçeada

Group 2: Dağlıç
Karayaka
Akkaraman

Group 3: Morkaraman
Herik-Iran
Azerbaijan

Table 3.9 AMOVA table for the three groups determined based on first and third genetic barriers in Delaunay Network Analysis that formed with Nei's D_A genetic distances.

Source of Variation	df	Variance Components	Percentage of Variation	Fixation Index
Among Groups	2	0.02483	8.22	0.08218* ϕ_{CT}
Among Breeds Within Groups	4	0.00674	2.23	0.02432* ϕ_{SC}
Within Breeds	297	0.27052	89.55	0.10450*** ϕ_{ST}
Total	**303**	**0.30290**		

*p<0.05 **p<0.01 ***p<0.001

Table 3.10 represents AMOVA results constructed by using three groups according to the first and second genetic barriers in Delaunay Network Analysis that formed with Nei's D_A genetic distances. These three groups used in AMOVA analysis were as follows;

Group 1: Gökçeada
Dağlıç
Karayaka
Akkaraman

Group 2: Morkaraman
Herik-Iran

Group 3: Azerbaijan

Table 3.10 AMOVA table for the three groups determined based on first and second genetic barriers in Delaunay Network Analysis that formed with Nei's D_A genetic distances.

Source of Variation	df	Variance Components	Percentage of Variation	Fixation Index
Among Groups	2	0.0369	11.75	0.11750** ϕ_{CT}
Among Breeds Within Groups	4	0.00203	-0.66	0.00746 ϕ_{SC}
Within Breeds	297	0.27052	89.59	0.12409*** ϕ_{ST}
Total	303	0.30885		

*p<0.05 **p<0.01 ***p<0.001

Table 3.11 represents AMOVA results constructed by using four groups according to the Delaunay Network Analysis that formed with Nei's D_A genetic distances. These four groups used in AMOVA analysis were as follows;

Group 1: Gökçeada

Group 2: Dağlıç
Karayaka
Akkaraman

Group 3: Morkaraman
Herik-Iran

Group 4: Azerbaijan

Table 3.11 AMOVA table for the four groups determined based on Delaunay Network Analysis that formed with Nei's D_A genetic distances.

Source of Variation	df	Variance Components	Percentage of Variation	Fixation Index
Among Groups	3	0.03164	10.49	0.10495*** ϕ_{CT}
Among Breeds Within Groups	3	-0.00066	-0.22	-0.00244 ϕ_{SC}
Within Breeds	297	0.27052	89.72	0.10277*** ϕ_{ST}
Total	**303**	**0.30151**		

*p<0.05 **p<0.01 ***p<0.001

In all of the AMOVA tables the highest portion of the total variation was within the breeds, the lowest portion of the total variation was among breeds within groups and difference within breeds were highly significant (p<0.001). According to the first and third AMOVA table, differences among groups and among breeds within groups were all significant (p<0.05). According to the second AMOVA table differences among groups and among breeds within groups were all were all significant (p<0.05) where the difference among groups within breeds were highly significant (p<0.001). According to the fourth and fifth AMOVA table the portion of the total variation among breeds within groups was null as implied by the negative value and difference among breeds within groups is more significant in fifth AMOVA table (p<0.001) that the fourth AMOVA table (p<0.01). Hence, grouping considered by the fifth AMOVA revealed homogenous groups such that at least one of them is significantly (p<0.001) different than the others.

CHAPTER 4

DISCUSSION

There has been many studies covering genetics of domestic sheep (e.g. Hiendleder *et al.*, 2002; Pereira *et al.*, 2006; Wang *et al.*, 2007; Bruford and Townsend, 2006; Peter *et al.*, 2007, Koban *et al.*, 2008). Data obtained from these studies are used mainly in resolving evolutionary history of sheep. Understanding of evolutionary history of sheep, in turn, will help to prioritize the sheep breeds for their conservation. For example the most ancient and the least admixed breed may have the prime importance to be conserved or among the breeds having similar evolutionary history, only one of them may be considered to be conserved (Koban *et al.*, 2008). Similarly evolutionary history may indicate how differentiated samples of the breeds are in state farms. Then, they can be managed such that they will not be diverged from the gene pool of the breed. Also evolutionary history of the breeds provides information about the uniqueness of the breed. This information can be used to find the best candidate breed to start searching new alleles related with productivity or resistance traits.

First step of making accurate genetic characterization is to collect a sample which characterizes the whole gene pool of the breed. For this purpose sampling must be composed of unrelated individuals and must be coming from the whole distribution area of the breed. Sampling from small geographical locations or from a single herd can cause bias in representing the breed. Because, in that case, some genotypes may be over represented, yet, some others will not be represented at all (Koban *et al.*, 2008). Inaccurate genetic characterization, for example, in conservation studies can

lead wrong priorization of the breeds under consideration and hence, irreversible loss of genes and waist of effort and money may result.

Choosing the proper marker and test method for genetic characterization is also important. Different markers are used to answer different genetic questions. Best marker must be selected by considering properties of the marker and the objective of the study (Allendorf and Luikart, 2007). If tracking the evolutionary history of the domestic sheep is the main objective of the study, mtDNA is the type of DNA to be used (Bruford *et al.*, 2003; Bruford, 2004). mtDNA sequencing is preferred in many studies, but it is relatively expensive and it demands an expensive automated machine. Method for analyzing the chosen marker must be efficient, cheap and easy to apply. There were two suggested methods in searching the polymorphism of mtDNA of sheep; analysis of ND4 region with SSCP as described in Guo *et al.* (2005) and analysis of CR region with RFLP as described in Bruford and Townsend (2006). The latter method was used by Koban *et al.* (2008) to analyze Turkish sheep mtDNA. The latter method was used by Koban *et al.* (2008) to analyze Turkish sheep mtDNA. In this study HPG-D is identified by sequencing and HPG-E gave HPG-B/HPG-C band pattern with RFLP. Hence it is also identified by sequencing.

4.1 Consistency of Results Obtained by Two Methods

By the present study, first compatibility of the SSCP and RFLP approaches for the mtDNA haplogroup determination was carried out. They are found to be highly compatible: results were 96% consistent. The differences are mainly due to the unidentified band patterns. Unidentified band patters revealed by RFLP method can arise due to minor mutations in the restriction sites or tandem repeats which can exist in vary number in CR region as explained in introduction of the present manuscript. Similarly, unidentified band patterns revealed by SSCP method might be originated because of mutations in the sequence of ND4 region. In the present study, in relation to identifying the major haplogroups (HPG-A, B and C), SSCP was found to be more

efficient in the sense that higher number of individuals can be screened per unit effort. Furthermore, SSCP can reveal three haplogroups: A, B and C, but RFLP can differentiate A, B or B/C haplogroups. Hence, again SSCP can be preferred. Yet, rare haplogroups: HPG-D and HPG-E identified by Tapio *et al.* (2006); Meadows *et al.* (2007); Koban *et al.* (2008), gave the same band patterns as HPG-A and HPG-B respectively by SSCP method. HPG- E can not be detected by RFLP but one HPG-D individual can be differentiated. To detect HPG-D and HPG-E, sequencing of mtDNA control region sequencing was used before (Tapio *et al.*, 2006; Meadows *et al.*, 2007; Koban *et al.*, 2008). Additionally, results of the present study suggests that individuals exhibiting HPG-B by SSCP may be further tested by RFLP and those revealing HPG-B/C by the second method can be suspected as an individual possessing HPG-E. This suggestion must be verified by extensive sequencing. RFLP analysis must be also repeated with different HPG-D's to verify that RFLP can separate HPG-D.

Also it is revealed that unidentified banding patterns observed by SSCP and RFLP are not HPG-D and HPG-E. Still, these unidentified bands might be representing a new candidate haplogroup H1540 that is uniquely identified by Guo *et al.* (2005) or a haplogroup that is not identified before. Especially, Akkaraman individual number 13, number 38, Dağlıç individual number 9, number 32 and number 55 are good candidates of revealing a new HPG, because very similar unidentified band patterns were obtained by SSCP method for these individuals and they all give HPG-B/HPG-C with RFLP method. This can hardly be because of some minor random mutation.

4.2 Distribution of HPG Frequencies in Different Studies

Results of the present study were compared with those of the previous studies (Pedrosa *et al.*, 2005; Meadows *et al.*, 2007; and Koban *et al.*, 2008) to test the robustness of the HPG frequency results under different samplings. Breeds that are common in current study and Pedrosa *et al.*'s (2005); Meadows *et al.*'s (2007) and

Koban *et al.*'s (2008) were tested with X^2 test of homogeneity by Minitab package program (Minitab Inc., State College, PA, USA), to observe if the samples from present study and samples from those studies can be excepted as the samples of the same breeds.

Three breeds (Karayaka, Akkaraman, Morkaraman) in Pedrosa *et al.*'s (2005) study, two breeds (Karayaka, Morkaraman) in Meadows *et al.*'s (2007) study and four breeds (Karayaka, Akkaraman, Dağlıç, Morkaraman) in Koban *et al.*'s (2008) study are found to be common when they are compared with the breeds of the current study. Figure 4.1 represents comparison of frequency data as shown by pie charts and Table 4.1 represents computed X^2 values of the breeds with their significance level.

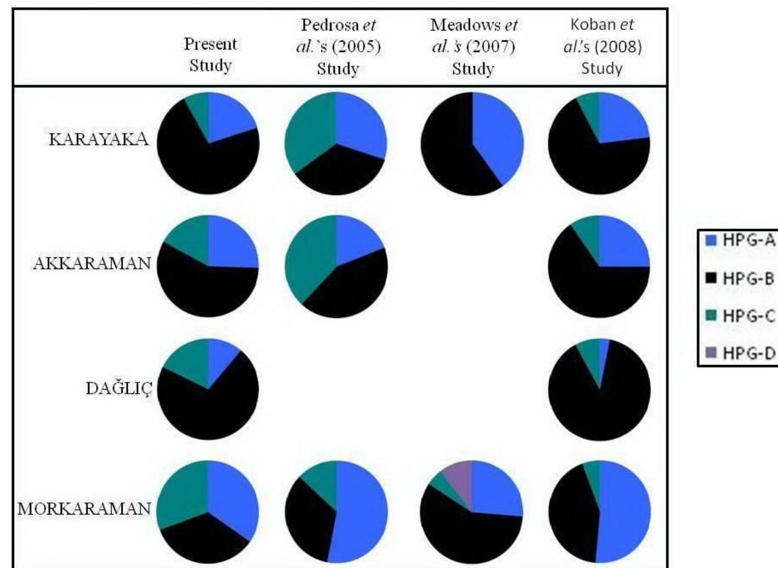


Figure 4.1 Comparison of the data in the present study and other studies (Pedrosa *et al.*, 2005; Meadows *et al.*, 2007; and Koban *et al.*, 2008), represented by pie charts.

Table 4.1 Computed X^2 values of the X^2 test of homogeneity for the data of present study and Pedrosa *et al.*'s (2005); Meadows *et al.*'s (2007); Koban *et al.'s* (2008) study.

	BREED	COMPUTED X^2 VALUE
Pedrosa *et al.*'s (2005) Study and Present Study	KARAYAKA	9,54625**
	AKKARAMAN	2,88606
	MORKARAMAN	2,31401
Meadows *et al.* 's (2007) Study and Present Study	KARAYAKA	3,31286
	MORKARAMAN	10,97360*
Koban *et al.*'s (2008) Study and Present Study	KARAYAKA	0,14188
	AKKARAMAN	1,28836
	DAĞLIÇ	5,66792
	MORKARAMAN	7,98004*

*p < 0.05, **p < 0.01, *** p< 0.001

When the data of the Koban *et al.*'s (2008) study and the present study is compared, out of four compared breeds, three of them revealed similar haplogroup compositions and with the fourth breed difference was minor ($p<0.05$). When the data of the Pedrosa *et al.*'s (2005) study and the present study is compared, out of three breeds two of them reveal similar haplogroup compositions and the third breed have a significant difference ($p<0.01$). When the data of the Meadows *et al.*'s (2007) study and the data of the present study is compared, one of the two breeds are found to be similar and the difference of the other breed was minor ($p<0.05$). Sampling strategy was similar, between the present study and that of Koban *et al.'* s (2008). Samplings were independent, 2-3 individuals / herd were sampled, span of the breed was tried to be covered. Perhaps that is why relatively high degree of similar HPG composition was observed in these two studies but not between the present study and that of Pedrosa *et al.*'s (2005) and Meadows *et al.*'s (2007).

Observed and expected number of individuals for the X^2 test of homogeneity was given in Appendix D.

4.3 Interpretations of Statistical Analyses

All native breeds except Gökçeada and all foreign breeds are composed of all three major haplogroups. This is expected since Turkey is harboring the earliest domestication center and it is close to other domestication centers (Bruford *et al.*, 2003; Pedrosa *et al.*, 2005; Meadows *et al.* 2007; Koban *et al.*, 2008; Zeder, 2008) This diversity is signature of animal domestication in the Central Anatolia-Zagros Mountains region and expansion of domesticated animals by Neolithic expansion (Zeder, 2008).

Haplogroup B is the most common among native breeds (60.42%). Since it is also the most common haplogroup in Europe and since wild sheep was absent in Europe during the time of sheep domestication (Clutton-Brock, 1981) it can be assumed that first domestic sheep in Anatolia was carrying haplogroup B (Koban *et al.* 2008). Also haplogroup B is can be the sign of the earliest sheep domestication at the Southeastern Turkey, This hypothesis must be checked by ancient DNA studies from Nevali Çöri and Çayönü area (Peters *et al.*, 1999) which is claimed to be harboring the earliest sheep domestication area. Karayaka and Dağlıç breeds have highest frequency of haplogroup B as was also observed by Koban *et al.* (2008). For this reason once more Karayaka and Dağlıç seemed to be the closest relatives of first domestics, at least maternally.

When haplogroup compositions of the breeds from the neighbor countries are studied it is observed that HPG-A is the most common haplogroup in Azerbaijan sheep. Since HPG-A is the most common haplogroup that is found in Azerbaijan probably migration to Anatolia through the Azerbaijan introduced HPG-A to Anatolia as was previously suggested (Koban *et al.*, 2008).

Herik was an Anatolian breed which will also be studied in the context of TURKHAYGEN-1 project. Current data suggest that Herik is taken to Iran and in the

absence of breed isolation in Iran, Iranian sheep were intogressed. As a result of introgression HPG-C frequency increased. Similarly in Koban *et al.*'s (2008) study HPG-C frequency of Iraqi breed Hamdani was high. Rich HPG-C in Herik-Iran supports the introduction of HPG-C to Anatolia from south east Anatolia (Koban *et al.*, 2008).

These frequency differences between breeds is exhibiting a cline as was revealed by the Spearman Rank Correlation Test. While conducting this test for HPG-A Gökçeada population is found to harbor high frequency of HPG-A contrary to expectations. Because, HPG-A was the HPG of our eastern neighbor Azerbaijan. Yet, Gökçeada is an island population and island population may experience high degree of drift due to the isolation and presence of small population size. It is known that, genetic drift causes changes in allele (in this study haplogroup) frequency. Frequency of a certain haplogroup can increase or decrease from one generation to another, due to drift just by change (Allendorf and Luikart, 2007). So, high frequency of HPG-A can be due to the drift. So Spearman Rank Correlation Test was repeated by excluding the Gökçeada population. Results of the test shows Azerbaijan can be the entrance point for HPG-A and Iran can be the entrance point for HPG-C to the Anatolia. These data adds another evidence to the scenario that first domestic sheep that carry haplogroup A migrated to Anatolia though Azerbaijan and first domestic sheep that carry haplogroup C migrated to Anatolia through south east Anatolia and these domestication products were mixed with the products of sheep domestication study in Nevali Çöri and Çayönü (Peters *et al.*, 1999) which was the haplogroup B (Koban *et al.*, 2008).

In this study, both pairwise F_{ST} values and Nei's genetic distance values are used to observe genetic relation and differentiation among populations. Pairwise F_{ST} values are good estimators when measuring the amount of differentiation between two populations. Significance of this differentiation can also be measured. But pairwise F_{ST} values can be biased when there are high levels of within group variation because

total variation between populations can not be high when within group variation is high (Hendrick, 1999). Additionally, F_{ST} measures do not consider the degree of relatedness between alleles which makes it usually a poor estimator for conducting a phylogenetic tree. There are many genetic distance measures developed to be used in constructing phylogenetic trees. Nei's D_A genetic distance is one of them (Allendorf and Luikart, 2007).

When pairwise F_{ST} results in the study are observed, it can be seen that foreign breeds are significantly different from native breeds with the exception of Iran- Herik and Morkaraman. Furthermore, Morkaraman is found to be significantly different from other native breeds and it is geographically closest to the breeds from neighboring countries when all of the native breeds are considered. These results indicate the existence of possible subgroups among breeds according to their geographical locations. Also possible genetic barriers between breeds in East and West can be suggested. When Nei's D_A genetic distance values are considered, the most diverged breeds are Dağlıç and Azerbaijan which are the breeds of far west and far east respectively, while genetically closest breeds are Herik and Morkaraman which are both in the east. Therefore; a possible genetic differentiation, which seems to be correlated with the geographical locations of the breeds, can be observed also with Nei's D_A genetic distance values. NJ trees, Delaunay Network Analyses, PCA and AMOVA were used to have a better understanding about the groupings of the breeds based on their genetic similarities/dissimilarities.

NJ tree is constructed by considering unequal rate of evolution and length of the branches are proportional to the amount genetic change. Different branch lengths can be seen and difference between neighboring braches can be observed. (Jobling *et al.*, 2004; Allendorf and Luikart, 2007) NJ tree is especially useful for revealing the relationship between populations which evolves in different rates. (Allendorf and Luikart, 2007) Bootstrap values are used to evaluate the significance of the constructed tree. Bootstrap is conducted by resampling the data as sub data sets for

many times with replacement. As a result, different trees are constructed. Bootstrap values show how many times the same node is constructed from these replications (Jobling, 2004). Higher bootstrap values mean the separation of the branch is more definite. However, in the present study 3 HPG frequencies where only two of them are independent does not allow us to construct bootstrap values. Therefore, two independent measures of genetic relatedness, F_{ST} and D_A were used to construct NJ tree.

Two NJ trees constructed by F_{ST} and Nei's D_A genetic distance values indicated that there are two different geographical groups namely as western breeds; Gökçeada, Akkaraman, Karayaka, Dağlıç and eastern breeds; Morkaraman, Azerbaijan and Herik-Iran.

Phylogenetic trees display relationship only in one dimension. PCA is used to visualize relationship between populations in multiple dimensions. Populations were also separated as eastern and western breeds according to the first component of the PCA and in this analysis, by examining the weights on the principle components it can be seen that this separation is due the HPG-B. In the second axis, Azerbaijan and Gökçeada breeds are found to be more distant from other breeds and this is due to the HPG-C and HPG-A. Azerbaijan has higher HPG-A composition because HPG-A entered to Anatolia though Azerbaijan. Moreover there is a possibility that HPG-A in the Gökçeada population might be increased due to the drift, which make this population genetically closer to Azerbaijan population but they remain different because of the difference between HPG-B frequencies. Also these breeds are separated on this axis by the HPG-C because, Gökçeada breed has no HPG-C, due to the gradual diffusion of HPG-C from Iran to west and HPG-C is probably introduced to Azerbaijan in a lower rate than it was to Anatolian breeds. In PCA, these two axes totally explain 100% of the variation. Mostly PCA in two dimensions can not explain the whole variation, but in the current study PCA results were based on three

variables, HPG- A, B and C frequencies, therefore only 2 axes could be constructed and 100% of the variation can be displayed by these axes.

Delaunay Network Analysis is another method indicating the genetic barriers between the breeds. Results of Delaunay Network Analyses which was made by using Nei's D_A genetic distances also show possible grouping between populations in the east and the west. Genetic barrier that makes this separation is also the most important genetic barrier, which means separation between breeds as East and West is mostly significant. This separation was also observed with the NJ tree and the first axis of the PCA. So it can be said that these seven breeds under consideration are mostly differentiated as eastern breeds and western breeds. Furthermore Azerbaijan was separated from other eastern breeds with the second important genetic barrier and Gökçeada was separated from other west populations with the third important genetic barrier. These two additional groups are parallel to the groups revealed by the separation of the second axis in PCA. Additionally these two groups cannot be observed in phylogenetic trees. There is another barrier that separates Morkaraman from Herik-Iran in the Delaunay Network Analysis constructed with F_{ST} values. However, F_{ST} value between Morkaraman and Herik-Iran is found to be non-significant. This genetic barrier is less reliable because F_{ST} values can not be compared with each other based on their high or low numerical values. The level of difference is measured by permutation based significance test. However, Delaunay Network Analysis is based on comparison of distances with each other. Results indicate that constructing a Delaunay Network with F_{ST} values are less reliable and Nei's D_A genetic distances must be preferred.

As a summary, populations are separated in four groups according to the genetic barriers between geographical locations. Strongest differentiation is present in between two groups (eastern and western breeds) and this differentiation can be visualized in only one dimension. Therefore, these two groups can be observed by the NJ tree. Also Azerbaijan is further differentiated from other eastern breeds in

East and Gökçeada is further differentiated from other western breeds in West. These separations explain the total variation in the second dimension, so it can not be observed in phylogenetic tree. This is an example of using appropriate and sufficient statistical analyses for interpretation of the results. In this case, if only NJ tree was used to obtain relationship between populations, other groupings could be missed.

Results obtained by AMOVA are totally confirming and enhancing conclusions given above. When only two groups, eastern and western breeds, were subjected to AMOVA breeds within the groups were significantly different from each other indicating that there were more groupings within the groups. When AMOVA was employed according to the separation between both axis of PCA, as Group 1: Karayaka, Dağlıç, Akkaraman; Group 2: Morkaraman, Herik-Iran and Group 3: Azerbaijan, Gökçeada; highly significant difference ($p<0.001$) was found within groups and when AMOVA was employed by separating breeds as Group 1: Gökçeada; Group 2: Dağlıç, Akkaraman, Karayaka; and Group 3: Morkaraman and Herik-Iran; still significant difference ($p<0.05$) can be observed. However, when breeds are grouped as Group 1: Gökçeada, Karayaka, Dağlıç, Akkaraman; Group 2: Morkaraman, Herik-Iran and Group 3: Azerbaijan; and when 4 groups as revealed by Delaunay Network Analysis and supported by PCA were subjected to AMOVA analysis there was no significant difference between the group members but significance of the difference among groups of AMOVA ($p<0.05$) with three groups is lower than the significance of the difference among groups ($p<0.01$) of the AMOVA with four groups. This indicates that, although Gökçeada breed is less differentiated from other western breeds, it harbors still harbors some variation. This result is also consistent with the Delaunay Network Analysis because the genetic barrier that separates the Gökçeada from other western breeds has priority level three. Hence, four defined group were homogenous groups. AMOVA not only confirmed the previously identified groups but also strengthen the presence of these four groups by presenting the significance tests for various portioned molecular variation.

All these data supports the scenario of migration of sheep from east (HPG-A) and south east (HPG-C) to HPG-B harboring Anatolia and hence data and analysis contribute to the evolutionary history of the domestic sheep.

It is known that genetic diversity is decreasing with migration from the domestication centers (Bradley *et al.*, 1996; Loftus *et al.*, 1999). So during the dispersion of HPG-B to Europe and possibly to Asia from Turkey, some of the important genetic information must be lost. However, eastern native Turkish sheep breeds might still be harboring some variability which is absent in extend breeds. It is known that growth of urban life and agricultural strengthening causes a threat to the today’s domestic sheep breeds. (Zeder, 2008) According to the Longworth, (2005)’s study number of sheep in Turkey is decreased by 47% in the last two decades. Also wrong agricultural policies causes “genetic pollution” and inbreeding which in turn, causes reduction in the effective population size. “Genetic pollution” is introgression or hybridization with a certain breed with different breeds (Bruford, 2004). Consequently invaluable gene pool of native Turkish domestic sheep is in danger of loss or admixture. For this reason an immediate action must be taken to conserve the Turkish domestic sheep.

CHAPTER 5

CONCLUSION

1) SSCP was observed to be more efficient than that of RFLP method on identifying the major haplogroups, HPG-A, HPG-B and HPG-C.

2) RFLP may identify HPG-D and joint use of SSCP and RFLP may help to identify-HPG-E. Unidentified bands revealed by RFLP analysis of CR region and SSCP analysis of the ND4 region must be identified by sequencing. If this identification can be made may be sequencing will not be necessary to differentiate five mtDNA haplogroups. RFLP and/or SSCP methods, which are easier and cheaper than sequencing, can be used instead.

3) Comparisons of the HPG frequencies of the present study with those of the previous ones suggested that the sampling strategy followed in the present study seemed to reveal robust estimates for the breeds.

4) All of the statistical analysis indicated presence of at least two but more precisely 4 genetically homogenous groups for the 7 sheep breeds from Anatolia-Azerbaijan- Iran region. These groups are as follows; Group 1: Gökçeada; Group 2: Akkaraman, Dağlıç, Karayaka; Group 3: Morkaraman, Herik-Iran and Group 4: Azerbaijan.

5) Results of the present study confirmed the previously suggested scenario about the evolutionary history of Turkish domestic native sheep breeds. The scenario was:

Oldest sheep in Anatolia was composed of HPG-B. HPG-A was introduced from the East via Azerbaijan, and HPG-C from the south-east of Anatolia (Koban *et al.*, 2008).

6) Dağlıç, Karayaka and Akkaraman breeds seemed to represent oldest Anatolian sheep and they may be the maternal ancestors of all European sheep hence they must have priority in conservation of Turkish native sheep breeds.

CHAPTER 6

RECOMMENDATIONS FOR FUTURE STUDIES

For the further studies following suggestions can be made.

- mtDNA data can only be interpreted for maternal lineage. Other markers like Y chromosome for paternal linage and microsattelite markers for genome evolution must be studied.

- Ancient DNA samples that can be obtained from bones and fossils in the domestication centers must be studied. This kind of study will further enlighten the evolution of the domestic sheep.

- Samples from wild sheep which are possible candidate ancestor of domestic sheep can be studied to fatherly understand the evolution of domestic sheep.

- Evolutionary histories and migration patterns of other species in Turkey can be observed and compared with evolutionary history of sheep. Especially cattle, sheep, goats, pigs and buffalo which are known to be domesticated also in Fertile Crescent (Bruford *et al*., 2003) More specifically in or very near Anatolia (Zeder, 2008). Furthermore, human migrations can be studied and obtained data can be interpreted to gain further knowledge of history of animal domestication in Anatolia.

REFERENCES

Akçapınar H. (2000) Koyun Yetiştiriciliği. Second Edition, İsmat Matbaacılık Ltd. Şti., Ankara, Turkey.

Allendorf F.W., Luikart G. (2007) Conservation and the Genetics of Populations. First Edition, Wiley-Blackwell, MA, USA

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

Bradley D.G., MacHugh D.E., Cunningham P., Loftus R.T. (1996) Mitochondrial diversity and the origins of African and European cattle. Proceedings of the National Academy of Sciences, 93:5131-5135.

Brassel and Reif (1979) A procedure to generate Thiessian polygons. Geographical Analysis, 11:289-303.

Bruford M.W., Bradley D.G., Luikart G. (2003) DNA markers reveal the complexity of livestock domestication. Nature Genetics, 4:2-12.

Bruford M.W. (2004) Conservation genetics of UK livestock: from molecules to management. In: Simm G., Villanueva B., Sinclair K.D., Townsend S. (Editors) Farm Animal Genetic Resources. First Edition, Nottingham University Press, Nottingham, UK.

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. (Editors) Documenting Domestication: New Genetic and Archaeological Paradigm. First Edition, University of California Press, CA, USA, 307-317.

Bökönyi S. (1976) Development of early stock rearing in the Near East. Nature, 264:19-23.

Clutton-Brock J. (1981) Domesticated animals from early times. Heinemann/British Museum (Natural History), London, UK.

Daniel W.W. (1999) Biostatistics: A Foundation for Analysis in the Health Sciences. Seventh Edition, John Wiley and Sons, NY, USA.

Dinç H. (2003) *Alu* insertion polymorphisms in Anatolian Turks. Master Thesis, Middle East Technical University, Ankara, Turkey.

Dytham, C. (2003) Choosing and Using Statistics: A Biologist's Guide. Second Edition, Blackwell Publishing Company, NJ, USA.

Excoffier L., Smouse P., Quattro J. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics, 136:343-359.

Excoffier L., Laval G., Schneider S. (2006) Arlequin: An Integrated Software Package for Population Genetics Data Analysis. Version 3.1, Institute of Zoology, University of Berne, Bern, Switzerland.

Felsenstein J. (1993) PHYLIP: Phylogeny Inference Package. Version 3.6c, Department of Genetics, University of Washington, Seattle, USA.

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

Hartl D.L., Clark A.G. (1997) Principles of Population Genetics. Third Edition, Sinauer Associates, MA, USA.

Hedrick P.W. (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution, 53:313-318.

Hermans W.A (1996) The European mouflon, *Ovis musimon*. Tijdschrift Voor Diergeneeskunde, 121:515-517.

[a]Hiendleder S., Lewalski H., Wassmuth R., Janke A. (1998) 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.

[b]Hiendleder S., Mainz K., Plante Y., Lewalski H. (1998) 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., Kaupe B., Wassmuth R., 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.

Jobling M.A., Hurles M.E., Tyler-Smith C. (2004) Human Evolutionary Genetics: Origins, Peoples and Disease. First Edition, Garland Publishing, NY, USA.

Klug W.S., Cummings M.R. (2000) Concepts of Genetics. Sixth Edition, Prentice Hall, New Jersey, USA.

Koban E., Trinidad P., Bruford MW., Togan I. (2008) A genetic analysis of marginal sheep breeds from Turkey. Molecular Phylogenetics and Evolution, Submitted Manuscript.

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

Lewis R. (2005) Human Genetics: Concepts and Applications. Sixth Edition, McGraw-Hill, New York, USA.

Loftus R.T., Ertuğrul O., Harba A.H., El-Barody M.A.A., MacHugh D.E., Park S.D.E., Bradley D.G. (1999) A microsatellite survey of cattle from a centre of origin the Near East. Molecular Ecology, 8:3-8.

Longworth N. (2005) Agricultural production, prices and trade in Turkey. In: Burrell A.M., Oskam O.J. (Editors) The European Union Implications for Agricultural, Food and Structural Policy, CABI Publishing, Wallingford and Cambridge, 57-87.

Luikart G., Gielly L., Excoffier L., Vigne J-D., Bouvet J., Taberlet P. (2001) Multiple maternal origins and weak phylogeographic structure in domestic goats. Proceedings of the National Academy of Sciences, 98:5927-5932.

Luikart G., Fernandez H., Mashkour M., England P.R., Taberlet P. (2006) Origins and diffusion of domestic goats inferred from DNA markers. In: In: Zeder M.A., Bradley D.G., Emshwiller E., Smith B.D. (Editors) Documenting Domestication: New Genetic and Archaeological Paradigm. First Edition, University of California Press, CA, USA, 307-317.

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

MINITAB Package Program. Version 13, Minitab Inc., State College, PA, USA.

Monmonier M. (1973) Maximum difference barriers: An alternative numerical regionalization method. Geographical Analysis, 3:245-261.

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

Nei M., and Kumar S. (2000) Molecular Evolution and Phylogenetics. First Edition, Oxford University Press, New York, USA.

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

Pereira F., Davis S.J.M., Pereira L., McEvoy B., Bradley D.G., Amorim A. (2006) Genetic signatures of a Mediterranean influence in Iberian Peninsula sheep husbandry. Molecular Biology and Evolution, 23:1420-1426.

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

Peters J., Helmer D., Von Den Driesch A., Sana Segui M. (1999) Early animal domestication in the Northern Levant. Paleorient, 25:27-47.

Renfrew C. (1991) Before Babel: Speculations on the origins of linguistic diversity. Cambridge Archaeological Journal, 01:3-223.

Rohlf F.J (2000) NTSYS: Numerical Taxonomy and Multivariate Analysis System. Version 2.10q, Appiled Biostatistics Inc., NY, USA.

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

Tapio M., Marzanov M., Ozerov M., Cinkulov M., Gonzarenko G., Kiselyova T., Murawski M., Viinalass H., Kantanen J. (2006) Sheep mitochondrial DNA variation in European, Caucasian and Central Asian areas. Molecular Biology and Evolution, 23:1776-1783.

Takezaki N., Nei M. (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics, 144:389-399.

Tamura K., Dudley J., Nei M., Kumar S. (2008) MEGA: Molecular Evolutionary Genetics Analysis Software. Version 4.0, Molecular Biology and Evolution 24:1596-1599.

Uzun M., Gutiérrez-Gil B., Arranz J.J., Primitivo F.S., Saatci M., Kaya M., Bayón Y. (2006) Genetic relationship among Turkish sheep. Genetics Selection Evolution, 38:513-524.

Wang X., Ma Y.H., Chen H., Guan W.J. (2007) Genetic and phylogenetic studies of Chinese native sheep breeds (*Ovis aries*) based on mtDNA D-loop sequences. Small Ruminant Research, 72:232-236.

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

Zeder M.A (1999) Animal domestication in the Zagros: A review of past and current research. Paleorient, 25:11-25.

Zeder M.A. (2008) Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion and impact. Proceedings of the National Academy of Sciences, 105:11597-111604.

APPENDIX A

CHEMICAL SOLUTIONS USED IN THIS STUDY

1) 10 X Lysis Buffer:
   770mM NH_4Cl
   46mM $KHCO_3$
   10mM EDTA

2) Salt-EDTA Buffer (pH:7.5):
   75mM NaCl
   25mM EDTA

3) SDS Solution:
   10% (w/v) SDS

4) Sodium Acetate (NaAc):
   3M NaAc

5) Tris-HCl Solution (pH:8.0):
   10mM Tris
   1M HCl

6) Acetic Acid Solution
   5% (v/v) Acetic Acid
   5% (v/v) Ethanol

7) Ethanol Solution

0.5% (v/v) Acetic Acid

10% (v/v) Ethanol

8) 10X $AgNO_3$ Solution

97mM $AgNO_3$

9) Developer Solution

75mM NaOH

37% Formaldehyde

10) 5X TBE Buffer

450mM Tris

450mM Boric Acid

9mM EDTA

APPENDIX B

Haplogroup results of each individual according to the SSCP analysis of ND4 region and RFLP analysis of CR region. ("?" represents the results of unidentified band patterns for both analyses.)

BREED	INDIVIDUAL NUMBER	SSCP RESULTS	RFLP RESULTS
KARAYAKAYA	1	HPG-B	HPG-B
KARAYAKAYA	2	HPG-B	HPG-B
KARAYAKAYA	3	HPG-B/C	HPG-B/C
KARAYAKAYA	4	HPG-B	HPG-B
KARAYAKAYA	5	HPG-B	HPG-B
KARAYAKAYA	6	HPG-A	HPG-A
KARAYAKAYA	7	HPG-B/C	HPG-B/C
KARAYAKAYA	8	HPG-B/C	HPG-B/C
KARAYAKAYA	9	HPG-A	HPG-A
KARAYAKAYA	10	HPG-B	HPG-B
KARAYAKAYA	11	HPG-A	HPG-A
KARAYAKAYA	12	HPG-B	HPG-B
KARAYAKAYA	13	HPG-B	HPG-B
KARAYAKAYA	14	HPG-A	HPG-A
KARAYAKAYA	15	HPG-B	HPG-B
KARAYAKAYA	16	HPG-B	HPG-B
KARAYAKAYA	17	HPG-A	HPG-A
KARAYAKAYA	18	HPG-A	HPG-A
KARAYAKAYA	19	HPG-B	HPG-B
KARAYAKAYA	20	HPG-B	HPG-B
KARAYAKAYA	21	HPG-B	HPG-B
KARAYAKAYA	22	HPG-B	HPG-B
KARAYAKAYA	23	HPG-B	HPG-B
KARAYAKAYA	24	HPG-B	HPG-B
KARAYAKAYA	25	HPG-B	HPG-B
KARAYAKAYA	26	HPG-B	HPG-B
KARAYAKAYA	27	HPG-A	HPG-A

KARAYAKAYA	28	HPG-B	HPG-B
KARAYAKAYA	29	HPG-B	HPG-B
KARAYAKAYA	30	HPG-B	HPG-B
KARAYAKAYA	31	HPG-B	HPG-B
KARAYAKAYA	32	HPG-B	HPG-B
KARAYAKAYA	33	HPG-B	HPG-B
KARAYAKAYA	34	HPG-B	HPG-B
KARAYAKAYA	35	HPG-A	HPG-A
KARAYAKAYA	36	HPG-B	HPG-B
KARAYAKAYA	37	HPG-B	HPG-B
KARAYAKAYA	38	HPG-A	HPG-A
KARAYAKAYA	39	HPG-B	HPG-B
KARAYAKAYA	40	HPG-B	HPG-B
KARAYAKAYA	41	HPG-A	HPG-A
KARAYAKAYA	42	HPG-B	HPG-B
KARAYAKAYA	43	HPG-B	HPG-B
KARAYAKAYA	44	HPG-B/C	HPG-B/C
KARAYAKAYA	45	HPG-B	HPG-B
KARAYAKAYA	46	HPG-B	HPG-B
KARAYAKAYA	47	HPG-B	HPG-B
KARAYAKAYA	48	HPG-B	HPG-B
KARAYAKAYA	49	HPG-B	HPG-B
KARAYAKAYA	50	HPG-B	HPG-B
AKKARAMAN	1	HPG-B	HPG-B
AKKARAMAN	2	HPG-A	HPG-A
AKKARAMAN	3	HPG-A	HPG-A
AKKARAMAN	4	HPG-B	HPG-B
AKKARAMAN	5	HPG-B	HPG-B
AKKARAMAN	6	HPG-B	HPG-B
AKKARAMAN	7	HPG-B	HPG-B
AKKARAMAN	8	HPG-B	HPG-B
AKKARAMAN	9	HPG-B	HPG-B
AKKARAMAN	10	HPG-C	HPG-B/C
AKKARAMAN	11	HPG-B	HPG-B
AKKARAMAN	12	HPG-C	HPG-B/C
AKKARAMAN	13	?	HPG-B/C
AKKARAMAN	14	HPG-B	HPG-B
AKKARAMAN	15	HPG-B	HPG-B
AKKARAMAN	16	HPG-C	HPG-B/C
AKKARAMAN	17	HPG-A	HPG-A

AKKARAMAN	18	HPG-A	HPG-A
AKKARAMAN	19	HPG-A	HPG-A
AKKARAMAN	20	HPG-A	HPG-A
AKKARAMAN	21	HPG-B	HPG-B
AKKARAMAN	22	HPG-A	HPG-A
AKKARAMAN	23	HPG-C	HPG-B/C
AKKARAMAN	24	HPG-A	HPG-A
AKKARAMAN	25	HPG-B	HPG-B
AKKARAMAN	26	HPG-B	HPG-B
AKKARAMAN	27	HPG-B	HPG-B
AKKARAMAN	28	HPG-A	HPG-A
AKKARAMAN	29	HPG-B	HPG-B
AKKARAMAN	30	HPG-A	HPG-A
AKKARAMAN	31	HPG-B	HPG-B
AKKARAMAN	32	HPG-C	HPG-B/C
AKKARAMAN	33	HPG-A	HPG-A
AKKARAMAN	34	HPG-B	HPG-B
AKKARAMAN	35	HPG-C	HPG-B/C
AKKARAMAN	36	HPG-B	HPG-B
AKKARAMAN	37	HPG-B	HPG-B
AKKARAMAN	38	?	HPG-B/C
AKKARAMAN	39	HPG-C	HPG-B/C
AKKARAMAN	40	HPG-B	HPG-B
AKKARAMAN	41	HPG-B	HPG-B
AKKARAMAN	42	HPG-B	HPG-B
AKKARAMAN	43	HPG-B	HPG-B
AKKARAMAN	44	HPG-B	HPG-B
AKKARAMAN	45	HPG-C	HPG-B/C
AKKARAMAN	46	HPG-A	?
AKKARAMAN	47	HPG-A	HPG-A
AKKARAMAN	48	HPG-B	HPG-B
AKKARAMAN	49	HPG-B	HPG-B
AKKARAMAN	50	HPG-B	HPG-B
GÖKÇEADA	1	HPG-A	HPG-A
GÖKÇEADA	2	HPG-A	HPG-A
GÖKÇEADA	3	HPG-A	HPG-A
GÖKÇEADA	4	HPG-B	HPG-B
GÖKÇEADA	5	HPG-B	HPG-B
GÖKÇEADA	6	HPG-A	HPG-A
GÖKÇEADA	7	HPG-B	HPG-B

GÖKÇEADA	8	HPG-B	HPG-B
GÖKÇEADA	9	HPG-B	HPG-B
GÖKÇEADA	10	HPG-B	HPG-B
GÖKÇEADA	11	HPG-A	HPG-A
GÖKÇEADA	12	HPG-B	HPG-B
GÖKÇEADA	13	HPG-B	HPG-B
GÖKÇEADA	14	HPG-A	HPG-A
GÖKÇEADA	15	HPG-A	HPG-A
GÖKÇEADA	16	HPG-B	HPG-B
GÖKÇEADA	17	HPG-A	HPG-A
GÖKÇEADA	18	HPG-A	HPG-A
GÖKÇEADA	19	HPG-A	HPG-A
GÖKÇEADA	20	HPG-A	HPG-A
GÖKÇEADA	21	HPG-A	HPG-A
GÖKÇEADA	22	HPG-A	HPG-A
GÖKÇEADA	23	HPG-B	HPG-B
GÖKÇEADA	24	HPG-B	HPG-B
GÖKÇEADA	25	HPG-B	HPG-B
GÖKÇEADA	26	HPG-B	HPG-B
GÖKÇEADA	27	HPG-A	HPG-A
GÖKÇEADA	28	HPG-B	HPG-B
GÖKÇEADA	29	HPG-B	HPG-B
GÖKÇEADA	30	HPG-B	HPG-B
GÖKÇEADA	31	HPG-B	HPG-B
GÖKÇEADA	32	HPG-B	HPG-B
GÖKÇEADA	33	HPG-B	HPG-B
GÖKÇEADA	34	HPG-B	HPG-B
GÖKÇEADA	35	HPG-B	HPG-B
GÖKÇEADA	36	HPG-B	HPG-B
GÖKÇEADA	37	HPG-B	HPG-B
GÖKÇEADA	38	HPG-C	HPG-B
GÖKÇEADA	39	HPG-B	HPG-B
GÖKÇEADA	40	HPG-B	HPG-B
GÖKÇEADA	41	HPG-A	HPG-A
GÖKÇEADA	42	HPG-B	HPG-B
GÖKÇEADA	43	HPG-B	HPG-B
GÖKÇEADA	44	HPG-B	HPG-B
GÖKÇEADA	45	HPG-B	HPG-B
GÖKÇEADA	46	HPG-B	HPG-B
GÖKÇEADA	47	HPG-B	HPG-B

GÖKÇEADA	48	HPG-B	HPG-B
GÖKÇEADA	49	HPG-B	HPG-B
GÖKÇEADA	50	HPG-A	HPG-A
DAĞLIÇ	1	HPG-C	HPG-B/C
DAĞLIÇ	2	HPG-B	HPG-B
DAĞLIÇ	3	HPG-B	HPG-B
DAĞLIÇ	4	HPG-B	HPG-B
DAĞLIÇ	5	HPG-B	HPG-B
DAĞLIÇ	6	HPG-B	HPG-B
DAĞLIÇ	7	HPG-B	HPG-B
DAĞLIÇ	8	HPG-B	HPG-B
DAĞLIÇ	9	?	HPG-B/C
DAĞLIÇ	10	HPG-C	HPG-B/C
DAĞLIÇ	11	HPG-B	HPG-B
DAĞLIÇ	12	HPG-C	HPG-B/C
DAĞLIÇ	13	HPG-B	HPG-B
DAĞLIÇ	14	HPG-B	HPG-B
DAĞLIÇ	15	HPG-B	HPG-B
DAĞLIÇ	16	HPG-B	HPG-B
DAĞLIÇ	17	HPG-B	HPG-B
DAĞLIÇ	18	HPG-B	HPG-B
DAĞLIÇ	19	HPG-B	HPG-B
DAĞLIÇ	20	HPG-B	HPG-B
DAĞLIÇ	21	HPG-A	HPG-A
DAĞLIÇ	22	HPG-B	HPG-B
DAĞLIÇ	23	HPG-B	HPG-B
DAĞLIÇ	24	?	HPG-B
DAĞLIÇ	25	HPG-A	HPG-A
DAĞLIÇ	31	HPG-B	HPG-B
DAĞLIÇ	32	HPG-C	HPG-B/C
DAĞLIÇ	33	HPG-B	HPG-B
DAĞLIÇ	34	HPG-B	HPG-B
DAĞLIÇ	35	?	HPG-B/C
DAĞLIÇ	36	HPG-C	HPG-B/C
DAĞLIÇ	37	HPG-B	HPG-B
DAĞLIÇ	38	HPG-B	HPG-B
DAĞLIÇ	39	HPG-C	HPG-B/C
DAĞLIÇ	40	HPG-C	HPG-B/C
DAĞLIÇ	41	HPG-B	HPG-B
DAĞLIÇ	42	HPG-B	HPG-B

DAĞLIÇ	43	HPG-B	HPG-B
DAĞLIÇ	44	HPG-C	HPG-B/C
DAĞLIÇ	45	HPG-B	HPG-B
DAĞLIÇ	46	HPG-B	HPG-B
DAĞLIÇ	47	HPG-B	HPG-B
DAĞLIÇ	48	HPG-B	HPG-B
DAĞLIÇ	49	HPG-B	HPG-B
DAĞLIÇ	50	HPG-A	HPG-A
DAĞLIÇ	55	?	HPG-B/C
DAĞLIÇ	56	HPG-A	?
DAĞLIÇ	1862	HPG-B	HPG-B
DAĞLIÇ	50281	HPG-A	HPG-A
MORKARAMAN	1	HPG-C	HPG-B/C
MORKARAMAN	2	HPG-A	HPG-A
MORKARAMAN	3	HPG-B	HPG-B
MORKARAMAN	4	HPG-A	HPG-A
MORKARAMAN	5	HPG-B	HPG-B
MORKARAMAN	6	HPG-A	HPG-A
MORKARAMAN	7	HPG-A	HPG-A
MORKARAMAN	8	HPG-C	HPG-B/C
MORKARAMAN	9	HPG-C	HPG-B/C
MORKARAMAN	10	HPG-A	HPG-A
MORKARAMAN	11	HPG-B	HPG-B
MORKARAMAN	12	HPG-B	HPG-B
MORKARAMAN	13	HPG-B	HPG-B
MORKARAMAN	14	HPG-C	HPG-B/C
MORKARAMAN	15	HPG-B	HPG-B
MORKARAMAN	16	HPG-B	HPG-B
MORKARAMAN	17	HPG-B	HPG-B
MORKARAMAN	18	HPG-C	HPG-B/C
MORKARAMAN	19	HPG-A	HPG-A
MORKARAMAN	20	HPG-B	HPG-B
MORKARAMAN	21	HPG-A	HPG-A
MORKARAMAN	22	HPG-A	HPG-A
MORKARAMAN	23	HPG-A	HPG-A
MORKARAMAN	24	HPG-C	HPG-B/C
MORKARAMAN	25	HPG-B	HPG-B
MORKARAMAN	26	HPG-A	HPG-A
MORKARAMAN	27	HPG-B	HPG-B
MORKARAMAN	28	HPG-C	HPG-B/C

MORKARAMAN	29	HPG-C	HPG-B/C
MORKARAMAN	30	HPG-C	HPG-B/C
MORKARAMAN	31	HPG-C	HPG-B/C
MORKARAMAN	32	HPG-A	HPG-A
MORKARAMAN	33	HPG-B	HPG-B
MORKARAMAN	34	HPG-B	HPG-B
MORKARAMAN	35	HPG-A	HPG-A
MORKARAMAN	36	HPG-A	HPG-A
MORKARAMAN	37	HPG-C	HPG-B/C
MORKARAMAN	38	HPG-A	HPG-A
MORKARAMAN	39	HPG-A	HPG-A
MORKARAMAN	40	HPG-C	HPG-B/C
MORKARAMAN	41	HPG-C	HPG-B/C
MORKARAMAN	42	?	HPG-A
MORKARAMAN	43	HPG-C	HPG-B/C
MORKARAMAN	44	HPG-B	HPG-B
MORKARAMAN	45	HPG-B	HPG-B
MORKARAMAN	46	HPG-B	HPG-B
MORKARAMAN	47	HPG-A	HPG-A
MORKARAMAN	48	HPG-C	HPG-B/C
MORKARAMAN	49	HPG-B	HPG-B
MORKARAMAN	50	HPG-A	HPG-A
HERIK-IRAN	1	HPG-B	HPG-B
HERIK-IRAN	2	HPG-B	HPG-B
HERIK-IRAN	3	HPG-B	HPG-B
HERIK-IRAN	4	HPG-A	HPG-A
HERIK-IRAN	5	HPG-C	HPG-B/C
HERIK-IRAN	6	HPG-C	HPG-B/C
HERIK-IRAN	7	HPG-C	HPG-B/C
HERIK-IRAN	8	HPG-A	HPG-A
HERIK-IRAN	9	HPG-C	HPG-B/C
HERIK-IRAN	10	HPG-B	HPG-B
HERIK-IRAN	11	HPG-A	HPG-A
HERIK-IRAN	12	HPG-C	HPG-B/C
HERIK-IRAN	13	HPG-A	HPG-A
HERIK-IRAN	14	HPG-C	HPG-B/C
HERIK-IRAN	15	HPG-A	HPG-A
HERIK-IRAN	16	HPG-B	HPG-B
HERIK-IRAN	17	HPG-C	HPG-B/C
HERIK-IRAN	18	HPG-B	HPG-B

HERIK-IRAN	19	HPG-B	HPG-B
HERIK-IRAN	20	HPG-A	HPG-A
AZERBAIJAN	1	HPG-B	HPG-B
AZERBAIJAN	2	HPG-A	HPG-A
AZERBAIJAN	3	HPG-A	HPG-A
AZERBAIJAN	4	HPG-A	HPG-A
AZERBAIJAN	5	HPG-B	HPG-B
AZERBAIJAN	6	HPG-A	HPG-A
AZERBAIJAN	7	HPG-A	HPG-A
AZERBAIJAN	8	HPG-B	HPG-B
AZERBAIJAN	9	HPG-B	HPG-B
AZERBAIJAN	10	HPG-A	HPG-A
AZERBAIJAN	11	HPG-C	HPG-B/C
AZERBAIJAN	12	HPG-A	HPG-A
AZERBAIJAN	13	HPG-A	HPG-A
AZERBAIJAN	14	HPG-A	HPG-A
AZERBAIJAN	15	HPG-A	HPG-A
AZERBAIJAN	16	HPG-A	HPG-A
AZERBAIJAN	17	HPG-B	HPG-B
AZERBAIJAN	18	HPG-C	HPG-B/C
AZERBAIJAN	19	HPG-A	HPG-A
AZERBAIJAN	20	HPG-B	HPG-B
AZERBAIJAN	21	HPG-B	HPG-B
AZERBAIJAN	22	HPG-A	HPG-A
AZERBAIJAN	23	HPG-B	HPG-B
AZERBAIJAN	24	HPG-A	HPG-A
AZERBAIJAN	25	HPG-A	HPG-A
AZERBAIJAN	26	HPG-A	HPG-A
AZERBAIJAN	27	HPG-B	HPG-B
AZERBAIJAN	28	HPG-B	HPG-B
AZERBAIJAN	29	HPG-B	HPG-B
AZERBAIJAN	30	HPG-B	HPG-B
AZERBAIJAN	31	HPG-B	HPG-B
AZERBAIJAN	32	HPG-B	HPG-B
AZERBAIJAN	33	HPG-A	HPG-A
AZERBAIJAN	34	HPG-B	HPG-B
AZERBAIJAN	35	HPG-A	HPG-A
AZERBAIJAN	36	HPG-A	HPG-A
AZERBAIJAN	37	HPG-A	HPG-A
AZERBAIJAN	38	HPG-A	HPG-A

AZERBAIJAN	39	HPG-B	HPG-B
AZERBAIJAN	40	HPG-A	HPG-A
AZERBAIJAN	41	HPG-B	HPG-B
AZERBAIJAN	42	HPG-A	HPG-A
AZERBAIJAN	43	HPG-A	HPG-A
AZERBAIJAN	44	HPG-C	HPG-B/C

APPENDIX C

Table 1 Ranked values of the X and Y components and the di values of the HPG-A Spearman Rank Correlation Test. (Gökçeada breed included.)

X Component (Frequency)	Y Component (Geographical Distance)	di Values
1	5	-4
2	3	-1
3	4	-1
4	6	-2
5	2	3
6	1	5

Table 2 Ranked values of the X and Y components and the di values of the HPG-A Spearman Rank Correlation Test. (Gökçeada breed excluded.)

X Component (Frequency)	Y Component (Geographical Distance)	di Values
1	5	-4
2	3	-1
3	4	-1
4	2	2
5	1	4

Table 3 Ranked values of the X and Y components and the di values of the HPG-C Spearman Rank Correlation Test.

X Component (Frequency)	Y Component (Geographical Distance)	di Values
1	6	-5
2	3	-1
3	4	-1
4	5	-1
5	2	3
6	1	5

APPENDIX D

Observed and expected values of the X^2 Test of Homogeneity.

BREEDS	HAPLOGROUPS	Present Study		Pedrosa *et al.*'s (2005) Study	
		Observed	**Expected**	**Observed**	**Expected**
KARAYAKA	HPG-A	10	11.19	5	3.81
	HPG-B	36	31.34	6	10.66
	HPG-C	4	7.46	6	2.54
AKKARAMAN	HPG-A	12	11.19	3	3.81
	HPG-B	27	25.37	7	8.63
	HPG-C	8	10.44	6	3.56
MORKARAMAN	HPG-A	17	19.14	8	5.86
	HPG-B	17	16.84	5	5.17
	HPG-C	15	13.02	2	3.98

BREEDS	HAPLOGROUPS	Present Study		Meadows *et al.* 's (2007)	
		Observed	**Expected**	**Observed**	**Expected**
KARAYAKA	HPG-A	10	12.31	6	3.69
	HPG-B	36	34.62	9	10.38
	HPG-C	4	3.08	0	0.92
MORKARAMAN	HPG-A	17	15.85	5	6.15
	HPG-B	17	20.18	11	7.82
	HPG-C	15	11.53	1	4.47
	HPG-D	0	1.44	2	0.56

BREEDS	HAPLOGROUPS	Present Study		Koban *et al.*'s (2008) Study	
		Observed	**Expected**	**Observed**	**Expected**
KARAYAKA	HPG-A	10	10.78	12	11.22
	HPG-B	36	35.29	36	36.71
	HPG-C	4	3.92	4	4.08
AKKARAMAN	HPG-A	12	11.87	13	13.13
	HPG-B	27	28.96	34	32.04
	HPG-C	8	6.17	5	6.83
DAĞLIÇ	HPG-A	5	2.92	2	4.08
	HPG-B	32	36.67	56	51.33
	HPG-C	8	5.42	5	7.58
MORKARAMAN	HPG-A	17	20.42	18	14.58
	HPG-B	17	18.76	15	13.33
	HPG-C	15	9.92	2	7.08