UNDERSTANDING DOMESTICATION PROCESS OF SHEEP ACROSS CENTRAL AND WESTERN ANATOLIA BY USING ANCIENT DNA

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

BY

ONUR ÖZER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MOLECULAR BIOLOGY AND GENETICS

OCTOBER 2017

Approval of the thesis:

UNDERSTANDING DOMESTICATION PROCESS OF SHEEP ACROSS CENTRAL AND WESTERN ANATOLIA BY USING ANCIENT DNA

submitted by **ONUR ÖZER** in partial fulfillment of the requirements for the degree of **Master of Science in Molecular Biology and Genetics Department, Middle East Technical University by,**

Prof. Dr. Gülbin Dural Ünver ____________
Dean, **Graduate School of Natural and Applied Sciences**

Prof. Dr. Orhan Adalı ____________
Head of Department, **Biology**

Assoc. Prof. Dr. Mehmet Somel ____________
Supervisor, **Biology Dept., METU**

Examining Committee Members:

Assoc. Prof. Dr. Can Bilgin ____________
Biology Dept., METU

Assoc. Prof. Dr. Mehmet Somel ____________
Biology Dept., METU

Assoc. Prof. Dr. Evangelia Pişkin ____________
Settlement Archaeology Dept., METU

Assoc. Prof. Dr. Çiğdem Atakuman ____________
Settlement Archaeology Dept., METU

Assist. Prof. Dr. Ali Metin Büyükkarakaya ____________

Anthropology Dept., Hacettepe University

Date: 26.10.2017

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

Name, Last name: Onur Özer

Signature :

ABSTRACT

UNDERSTANDING DOMESTICATION PROCESS OF SHEEP ACROSS CENTRAL AND WESTERN ANATOLIA BY USING ANCIENT DNA

Özer, Onur
M.Sc., Department of Molecular Biology and Genetics
Supervisor: Assoc. Prof. Dr. Mehmet Somel

October 2017, 101 pages

Several archaeological and genetic studies indicated that Southeastern Anatolia was the only center of domestication for sheep. The study presented here aims to understand how and when domestic sheep were transported across Anatolia into west from the domestication center by using ancient DNA. In order to achieve that, ancient DNA was extracted from 234 sheep bone samples dating between Epipaleolithic and 2800 BCE from 9 archaeological excavations (Tepecik-Çiftlik, Yeşilova, Ulucak, Aktopraklık, Barcın, Çatalhöyük, Boncuklu, Canhasan III, Pınarbaşı,). A 144 base pair (bp) long fragment of sheep mtDNA was successfully amplified for 121 of these samples yielding a success rate of 52%. The targeted 144 base pair long fragment was shown to be able to identify five mtDNA haplogroups (A-E) observed in modern sheep breeds.

Domestic sheep of Central and Western Anatolia within the mentioned time intervals were found to be dominated by HPG B. Temporal analysis of haplogroup diversity indicates a wave of migration into Central Anatolia at around 7000 BCE most likely from the east.

Haplogroup distribution of initial phases of Yeşilova Höyük shows a strong deviation from the general trend with a high frequency of HPG A (75%). This deviation may be the result of "maritime route" expansion through which seafaring voyagers migrates into western Anatolia by following the southern coasts.

Temporal analysis of haplotype and nucleotide diversity within the HPG B individuals revealed a strong domestication bottleneck and loss of within-haplogroup diversity after 7000 BC.

Results of the present study provides information on spatial and temporal distribution of mtDNA haplogroups of sheep Anatolia mainly for the Neolithic Period and contributes to the understanding of initial phases of domestication process of sheep across central and western Anatolia.

Keywords: ancient DNA, mtDNA haplogroup, Neolithic, sheep,
domestication

ÖZ

ANTİK DNA KULLANILARAK KOYUNLARIN ORTA VE BATI ANADOLU'DAKİ EVCİLLEŞME SÜRECİNİN ORTAYA ÇIKARILMASI

Özer, Onur
Yüksek Lisans, Moleküler Biyoloji ve Genetik Bölümü
Tez Yöneticisi: Doç. Dr. Mehmet Somel

Ekim 2017, 101 sayfa

Birçok arkeoloji ve genetik çalışması Doğu Anadolu'nun koyunun tek evcilleştirme merkezi olduğunu işaret etmektedir. Burada sunulan araştırmada antik DNA kullanılarak evcil koyunların Anadolu üzerinden ilk evcilleştirme merkezinden batıya nasıl ve ne zaman götürüldüğünün anlaşılması amaçlanmıştır. Bu amaç doğrultusunda 9 arkeolojik kazıdan (Yeşilova, Ulucak, Aktopraklık, Barcın, Çatalhöyük, Boncuklu, Canhasan III, Pınarbaşı, Tepecik-Çiftlik) toplanan ve Epipaleolitik Dönem ile M.Ö. 2800 tarihleri arasına tarihlendirilmiş 234 koyun kemiği örneğinden antik DNA izole edilmiştir. Bu örneklerin 121 tanesinden yani %52'lük bir başarı oranıyla 144 baz çifti uzunluğunda koyun mitokondriyal DNA parçası yükseltgenebilmiştir. Geçmiş çalışmalar, hedeflenen 144 baz çifti uzunluğundaki parçanın modern koyun ırklarında görülen beş mtDNA haplogrubunu tanımlayabildiğini göstermiştir.

Belirtilen tarih aralıklarında Orta ve Batı Anadolu'ya B haplogrubu koyunların hakim olduğu ortaya çıkarılmıştır. Genetik çeşitliliğin zamansal analizi M.Ö. 700 yılında Orta Anadolu'nun büyük ihtimalle evcilleştirme merkezinden bir göç aldığına işaret etmektedir.

Yeşilova Höyük'ün ilk aşamalarından elde edilen haplogroup dağılımı %75'lik HPG A frekansı ile genel yönelimden ciddi bir sapma göstermiştir. Bu sapma, deniz yolunu kullanarak Anadolu'nun güney kıyılarından batıya göç eden toplumlar tarafından ortaya çıkarılmış olabilir.

HPG B bireyler arasındaki nükleotid ve haplotip çeşitliliğinin zamansal analizi, M.Ö. 7000 yılından sonra çeşitliliğin azaldığını göstererek büyük bir evcilleştirmeye bağlı darboğaz etkisine işaret etmiştir.

Bu çalışmada sunulan sonuçlar Anadolu'da Neolitik Dönem'deki koyun toplumlarının mtDNA haplogruplarının uzay-zamansal dağılımı hakkında bilgi vermekte ve evcil koyunların Anadolu üzerinden Batı'ya yayılmalarının ilk aşamalarının anlaşılmasına katkı sağlamaktadır. Ayrıca Anadolu içerisinde yaban koyunundan evcil koyuna gen akışının da izleri tespit edilmiştir.

Anahtar kelimeler: antik DNA, mtDNA haplogrubu, Neolitik, koyun, evcilleştirme

To my family

ACKNOWLEDGMENTS

Firstly, I would like to thank to Prof. Dr. İnci Togan for introducing me to the ancient DNA research, for sharing her scientific enthusiasm and guidance with me. Every conversation we had replenished my motivation, my curiosity on science and also my respect to her.

I would like to thank my thesis committee members, Assoc. Prof. Dr. Can Bilgin, Assoc. Prof. Dr. Mehmet Somel, Assoc. Prof. Dr. Evangelia Pişkin, Assoc. Prof. Dr. Çiğdem Atakuman and Assist. Prof. Dr. Ali Metin Büyükkarakaya for their valuable contributions.

I would like to thank to Dr. Füsun Özer, Nihan Dilşad Dağtaş Kılıç and Dr. Eren Yüncü who patiently taught me how to work in the clean ancient DNA laboratory as well as how to analyze genetic data. Writing this thesis wouldn't be possible without their generous support on every step. Similarly, my remaining lab mates Reyhan Yaka, Ali Akbaba, Mustafa Özkan and my honorary lab mate Nurbahar Usta were always supportive and eager to help me solve any problem I have.

I am thankful to my current supervisor Assoc. Prof. Dr. Mehmet Somel and my previous supervisor Assist. Prof. Dr. Ayşegül Birand for their academic support. Together with Dr. Sinan Can Açan, their help on producing coalescent simulations and their comments on results were invaluable for my thesis.

I am grateful to Pelin Kasap for her love, friendship and companion in several parts of the world.

Finally, I would like to express my deepest gratitude to my parents, Gülten Özer and Mehmet Ali Özer. Their wholehearted love and endless

support carried me until this point and knowing that their support will continue makes me feel invincible.

This study was supported by the Scientific and Technical Research Council of Turkey (TÜBİTAK) as a part of the project "Güneydoğu Anadolu'da evcilleştirildiği bilinen koyunun, evcilleştirme merkezinden Batı Anadolu'ya doğru gidiş yolunda geçirdiği evcilleştirme sürecinin antik mtDNA kullanılarak araştırılması" under the grant number 114Z356.

TABLE OF CONTENTS

ABSTRACT...v

ÖZ vii

ACKNOWLEDGMENTS x

TABLE OF CONTENTS.. xii

LIST OF FIGURES xiv

LIST OF TABLES xvii

LIST OF ABBREVATIONS xvii

INTRODUCTION 1

1.1 Domestication and human-mediated migration of sheep..................1

1.2 Sheep mitochondrial DNA haplogroups 5

1.3 Ancient DNA research 7

1.4 Ancient DNA research on sheep 11

1.5 Objectives of the study...13

MATERIALS AND METHODS..15

2.1 Material 15

2.2 Laboratory Experiments..17

2.2.1. Working in dedicated clean ancient DNA laboratory and precautions to prevent contamination 17

2.2.2. Extraction of aDNA 17

2.2.3. DNA amplification...19

2.2.4. Agarose gel electrophoresis 22

2.2.5. Purification of amplified DNA and sequencing...........................22

2.3 Analyses..23

2.3.1. Alignment and editing of sequences 23
2.3.2. Assignment of haplogroups 23
2.3.3. Genetic diversity measurements 24
2.3.4. Coalescent simulations of population continuity 26
RESULTS 29
3.1 Ancient DNA extraction and amplification 29
3.2 Editing and aligning sequences 32
3.3 mtDNA haplogroup determination and spatial-temporal distribution of sheep haplogroups 33
3.4 mtDNA diversity calculations 36
3.4.1 mtDNA haplogroup diversity across Neolithic Period for central and western Anatolia 36
3.4.2. Molecular diversity measurements 41
3.1 Coalescent simulations for population continuity 44
3.5.1 Simulations with assumption of no growth 46
3.5.2 Simulations with assumption of bottleneck and exponential growth 53
DISCUSSION 61
4.1. Establishment of DNA extraction protocol 61
4.2. Spatial and temporal distribution and genetic diversities of mtDNA haplogroups in central and western Anatolia 62
4.3 Continuity tests with coalescent simulations 69
4.4 Temporal diversity changes of HPG B in central and western Anatolia 71
4.5 Conclusion 72
REFERENCES 75
APPENDIX A 87
APPENDIX B 101

LIST OF FIGURES

Figure 1. Figure showing the domestication center of four major livestock animals. 2
Figure 2 Haplogroup (Y1, Y2, Arm1T, Arm2T, European) distribution of ancient domestic pigs over excavation sites and different time periods.. 11
Figure 3. Map of Anatolia with locations of excavations from which samples are collected. 16
Figure 4. Sequences of forward (L15391) and reverse (H15534) primers.. 20
Figure 5. Image of an agarose gel. 31
Figure 6. Chromatogram of the sample BH24, sequenced with reverse primer. 32
Figure 7. Alignment of reverse and forward reads to the reference sequence. 33
Figure 8. Partial mtDNA sequence of individuals with ambiguous HPG identity.. 36
Figure 9. Pie charts showing the haplogroup composition of domestic sheep for different time periods for central and western Anatolia. 37
Figure 10. Significance test with 10000 permutations for the difference between Shannon diversity indices of Epipaleolithic – 7000 BCE and the successive period 7000 BCE – 6500 BCE. 39
Figure 11. Significance test with 10000 permutations for the difference between Shannon diversity indices of 7000 BCE – 6500 BCE and the successive period 6500 BCE – 6000 BCE.. 40
Figure 12. Significance test with 10000 permutations for the difference between Shannon diversity indices of 6500 BCE – 6000 BCE and the successive period 6000 BCE – 2800 BCE. 41

Figure 13. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and late Tepecik Çiftlik (TP 2) populations..46
Figure 14. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and Ulucak Höyük (UH) Populations..47
Figure 15. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and Yeşilova Höyük (YH) Populations..48
Figure 16. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and early Barcın Höyük (BH 1) populations..49
Figure 17. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Ulucak Höyük (UH 1) and early Barcın Höyük (BH 1) Populations. 50
Figure 18. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Ulucak Höyük (UH) and Yeşilova Höyük (YH) Populations..51
Figure 19. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early (BH 1) and (BH 2) late Barcın Höyük Populations. 52
Figure 20. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and late Tepecik Çiftlik (TP2) Populations. 53
Figure 21. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and Ulucak Höyük (UH) Populations..54

Figure 22. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and Yeşilova Höyük (YH) Populations....55
Figure 23. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 2) and early Barcın Höyük (BH 1) Populations. 56
Figure 24. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Ulucak Höyük (UH 1) and early Barcın Höyük (BH 1) Populations.. 57
Figure 25. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Ulucak Höyük (UH) and Yeşilova Höyük (YH) Populations....58
Figure 26. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early (BH 1) and late (BH 2) Barcın Höyük Populations. 59
Figure 27. Map of Turkey showing possible routes for human mediated sheep dispersal with black arrows....64

LIST OF TABLES

Table 1. Preparation of binding buffer..19
Table 2. Reagents for PCR master mix..21
Table 3. PCR amplification conditions..21
Table 4. mtDNA control region nucleotide positions used for haplogroup assignment. Table adapted from Demirci et al. (2013)..................................24
Table 5. Excavation sites, number of samples available from the sites and number of samples with successfully amplified and sequenced aDNA. 30
Table 6. mtDNA haplogroup compositions of domestic sheep from Anatolia 35
Table 7. Shannon Diversity Indices for specified time periods, number of samples (n) used and their excavation sites. YH (Yeşilova Höyük), UH (Ulucak Höyük), BH (Barcın Höyük), AT (Aktopraklık Höyük), CH (Canhasan III), CT (Çatalhöyük), TP (Tepecik Çiftlik Höyüğü), BK (Boncuklu Höyük), PB (Pınarbaşı Höyük) 38
Table 8. Molecular diversity measurements of individuals before and after 7000 BCE..42
Table 9. Parameters between populations used for simulations. Values within brackets indicate sample sizes 45
Table 10. Frequency of simulated populations with indicated significance values over all mutation rate and population size values...............................60

LIST OF ABBREVATIONS

DNA	deoxyribonucleic acid
aDNA	ancient DNA
BCE	Before Common Era
bp	base pair
BP	Before Present
CR	control region
dNTP	deoxyribonucleotide triphosphates
g	gravitational acceleration
HPG	haplogroup
min	minute
ml	milliliter
mm	millimeter
mtDNA	mitochondrial DNA
PCR	Polymerase Chain Reaction
rpm	revolutions per minute
s	second
$tRNA^{Pro}$	transfer RNA proline gene

µl: microliter

CHAPTER 1

INTRODUCTION

1.1 Domestication and human-mediated migration of sheep

Neolithic Transition is considered as one of the most important cornerstones in the history of humankind. This transition is marked by abandoning the foraging and hunting practices and adopting a farming-based sedentary life style (Zeder 2008). Changes leading to Neolithic Transition first occurred in the region known as “Fertile Crescent”, covering the Southeast Anatolia, Levant, Mesopotamia and western borders of Zagros Mountains. This new way of living was made possible by domestication of plants and animals. Four major livestock animals (sheep, goat, cattle, pig) are domesticated in the northern Fertile Crescent region and sheep together with goat were the first ones to be domesticated at around 11.000 BP (Peters et al. 2005; Zeder et al. 2005; Zeder 2008) (Figure 1).

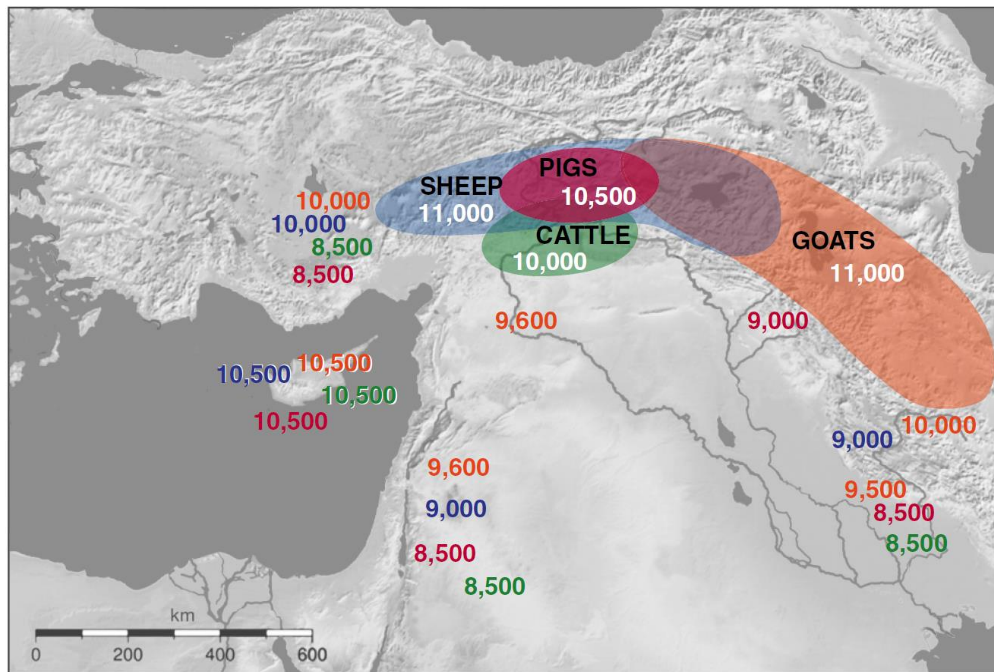


Figure 1. Figure showing the domestication center of four major livestock animals. Shaded areas indicates the domestication centers. Numbers indicate the dates (in years before present) of first observations of domestic animals out of the domestication center. Figure is taken from Zeder (2008)

After their initial domestication, sheep are carried into Europe (Vigne, 2014), to North Africa (Barker, 2002) and to Asia (Price, 2000; Lv et al., 2015) by migrating humans. These dispersion routes have been a topic of interest for scientists because the information can help to elucidate the origins of and relationship between modern breeds. Also, it indicates migration and trading routes of humans in Neolithic period. Archaeozoologists used several methods to identify the place and time of initial domestication as well as the migration of domestic sheep. Until recently the most useful marker of domestication is thought to be the reduction in body size (Uerpmann H-P, 1979; Meadow RH,

1989). This criterion is based on the idea that herded animals are much smaller than their wild counterparts. However recent research showed that the main factor affecting the body size in sheep is the sexual dimorphism between male and females (Zeder, 2006). Wild males are observed to be bigger than wild females. Thus, the observed difference between the size of animal bones found in hunting communities and bones found in herding communities can be explained by the different culling strategies between these communities. Accordingly, few males are reserved for breeding in herding communities and the rest of the males are slaughtered at very young ages whose fragile bones generally were not preserved in the archaeological sites. As a result, the animal bones found from hunting communities mainly belonged to large old males whereas the bones from herding communities mainly belonged to relatively small old females. So called "young male culling strategy" was assumed to apply for goats in several archaeological excavations from Iran and Iraq. Results indicated that the goats dating to 9900 BP from the site of Ganj Dareh showed the signatures of a herding community where the males were harvested early in their life but females were kept alive until they were old. (Zeder, 2006) These goats do not show any evidence of domestication-related morphological changes. Therefore, it was argued that the "young male culling strategy" can be the early archaeozoological marker of domestication.

Search of "young male culling strategy" in different archaeological excavation sites yielded interesting results. For example, in the beginning of 1990s several new sites were excavated in Cyprus (Guilaine, 2003). These sites dated between 10500-900 BP and all four major livestock animals were identified in them. Demographic profiles of these livestock animals were consistent with the "young male culling strategy" although they showed no morphological evidence of domestication. Domestic plants that were not native to Cyprus were also found in these sites. Altogether, these evidences indicated that the seafaring colonists were able to build boats that could carry big animals

for long distances at the beginning of Neolithic period (Vigne, 2014). Further investigation of other Mediterranean Islands by Vigne (2014) also showed rapid colonization of seafaring communities with a fully-formed Neolithic culture where four major livestock animals were carried. Therefore, a maritime route across the coasts and islands of the Mediterranean See was suggested as one of the routes for the westward dispersal of Neolithic Culture from Southwest Asia into Europe (Zilhao, 2001).

Arbuckle and Atıcı (2013) analyzed the demography of sheep from several archaeological sites across the Anatolia in order to observe when and where the young male culling strategy, thus sign of early domestication was observed in Anatolia. Their results indicated that the early husbandry systems in line with long lasting process of domestication (Larson and Burger, 2013), within Anatolia were highly diverse and probably shaped by local environmental pressures as well as by cultural traditions and social interactions. This diversity includes practicing both herding and hunting simultaneously (Arbuckle, 2008a) or culling of sheep within a narrow age interval without any focus on young males (Arbuckle, 2008b). These different husbandry systems are called "experimental" systems by Arbuckle and it is suggested that the way people handle their herds might involve introgression of wild sheep into domestic herds (Arbuckle et al. 2014).

Although sheep were domesticated primarily for their meat, the secondary products of sheep also gained importance as humans lived together with them for longer times (Sherratt, 1983). A secondary migration through which a new type of sheep with a high wool quality was spread into Europe, Asia and Africa at around 3000 BCE was proposed by Chessa et al. (2009). Studies of endogenous retrovirus integration sites by Chessa et al. (2009) indicated that the effect of this secondary migration was visible by the existence of genetic marker enJSRV-18 in almost all breeds around the world. Markers of the first

migration presumably the one present study is focusing on were R0 (no insertionally polymorphic enJSRVs) and R1 (enJSRV-7) were observed, only in the periphery of Europe such as in Cyprus, Sicily, Corsica, Soay, Iceland, Norway. Almost of all of these sheep with R0 and R1 are wild sheep with morphologies very different than modern domestic sheep.

1.2 Sheep mitochondrial DNA haplogroups

Mitochondrial DNA (mtDNA) is a commonly used marker for studying the history of domestic animals and process of domestication. Bruford et al. (2003) suggested that the mtDNA is suitable for domestication studies in several ways. First, it is conserved enough to identify the wild ancestor from which the domestic animal is derived. Second, it is variable enough across the geographic distribution of species so that the place of domestication can be approximated. Moreover, mammalian mtDNA is haploid, does not undergo recombination. Therefore, interpretation of the phylogenetic analyses is relatively straightforward. It is inherited maternally, thus maternal history can be deduced. Finally, it evolves at a rapid but constant rate so that the origin of specific polymorphisms can be dated. Besides its low cost of study, the suitability, of mtDNA for domestication studies made it the most widely used molecular marker. The accumulated mtDNA data can be used in comparative studies.

Initial studies on sheep mtDNA found two lineages, haplogroup (HPG) A and HPG B, within the modern sheep breeds (Hiendleder et al. 2002). These two lineages are the most common ones among the modern sheep breeds. Despite the fact that both of them are found in almost every breed, the relative frequency of HPG A is high in Asia whereas the relative frequency of HPG B

is high in Europe. In 2005 another HPG was described in several breeds from Turkey (Pedrosa et al. 2005) and China (Guo et al. 2005) and called HPG C. HPG C was later observed in the Iberian Peninsula (Pereira et al. 2006) and in Caucasus as well as in the Central Asia (Tapio et al. 2006). Tapio et al. (2006) also defined a fourth clade within the breeds of the Northern Caucasus and named it as HPG D. Two more HPG D individuals are identified within the Morkaraman breed (in the northeastern part of Turkey) of Turkey by Meadows et.al. (2007). The same study of Meadows et al. also defined a fifth clade namely HPG E. HPG E was first observed in breeds from Israel and Turkey. In total, 5 haplogroups are defined for sheep until now, with HPG A and HPG B being the most common ones and HPG C, HPG E being the rare haplogroups and HPG D being the least frequent one. All of these haplogroups are observed in modern sheep of Turkey (Meadows 2007, Demirci 2013).

All of these haplogroup assignments were based on control region or cytochrome B region of mtDNA. Meadows (2011) analyzed the whole mitogenome of sheep from all five haplogroups. The study revealed that the control region was indeed one of the most powerful fragments of the mtDNA resolving the relationship of different haplogroups. Further studies on the control region of sheep mtDNA showed that specific sequence motifs within a small region can differentiate A, B and C haplogroups (Luo et al. 2005, Cai et al. 2007). The same motifs are further shown to be useful for identifying the remaining two haplogroups namely HPG D and HPG E (Demirci 2013, Dağtaş 2013) (Table 4).

1.3 Ancient DNA research

Pääbo et al. (2004) defined ancient DNA (aDNA) as the DNA retrieved from museum samples, fossil remains, archaeological findings or other unusual sources like parchments, clothes or even the cave sediments (Slon et al. 2017). Ancient DNA research is a growing field of study since the first isolation of aDNA from the museum specimens of an extinct *Equus quagga* (Higuchi et al. 1984). As more research accumulate on ancient DNA obtained from different species, critical evaluation of results and methods was also needed. The main problem of the ancient DNA research was the ability to assess the authenticity of ancient DNA. In other words, researchers had to be sure that the DNA sequence obtained from the sample belongs to the organism that is in question but not to other organisms whose DNA might have contaminated the sample. Importance of authenticating the results has been obvious when several studies were shown to be unrepeatable (Austin et al. 1997). In order to overcome this problem, researchers studying ancient DNA put strict rules into practice (Cooper and Poinar, 2000). These rules can be summarized as follows:

1. All steps before the PCR amplification of DNA should be carried out in an isolated and clean laboratory dedicated to ancient DNA isolation.
2. Surfaces within the clean laboratory should be wiped with 10% sodium hypochlorite solution and exposed to UV light (254 nm) in order to get rid of DNA remnants.
3. Special full-body clothing must be worn within the clean laboratory. Gloves should be changed or cleaned with 10% sodium hypochlorite solution frequently.
4. Positive controls should be avoided as they pose a further risk of contamination. Negative controls should be implemented in extraction

and PCR preparation steps to check any contamination coming from the air or equipment.

5. Minimum possible number of researchers should have access to the clean aDNA laboratory.
6. For human studies, DNA sequences of researchers should be obtained in order to detect possible contamination from the laboratory personnel.
7. Results should be reproducible in the subsequent DNA extraction of the same specimen. Long fragments of amplified DNA should be evaluated cautiously as the strength of PCR amplification is inversely related to the length of product.

With the advent of next generation sequencing, it became possible to obtain huge amounts of DNA from the ancient specimens. It even became possible to construct whole genome or whole mitochondrial genome from well-preserved samples (Orlando et al. 2015). These include wooly mammoth (Miller et al. 2008), Neanderthal (Green et al. 2010), *Yersinia pestis* (Bos et al. 2011) and Denisovan hominin (Krause et al. 2010, Meyer et al. 2012). It is clear that as the technology advances, more organisms and higher number of individuals will be sequenced with high coverage.

The extent to which we can recover DNA from an ancient sample is limited mainly by the age of the sample. Once the organism is dead, DNA comes under the attack of catabolic enzymes and quickly gets degraded. This harassment is further carried by bacteria and fungi which release degrading enzymes acting on macromolecules. If DNA survives through the enzymatic and microbial degradation, it slowly decomposes due to the chemical processes like oxidation, hydrolysis etc. (Pääbo et al. 2004). The most prominent result of this chemical processes is an increased frequency of C to T substitution within the clones of ancient DNA (Hofreiter, 2001). Reason of this high frequency of C

to T transitions is spontaneous deamination of cytosine into uracil. Uracil pairs with an adenine during a PCR amplification which leads to incorporation of a thymine instead of the original cytosine during the subsequent PCR steps. Uracil N-glycosylase, an enzyme which removes uracil from the DNA molecule, treatment of extracted DNA before the amplification decreases the C to T transition frequency dramatically.

The rate of DNA degradation seems to be dependent on environmental factors namely temperature, humidity and pH (Higgins et al. 2015). Among these factors, temperature is suggested to be the most influential one (Allentoft et al. 2012). Increased temperature is expected to affect the DNA survival negatively. This suggestion can be supported by the fact that the oldest sample from which DNA is obtained was a horse bone retrieved from permafrost (Orlando et al. 2013). The bone dated back to 560-780 thousand years before present. In that sense, Anatolia is expected to have medium survival of ancient DNA due to increased humidity and temperature in coastal regions (Haile et al. 2009).

The power of ancient DNA comes from its ability to directly test phylogenies among the modern and/or extinct populations as well as direct evaluation of the evolutionary forces acting on organisms. Two examples can highlight that power. First one is the discovery of Denisovan hominin. In 2008, a small part of a finger bone of a hominin was discovered in the Denisova Cave of Altai Mountains. Sequencing of the ancient DNA obtained from this bone revealed that the bone does not belong to a modern human but to a new hominin (Krause et al. 2010, Meter et al. 2012). Identification of this hominin would be impossible without advances in ancient DNA retrieval and sequencing methods. Further analysis of the sequenced Denisovan DNA showed 4-6% introgression of Denisovan genetic material into modern humans living in southeast Asia (Reich et al. 2010; Skoglund and Jakobsson, 2011) and the high

altitude adaptation observed in Tibeteans was the result of this introgression (Huerta-Sánchez et al. 2014).

Another example is related to the domestication and westward migration of pigs. Larson et al. (2005) showed that the European domestic pigs possess the mtDNA haplogroup of European wild boar without any affinity to Near Eastern wild boar. This data was interpreted as an evidence for a separate domestication event for pigs within the Europe. Later Larson et al. (2007) and Ottoni et al. (2013) studied 221 and 393 ancient pig specimen, respectively, from Eurasia. Their results suggested that pigs were first domesticated in Near East where the Arm1 haplogroup was dominant. These domestic pigs were introduced into western Anatolia where the mtDNA signature shifted to Y1 haplogroup, most likely as a result of introgression of female boar from the wild. Y1 is the haplogroup that was later introduced into Europe and traces of Y1 haplogroup was found all the way up in the Paris basin. Similar to what happened in western Anatolia, Y1 haplogroup was replaced by the local European haplogroup through the introgression of local wild female boar into domestic herds. Starting from the Late Bronze Age, European haplogroup was observed among the Anatolian domestic pigs and later in 5th century AD, local haplogroups of Y1 and Arm1 were completely replaced by the European haplogroup (Figure 2). Clearly, the history of pig domestication and migration explained here is too complex to be uncovered just by using modern data.

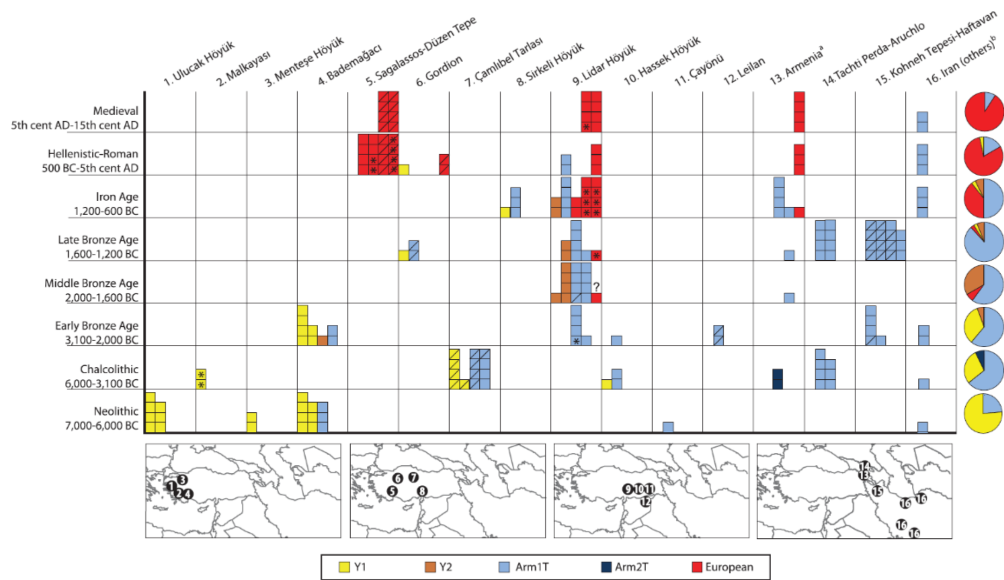


Figure 2 Haplogroup (Y1, Y2, Arm1T, Arm2T, European) distribution of ancient domestic pigs over excavation sites and different time periods. Numbers within the small maps indicate the excavation sites. Figure taken from Ottoni et al. (2013).

1.4 Ancient DNA research on sheep

Although aDNA studies with big sample sizes were conducted for domestic cattle (Edwards 2004, Cai et al. 2014, Scheu et al. 2015) and domestic pig (Larson et al. 2007, Ottoni et al. 2013) studies on sheep were rather limited. Cai et. al. (2007) extracted DNA from 8 samples from China which were dating back to 2100-1800 BCE and all of those samples were identified as HPG A. Later in 2011 the same group added 14 more samples into their analysis and identified a single individual with HPG B. Their results indicate that similar to

modern breeds, HPG A was dominant mtDNA haplogroup for sheep in Asia in Bronze Age.

In 2010, Horsburg and Rhines obtained DNA from 20 samples from South Africa dating back to 2000 before present (BP). All samples belonged to HPG B lineage in their study indicating that these sheep were probably carried from Southwest Asia into South Africa.

In 2013, Niemi et al. successfully amplified DNA from 26 sheep from Finland dating back to Iron Age, Medieval Age and post-Medieval Age. Four of them were assigned to HPG A while the remaining 22 were assigned to HPG B indicating the dominance of HPG B within Europe back in the Iron Age.

For the first time in ancient sheep DNA, a study from Turkey by Demirci et al. (2013) reported the existence of HPG C and HPG E among the 32 ancient samples from Oylum Höyük dating back to 1880-330 BCE. This study shows the importance of samples obtained from the domestication center of sheep in order to be able to understand domestication process comprehensively. If the origin of all haplogroups observed among the modern sheep breed was the center of domestication, then it is expected that rare haplogroups will be lost as the distance from the domestication center increases. This can explain why HPG C and E were observed in ancient sheep of Turkey but not from other regions of the world (Dağtaş, 2013). Furthermore, if some haplogroups were integrated to domestic sheep gene pool after the first domestication event the time of incorporation can also be seen from aDNA analysis.

Rannamäe et. al. (2016) published 102 ancient mtDNA sequences from Estonia dating from 688 BCE to 1900s. Similar to Niemi et. al.'s (2010) study, HPG B was dominant among the ancient Estonian sheep with very rare observation of HPG A.

A recent study by Dymova et al. (2017) reported 17 ancient mtDNA sequences from Altai, dating between 4000-1000 BCE. Their analysis revealed HPG A and HPG B lineages with similar frequencies (41% and 35.3% respectively) whereas some of their samples could not be assigned to a haplogroup within their phylogeny. Therefore, they suggested that the initial diversity of sheep was much higher than previously thought and new haplogroups should be defined to describe this diversity.

1.5 Objectives of the study

Anatolia harbors the domestication center of sheep and acts as a gateway into Europe for migrating communities. Therefore, revealing the genetic make-up of ancient sheep from Anatolia is crucial to the understanding of domestication and subsequent westward migration of sheep. This data can also be used in conjunction with the human ancient genome data to understand the trade routes or migration between the ancient communities.

In the present study, 234 sheep samples from 9 archaeological excavations from Turkey, dating between Epipaleolithic and 1800 BCE and 12 samples from one excavation from Iran were examined for their mtDNA haplogroups. Main objective of the study is to contribute to the understanding of domestication process of sheep by means of an aDNA study. More specifically, aims of this study are as follows:

1. To amplify 144 bp long CR sequence of sheep mtDNA from the samples and to identify their haplogroups in terms of HPG A, HPG B, HPG C, HPG D and HPG E.

2. To see if changes within or among haplogroup diversities in a time wise and space wise manner can be employed as markers of migration and or domestication status of sheep populations.
3. To understand the migration paths of sheep and see if there is a concordance between the domestic sheep and human migration paths across Anatolia for instance during the Neolithic expansion.

In the search of evolutionary history of domestic sheep within Anatolia, to check the possible continuity of some of the sheep populations under consideration.

CHAPTER 2

MATERIALS AND METHODS

2.1 Material

Sheep bone samples used in the study were collected mainly by Dr. Füsun Özer, Nihan Dilşad Dağtaş, Dr. Eren Yüncü or provided by the directors and zooarchaeologists of the excavations. All bone samples were kept at +4ºC in the ancient DNA laboratory in order to minimize the risk of further damage to DNA. In the present study, a total of 234 samples from 9 excavation sites were examined. Also haplogroup information of 12 samples from Tepe Khaleseh excavation from Iran was used for comparative analysis. DNA from these samples were extracted, amplified and analyzed by Foad Abazari at Togan Lab (Abazari, 2017). Excavations from which the samples were collected are shown in Figure 3. Details of the samples used in this study were provided in Table 5 and in Appendix A.

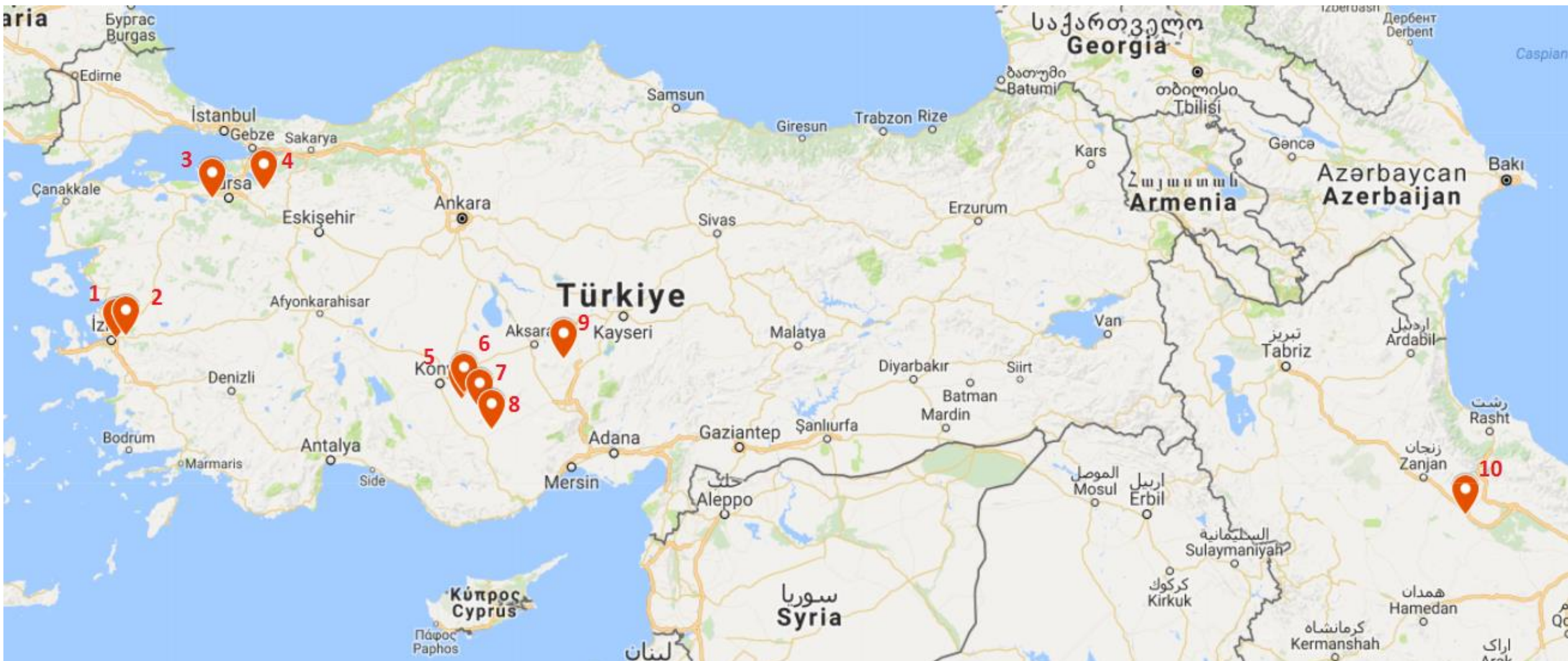


Figure 3. Map of Anatolia with locations of excavations from which samples are collected. 1. Yeşilova Höyük, 2. Ulucak Höyük, 3. Aktopraklık Höyük, 4. Barcın Höyük, 5. Çatal Höyük, 6. Boncuklu Höyük, 7. Pınarbaşı Höyük 8. Canhasan III Höyük, 9. Tepecik Çiftlik Höyüğü, 10. Tepe Khaleseh

2.2 Laboratory Experiments

2.2.1. Working in dedicated clean ancient DNA laboratory and precautions to prevent contamination

DNA extraction from bone samples and preparation of PCR reactions were performed in a dedicated ancient DNA laboratory. Although humans are not an important source of contamination for sheep DNA research, extensive care was given in order to produce reliable results. Before working in the clean lab, UV lights were turned on for overnight in order to get rid of any DNA remained on the surfaces as well on mortars and pestles that were used for grinding. Next day, right after entering the airlock of the clean laboratory, clothes that were only used within the laboratory were worn. Then after wearing a pair of regular non-sterile examination gloves, a face mask, a beard mask, a bonnet and protective clothing covering the whole body excluding the face was worn. When all the protective clothes were put on, a second pair of gloves were worn and hands were wiped with bleach. During the laboratory work hands and surfaces were wiped frequently with bleach and Thermo Scientific™ DNA AWAY™ Surface Decontaminant to prevent exogenous DNA contamination.

2.2.2. Extraction of aDNA

DNA extractions were carried out using the protocol by Dabney et al. (2013), with slight modifications. From each sample, 85 to 120 mg of bone powder was used to extract DNA. In order to obtain bone powder first a thin layer from the surface of the bone was removed and then with the help of a drill (Proxxon Micromot 40/E) a small part of the bone was cut and grounded into a fine powder in a clean, UV-exposed mortar with the help of a pestle. These

procedures were conducted in a class 2A Biosafety Cabinet (MetiSafe) The cabinet and all other tools were decontaminated with UV prior to the procedure. A sheet of aluminum foil was put on the surface of the biosafety cabinet and replaced with a new sheet each time a new bone was processed. The bone powder was weighted with the help of a small piece of aluminum foil and then transferred into screw cap tubes that were labeled with the ID of the sample, the date of extraction and the name of the person who performs the extraction. The remaining bone was put into sample bag together with the disc used for cutting the bone. Aluminum foil was discarded and the surface of the cabinet was cleaned with DNA AWAY. Second layer of gloves were removed off the hands and a new pair of gloves were worn for each sample.

One ml of the extraction buffer (0.45 M EDTA, 0.25 mg/ml Proteinase K, pH: 8.0) was added to each tube with bone powder. Then tubes were incubated overnight in a rotating incubator at 37°C. The following day, tubes were centrifuged at maximum speed (16000 rpm) for 2 minutes. Supernatant was transferred into a 15 ml conical centrifuge tube without disrupting the pellet. For the binding step, 13 ml of binding buffer (Table 1) was added to the same 15 ml centrifuge tube. A 20 ml extension reservoir (Zymo Research) was fitted forcefully onto a Qiaquick spin column. The extension reservoir – column assembly is placed into a 50 ml centrifuge tube (which will be called 'binding assembly' from here on). The mixture of supernatant and binding buffer was poured into extension reservoir. The binding assembly was centrifuged 4 minutes at 1000 g, rotated 90° by using a swing-bucket rotor and centrifuged for 2 more minutes. Fifty ml centrifuge tube was removed from the assembly; the spin column was separated from the reservoir and then placed into a 2 ml collection tube. Spin columns were centrifuged 1 min at 6000 rpm (dry-spin) in a bench-top microcentrifuge. Flow-through within the collection tube was discarded. Then 750 μL of washing buffer (Qiagen PE buffer) was added to spin columns, centrifuged at 3300 g and the flow-through was discarded. This

step was repeated one more time. After discarding the flow-through in the second washing step, tubes were centrifuged 1 minute at 16000 rpm (dry-spin). Spin columns were removed from collection tubes and placed into 1,5 ml microcentrifuge tubes. In each spin column, 12,5 µL of elution buffer (Qiagen EB buffer) is dispensed on the filter. After 2 – 5 minutes of incubation, tubes were centrifuged at maximum speed for 30 second. The collection step was repeated after adding 12.5 µL of elution buffer, yielding a final volume of 25 ml DNA extract. Tubes were kept at -20ºC unless they are used immediately for PCR solution preparation.

Table 1. Preparation of binding buffer

Chemical	Final Concentration
Guanidine Hydrochloride	5 M
Isopropanol	40% (vol/vol)
Tween – 20	0.05%
Sodium Acetate Buffer (pH 5.2)	90 mM

2.2.3. DNA amplification

Polymerase Chain Reaction (PCR) was used for the enrichment of target fragment of ancient mtDNA. Pair of primers were used in order to amplify a part of $tRNA^{pro}$ and the control region of mtDNA (Figure 4). Primers

were designed by Cai et al. (2007) from the reference sequence AF010406 and employed successfully in previous studies in our laboratory (Demirci et al 2013, Dağtaş 2013).

mtDNA site	Primer	Primer sequence	Fragment length (bp)
$tRNA^{Pro}$ - HVRI	L15391[a]	5′-CCACTATCAACACCCAAAG-3′	144
	H15534	5′-AAGTCCGTGTTGTATGTTTG-3′	

Figure 4. Sequences of forward (L15391) and reverse (H15534) primers. Figure taken from Cai et al. (2007).

PCR mixture was prepared with 5 µL of DNA and 15 µL of master mix, yielding a final volume of 20 µL. Details of the PCR master mix reagents is shown in Table 2.

Table 2. Reagents for PCR master mix.

Molecular Grade Water	7.4 µL
Buffer (10X)	2.0 µL
Mg^{2+} (25 mM)	1.6 µL
Forward Primer (10 µM)	1.5 µL
Reverse Primer (10 µM)	1.5 µL
AmpliTaq Gold 360 (5 U/ µL)	0.6 µL
dNTP (25 mM)	0.2 µL
BSA (10 mg/ml)	0.2 µL

Amplification of DNA was performed by using a thermal cycler (The Applied Biosystems Veriti Thermal Cycler) with the PCR conditions specified in Table 3.

Table 3. PCR amplification conditions

Phase	Temperature	Duration	
Initial Denaturation	94ºC	10 min	
Denaturation	94 ºC	30 sec	60 cycles
Annealing	53 ºC	45 sec	60 cycles
Extension	72 ºC	45 sec	60 cycles
Final Extention	72 ºC	5 min	
Base Temperature	4 ºC	∞	

2.2.4. Agarose gel electrophoresis

In order to check whether the target 144 bp long region of mtDNA was amplified successfully or not, PCR products were run on 2% agarose gel (1.6 g of agarose, 80 ml 0.5X TBE buffer and 4 µL Ethidium Bromide). Five ml PCR product was mixed with 5 ml 1X Loading Dye (diluted from Thermo Scientific 6X DNA Loading Dye) and loaded on the gel. 50bp DNA Ladder (GeneRuler, Thermo Scientific) was loaded on the first and last wells of the gel as a reference. Gel was run at 110 V for 45 min and bands were visualized under UV light using Vilber imaging system (CN-3000 WL; Vilber-Lourmat, Torcy, France).

2.2.5. Purification of amplified DNA and sequencing

PCR products with positive results for 144 bp long mtDNA fragment were purified with GeneJET PCR Purification Kit (Thermo Fisher Scientific, USA) by following manufacturer's instructions. Purification was performed to get rid of remnants of PCR such as buffer salts, ions, polymerase and excess dNTPs.

Briefly, binding buffer was added to PCR reaction mixture in 1:1 volume ratio and mixed by pipetting. An orange color indicated optimum pH value for the mixture. This mix was transferred into GeneJet Purification Column, centrifuged for 1 min and flow through was discarded. Then, 700 µL washing buffer was added to the column, centrifuged for 1 min and flow through was discarded. Empty columns were centrifuged for an additional 1 min (dry-spin) in order to get rid of any residual ethanol which may interfere with the elution step. The column was placed into a clean 1.5 ml microcentrifuge tube. DNA was eluted with 20 µL of elution buffer. Eluted DNA is kept in +4°C until sequencing.

Sanger Sequencing was performed on ABI PRISM® 3100 Genetic Analyzer by RefGen, Gen Araştırmaları ve Biyoteknoloji Company. Primers used for sequencing were the same with the pair used in PCR amplification.

2.3 Analyses

2.3.1. Alignment and editing of sequences

Raw sequence files with ".ab1" extension were produced by ABI PRISM® 3100 Genetic Analyzer. These files were examined by using the software Geneious version 8.1.8 ((http://www.geneious.com, Kearse et al., 2012). All reads were imported into the program and then ends of forward and reverse sequences for each sample were trimmed due to low quality of base assignments. After trimming, these pair of reads were aligned together to a reference sequence AF010406.1 (Hiendleder et al., 1998a) and any inconsistency was checked manually by visually inspecting the peaks of chromatogram. Consensus sequence for each sample was obtained from alignments and extracted as a separate file. These consensus sequences were further aligned to the reference and then used in analysis.

2.3.2. Assignment of haplogroups

The strength of the 144 bp long target DNA sequence comes from its ability to differentiate all 5 haplogroups observed in modern domestic sheep. On Table 4, five base positions that are used for HPG assignment are shown (Demirci et al. 2013). Haplogroup of each individual was assigned by visually checking these 5 nucleotide positions at each consensus sequences obtained from the previous steps.

Table 4. mtDNA control region nucleotide positions used for haplogroup assignment. Table adapted from Demirci et al. (2013)

Haplogroup	Reference Sequence	Positions (on AF010406) and identities of bases used for haplogroup assignment				
		15459	15476	15484	15509	15512
HPG A	HM236174	T	T	A	A	T
HPG B	HM236176	C	T	G	A	T
HPG C	HM236178	C	T	G	G	T
HPG D	HM236180	C	T	G	A	C
HPG E	HM236182	C	C	G	G	T

2.3.3. Genetic diversity measurements

2.3.3.1. Nucleotide diversity (π)

Nucleotide diversity, π, is defined by Nei (1987) as the average number of nucleotide differences per site between two sequences. The formula for nucleotide diversity (Nei 1987) is given in the Equation 2.1.

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

Equation 2.1

In the equation 2.1. n is the total number of sequences and π_{ij} is the proportion of nucleotide differences per site between the ith and jth sequences. Combinatorial $\binom{n}{2}$ represents the total number of pairwise comparisons.

Nucleotide diversity and the standard deviation for nucleotide diversity is was calculated by using the software DnaSP ver5 (Librado and Rozas, 2009).) software.

Statistical significance of difference between the nucleotide diversity measurements were assessed by Student's t test.

2.3.3.2. Haplotype diversity

Haplotype diversity is another commonly used measurement and it represents the probability that within a set of samples, two randomly chosen sequences are different from each other (Nei, 1987). The formula for haplotype diversity (Nei, 1987) is given in the Equation 2.2.

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

Equation 2.2

In equation 2.2. n is the total number of sequences. x_i is the frequency of the ith haplotype within the population.

Haplotype diversity and the standard deviation for haplotype diversity is calculated by using the software DnaSP ver5 (Librado and Rozas, 2009).

Statistical significance of difference between the haplotype diversity measurements were assessed by Student's t test.

2.3.3.1. Haplogroup diversity

Shannon-Wiener Diversity Index (Shannon and Weaver, 1949) was used to calculate haplogroup diversity. Shannon-Wiener index is commonly used to evaluate the species richness in biodiversity studies. In the present study, it was adopted to calculate the haplogroup diversity of sheep populations. The formula for Shannon-Wiener Index (Shannon and Weaver, 1949) is given in the Equation 2.3.

$$H = -\sum p_i \ln(p_i)$$

Equation 2.3

In equation 2.3. p_i represents the frequency of a species which corresponds to haplogroups in the present study.

Statistical significance of difference between the haplogroup diversity measurements were assessed by a permutation test. A Python script was written and used for permutation test.

2.3.4. Coalescent simulations of population continuity

In order to understand whether the genetic drift and mutation alone can explain the observed differentiation of two populations, coalescent simulations were produced using fastsimcoal2 software. For mutation rate and population size, a range of parameters were used because the mutation rate for sheep mtDNA and the effective population size is not known. Mutation rate range of

10^{-8} – 10^{-6} per generation per base pair was used whereas the effective population range of 50-1500 individuals (Leroy et al. 2013) was used. For each pair of population, 1000 simulations were run for every parameter combination to produce pairwise F_{ST} values. F_{ST} is a measure of population differentiation ranging between 0 (completely panmictic populations) and 1 (completely isolated and differentiated populations). F_{ST} values obtained from simulations were then compared with the observed F_{ST} value as calculated by using Arlequin 3.5.2 (Excoffier and Lischer 2010). Proportions of simulated F_{ST} values which were higher than the observed F_{ST} were plotted on a heatmap by using the Octave software.

CHAPTER 3

RESULTS

3.1 Ancient DNA extraction and amplification

Ancient DNA extraction was performed by Onur Özer for Barcın, Canhasan 3, Boncuklu and Pınarbaşı excavations. On the other hand, aDNA of Tepecik Çiftlik, Çatalhöyük and Aktopraklık samples were extracted by Nihan Dilşad Dağtaş Kılıç and Dr. Füsun Özer whereas aDNA of samples from Ulucak and Yeşilova were extracted by Dr. Eren Yüncü (Table 5). mtDNA haplogroup information of sheep from Tepe Khalese excavation of Iran was obtained from the thesis of Foad Abazari (Abazari, 2017). After the aDNA extraction, 144 bp long fragment of mtDNA was amplified via Polymerase Chain Reaction (PCR). Some of the isolates did not reveal the full length amplification. Some isolates were contaminated by foreign DNA which generally belonged to microbes even before the samples were unearthed. In order to check whether the target fragment was amplified successfully or not, PCR products were visualized on 2% agarose. An example of an agarose gel image can be seen in Figure 5. After visualization, PCR products that have the band of 144 bp long fragment were sent for sequencing with both forward and reverse primers to RefGen, Gen Araştırmaları ve Biyoteknoloji Company. Remaining samples were not sequenced. The total success rate for aDNA

retrieval with sequences was 52% with the highest rate at Canhasan Höyük (100%) and the lowest rate at Aktopraklık Höyük (20%).

Table 5. Excavation sites, number of samples available from the sites and number of samples with successfully amplified and sequenced aDNA.

Excavation Site	Number of samples studied	Number of samples with successfully amplified and sequenced aDNA
Yeşilova	24	9 (37,5%)
Ulucak	40	15 (37,5%)
Barcın	53	27 (51%)
Aktopraklık	8	2 (25%)
Canhasan III	5	5 (100%)
Çatalhöyük	13	5 (38%)
Tepecik Çiftlik	67	48 (72%)
Boncuklu	5	2 (40%)
Pınarbaşı	17	8 (47%)
Tepe Khalese	-	12
TOTAL	232	121 (52%) + 12

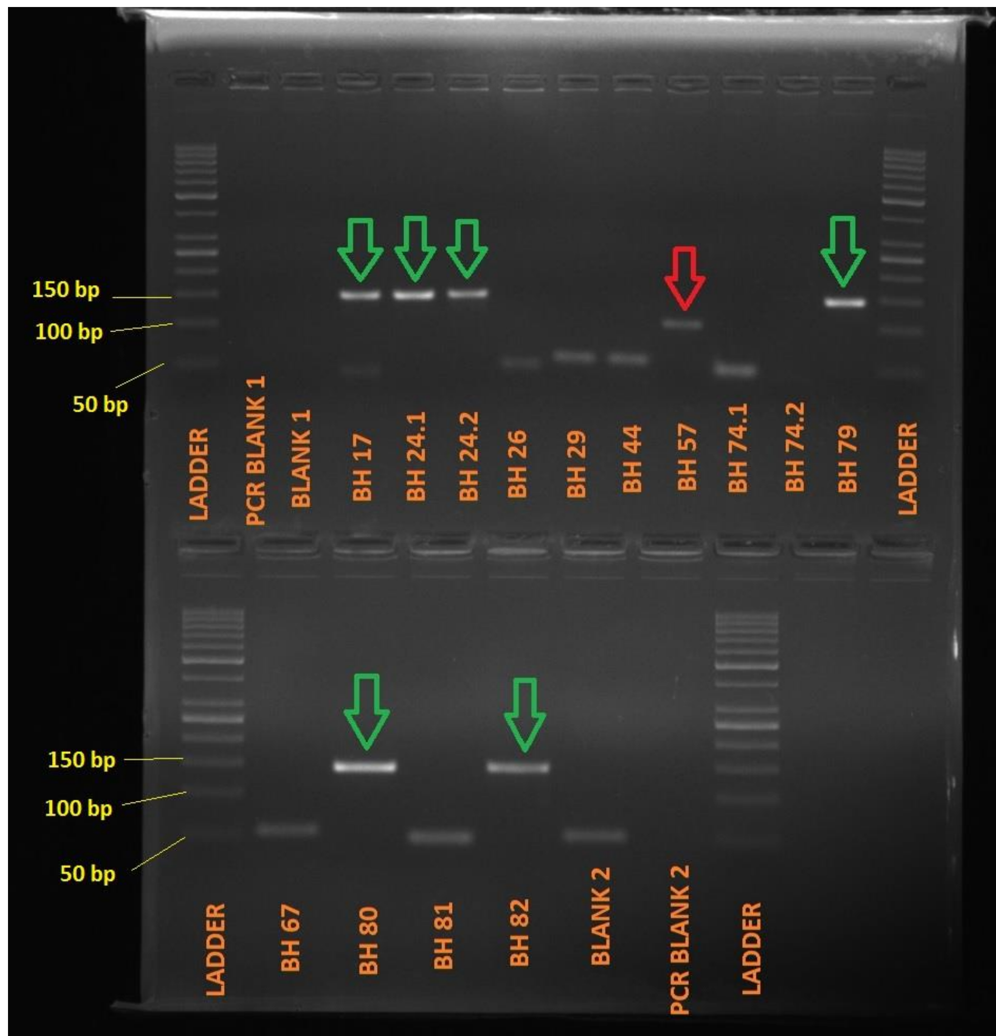


Figure 5. Image of an agarose gel. Green arrows indicate successfully amplified 144 bp long fragment of sheep mtDNA. Red arrow shows a band of non-specific amplification.

3.2 Editing and aligning sequences

Raw sequence files produced by ABI PRISM® 3100 Genetic Analyzer are examined by using the software Geneious version 8.1.8 ((http://www.geneious.com, Kearse et al. 2012). An example of a chromatogram obtained from these files can be seen in Figure 6. Bases with low quality readings at the ends of the reads are trimmed. After trimming, reads of the same sample obtained by reverse and forward primers are aligned to a reference and any inconsistency is checked manually by evaluating the peaks of chromatogram (Figure 7).

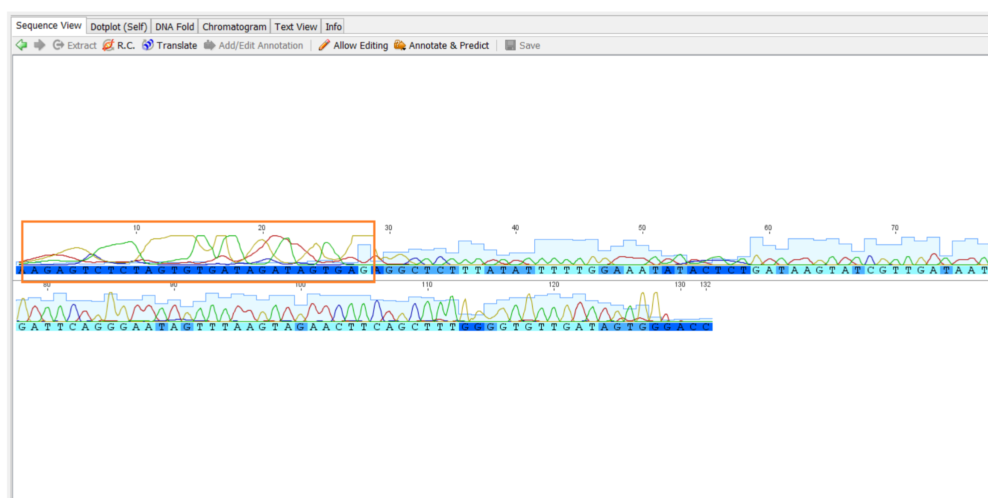


Figure 6. Chromatogram of the sample BH24, sequenced with reverse primer. Orange box indicates the region of low quality base assignments at the end of the read.

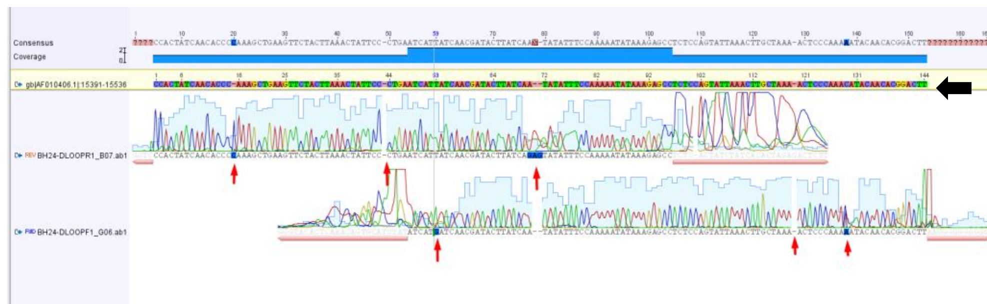


Figure 7. Alignment of reverse and forward reads to the reference sequence. Reference sequence is shown by black arrow. Pink boxes at the ends of the reads indicate the trimmed regions with low quality bases. Red arrows indicate regions that are inconsistent with the reference and the complementary read.

Consensus sequence for each sample was obtained from alignments and extracted as a separate file. These consensus sequences were further aligned to the reference and then used in analysis (Appendix A).

3.3 mtDNA haplogroup determination and spatial-temporal distribution of sheep haplogroups

Haplogroups were determined according to the identity of bases at 5 specific positions (Table 4.). Haplogroups were than considered at each site at different time periods. Time periods were arbitrarily considered as 500 years. B Haplogroup was observed as the major haplogroup for central and western Anatolia (Table 6). It was the most prominent one at each site in each time period except at Yeşilova in the time period 6500-6000 BCE. No individual with mtDNA haplogroup C was observed among our samples. Another interesting point was the observation of three individuals from Ulucak Höyük

(UH7, UH26, UH30) with HPG identity in between HPG A and HPG B. Considering the 5 positions used for HPG determination, these individuals have a T at position 15459 which was indicative of HPG A and a G at position 15484 which was indicative of HPG B (Figure 8). Identity of bases were validated by sequencing these individuals at least 2 times separately (Appendix A).

Table 6. mtDNA haplogroup compositions of domestic sheep from Anatolia

	Yeşilova	Ulucak	Barcin	Aktopraklık	Canhasan	Catalhoyuk	Tepecik - Çiftlik	Boncuklu	Pınarbaşı	Tepe Khalese
Middle Bronze 2500-1800 BC										
Early Bronze 3000-2500 BC										
Chalcolithic 3500-3000 BC										
Chalcolithic 3800-3500 BC										
Chalcolithic 5500-3800 BC										
Neolithic 6000-5500 BC										
Neolithic 6500-6000 BC										
Neolithic 7000-6500 BC										
Neolithic 7500-7000 BC										
Neolithic >7500 BC										
Total Observation	9	15	27	2	5	5	48	2	8	12

HPG A

HPG B

HPG C

HPG D

HPG E

HPG A/B

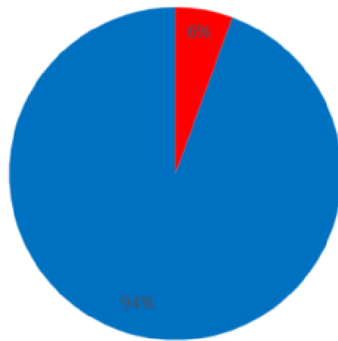
Figure 8. Partial mtDNA sequence of individuals with ambiguous HPG identity. Individuals in the first two row (having HPG B) and last two row (having HPG A) are displayed for comparison. Pink box on the left shows the position 15459, blue box on the right indicates position 15484.

3.4 mtDNA diversity calculations

3.4.1 mtDNA haplogroup diversity across Neolithic Period for central and western Anatolia

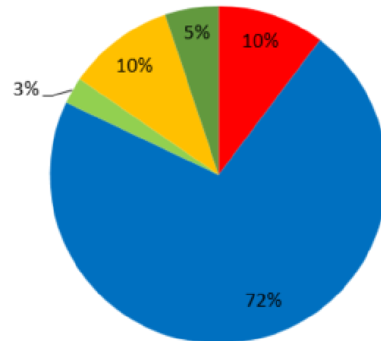
In order to understand changes in mtDNA haplogroup diversity of sheep across the Neolithic Period, Shannon Diversity Index was calculated for four time periods: Epipaleolithic – 7000 BCE, 7000 BCE – 6500 BCE, 6500 BCE – 6000 BCE, 6000 BCE – 2800 BCE (Table 7) and frequencies of haplogroups were summarized by using pie charts in Figure 9. Haplogroup diversity indices for the same time periods were given in Table 7 in terms of Shannon Diversity index values.

Epipaeolithic - 7000 BC



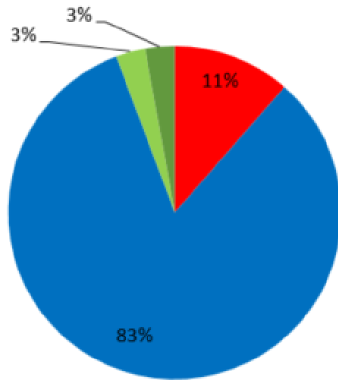
A B C D E Other

7000 BC - 6500 BC



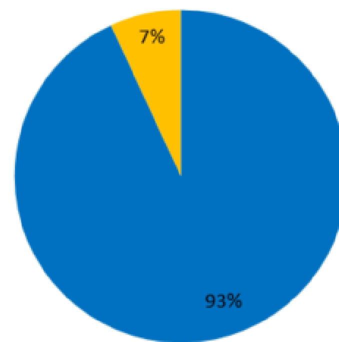
A B C D E Other

6500 BC - 6000 BC



A B C D E Other

6000 BC - 2800 BC



A B C D E Other

Figure 9. Pie charts showing the haplogroup composition of domestic sheep for different time periods for central and western Anatolia.

Table 7. Shannon Diversity Indices for specified time periods, number of samples (n) used and their excavation sites. YH (Yeşilova Höyük), UH (Ulucak Höyük), BH (Barcın Höyük), AT (Aktopraklık Höyük), CH (Canhasan III), CT (Çatalhöyük), TP (Tepecik Çiftlik Höyüğü), BK (Boncuklu Höyük), PB (Pınarbaşı Höyük)

Sample age and total number of samples	Excavation sites and number of samples	Shannon Diversity Index
Epipaleolithic - 7000 BCE n=18	CH(n=5), CT (n=2), TP (n=5), BH (n=2), PB (n=4)	0.214
7000 BCE – 6500 BCE n=39	UH (n=8), CT (n=1), TP (n=26), PB (n=4)	0.951
6500 BCE – 6000 BCE n=35	YH (n=4), UH (n=6), BH (n=19), AT (n=2), CT (n=2), TP (n=2)	0.606
6000 BCE – 2800 BCE n=29	YH (n=5), UH (n=1), BH (n=8), TP (n=15)	0.250

Shannon diversity indices of Epipaleolithic – 7000 BCE and the successive period 7000 BCE – 6500 BCE differed significantly ($p<0.05$) (Figure 10). There was a significant increase in haplogroups between the two considered time intervals. None of the successive periods differed significantly after 7000 BCE based on their haplogroups (Figure 11 and Figure 12). Yet, their diversity kept decreasing consistently and felled to a level of Epipaleolithic - 7000 BCE. When the compositions were examined it started with 94% HPG B with 6% HPG A in Epipaleolithic – 7000 BCE and returned

back to 93% HPG B with 7% HPG E in the period 6000 – 2800 BCE in western-central Anatolia.

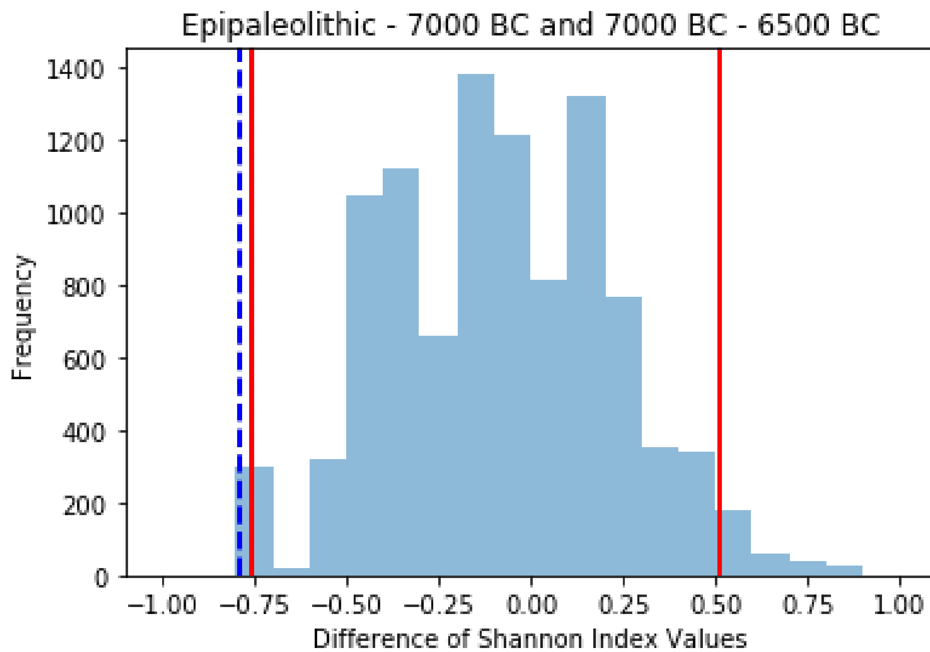


Figure 10. Significance test with 10000 permutations for the difference between Shannon diversity indices of Epipaleolithic – 7000 BCE and the successive period 7000 BCE – 6500 BCE. Red line indicates the value where p = 0.05. Blue line indicates the observed difference.

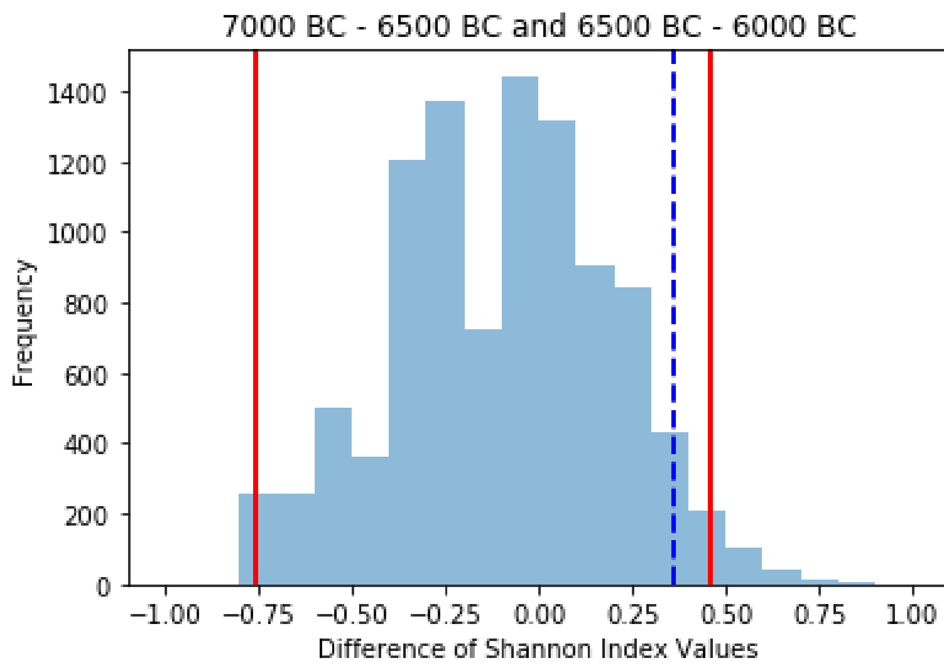


Figure 11. Significance test with 10000 permutations for the difference between Shannon diversity indices of 7000 BCE – 6500 BCE and the successive period 6500 BCE – 6000 BCE. Red line indicates the value where p = 0.05. Blue line indicates the observed difference.

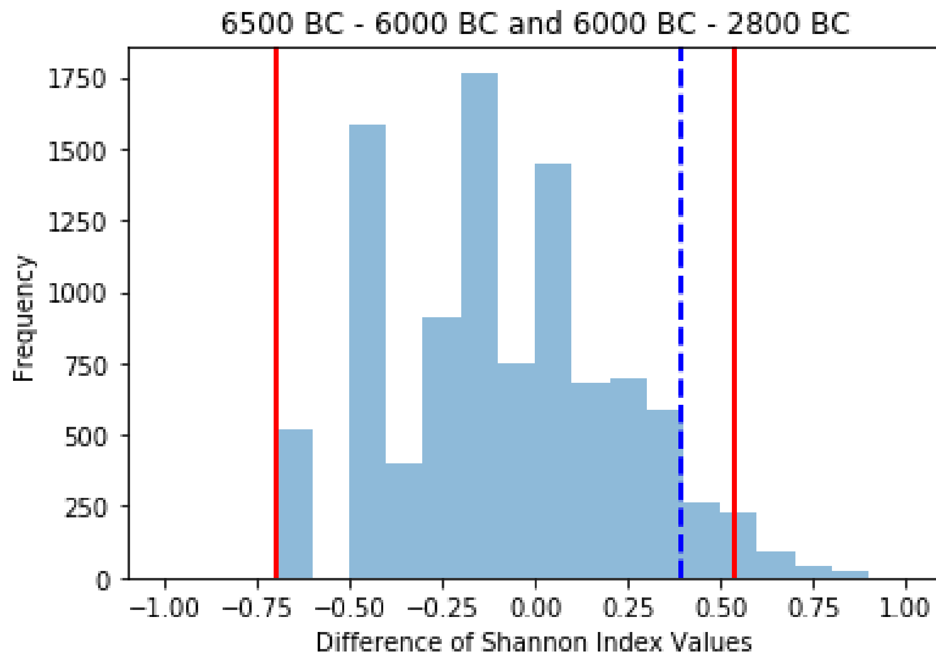


Figure 12. Significance test with 10000 permutations for the difference between Shannon diversity indices of 6500 BCE – 6000 BCE and the successive period 6000 BCE – 2800 BCE. Red line indicates the value where p = 0.05. Blue line indicates the observed difference.

3.4.2. Molecular diversity measurements

Since a significant change was observed for haplogroup diversity before and after 7000 BCE, haplotype and nucleotide diversity was calculated for individuals having B (the most prominent) haplogroup before and after 7000 BCE. To avoid obtaining false results due to mutations caused by post-mortem damage, only individuals with at least two sequence reads were used (n=59) (Appendix A). This criterion further limited the already low number of individuals older than 7000 BCE. Due to this limitation, HPG B samples of

Canhasan III (n=3), Pınarbaşı (n=3), Boncuklu (n=1), Tepecik (n=1) were pooled for diversity calculations These sites were all from Central Anatolia. The pooling of samples from different excavation sites may result in an artificially increased diversity. In order to account for that possible bias, individuals dating between 7000 BCE – 6500 BCE from Ulucak (n=5) and Tepecik (n=11) (now the two sites were from two different geographic regions) and individuals younger than 6500 BCE from Ulucak (n=6), Yeşilova (n=3), Barcın (n=13), Tepecik (n=12) (samples were from 3 different geographic regions) were also pooled for comparisons in calculations. Results were summarized in Table 8. Nucleotide and haplotype diversity measurements were the highest for the Epipaleolithic – 7000 BC samples while they were the lowest for 7000 BC – 6500 BC samples. Significance of difference between the haplotype and diversity values were evaluated by Student's t test. Difference was significant between Epipaleolithic – 7000 BC samples and 7000 BC – 6500 BC samples ($p<0.001$) whereas the difference between 7000 BC – 6500 BC samples and 6500 BC – 6000 BC samples was not significant.

Table 8. Molecular diversity measurements of individuals before and after 7000 BCE. h: Number of haplotypes, Hd: Haplotype Diversity, Pi: Nucleotide Diversity

Sites	Dates	Number of samples	Total number of samples	h	Hd	pi
Canhasan III	7400 BCE – 7100 BCE	3	9	4	0,694 ± 0,147	0,00733 ± 0,00245
Pınarbaşı	Epipaleolithic – 7500 BCE	4				
Boncuklu	8400 BCE – 7800 BCE	1				
Tepecik	7500 BCE – 7000 BCE	1				
Ulucak	7000 BCE – 5600 BCE	5	21	3	0,186 ± 0,110	0,00476 ± 0,00289
Tepecik	6850 BCE – 5800 BCE	16				
Ulucak	6500 BCE – 5600 BCE	6	34	5	0,275 ± 0,099	0,00334 ± 0,00168
Barcın	6500 BCE – 6100 BCE	13				
Yeşilova	6250 BCE – 5800 BCE	3				
Tepecik	6500 BCE – 5800 BCE	12				

3.1 Coalescent simulations for population continuity

To contribute to the understanding of origins of domestic sheep in the Western Anatolia, continuity simulations were used. With these simulations, it was asked if the observed differentiation between populations can be explained by chance events, i.e. mutation and genetic drift. A positive answer to this question implies that two populations that are being compared can be continuous. Otherwise, other evolutionary forces like migration, natural selection should be considered to explain the observed differentiation. Broad intervals; 10^{-8} – 10^{-6} for mutation rate and 50 – 1500 for effective population size for the assumed parameters were used. Time difference between populations was calculated from the mean of ages of samples and a generation time of 2 years was assumed for the sheep. By using fastsimcoal2 coalescent simulator, 1000 simulations are produced for each pair of parameters with the assumption of no population growth. Another set of simulations were produced with the scenario where $1/5^{th}$ of the old population forms the new population and the new population expands exponentially into the population size specified by the parameter. F_{ST} was calculated for simulated populations (Table 9) and the proportion of F_{ST} values that are greater than the observed F_{ST} was plotted on a heatmap graph. If the calculated proportion is less than 0.05, the continuity was rejected for that pair of parameters. All F_{ST} calculations were done by using Arlequin ver 3.5.2.2 (Excoffier and Lischer, 2010). Computer codes used for simulations, calculations and plotting were provided by Assist. Prof. Dr. Ayşegül Birand and Dr. Sinan Can Açan. Overall results were summarized in Table 10. The simulations were repeated with the assumption of 4 years of generation time and results were given in Appendix B.

Table 9. Parameters between populations used for simulations. Values within brackets indicate sample sizes

Test Populations	Generation	F_{st}	P value of F_{st}
TP1 (n=24) – TP2 (n=13)	413	0.01206	0.31106 (ns)
TP1 (n=24) – UH (n=15)	119	0.05299	0.04089 (*)
TP1 (n=24) – YH (n=8)	361	0.10911	0.03673 (*)
TP1 (n=24) – BH1 (n=17)	208	0.01135	0.29641 (ns)
UH1 (n=7) – BH1 (n=17)	220	0.12545	0.01208(*)
UH (n=15) – YH (n=8)	142	0.04108	0.11108 (ns)
BH1 (n=17) – BH2 (n=6)	1630	-0.08008	0.99990 (ns)

ns: not significant, *: significant

TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population

TP 2: Tepecik Çiftlik (6000 BCE – 5500 BCE) population

YH: Yeşilova Höyük (6500 BCE – 5500 BCE) population

UH 1: Ulucak Höyük (7000 BCE – 6500 BCE) population

UH: Ulucak Höyük (7000 BCE – 5500 BCE) population

BH 1: Barcın Höyük (6500 BCE – 6000 BCE) population

BH 2: Barcın Höyük (3800 BCE – 1800 BCE) population.

3.5.1 Simulations with assumption of no growth

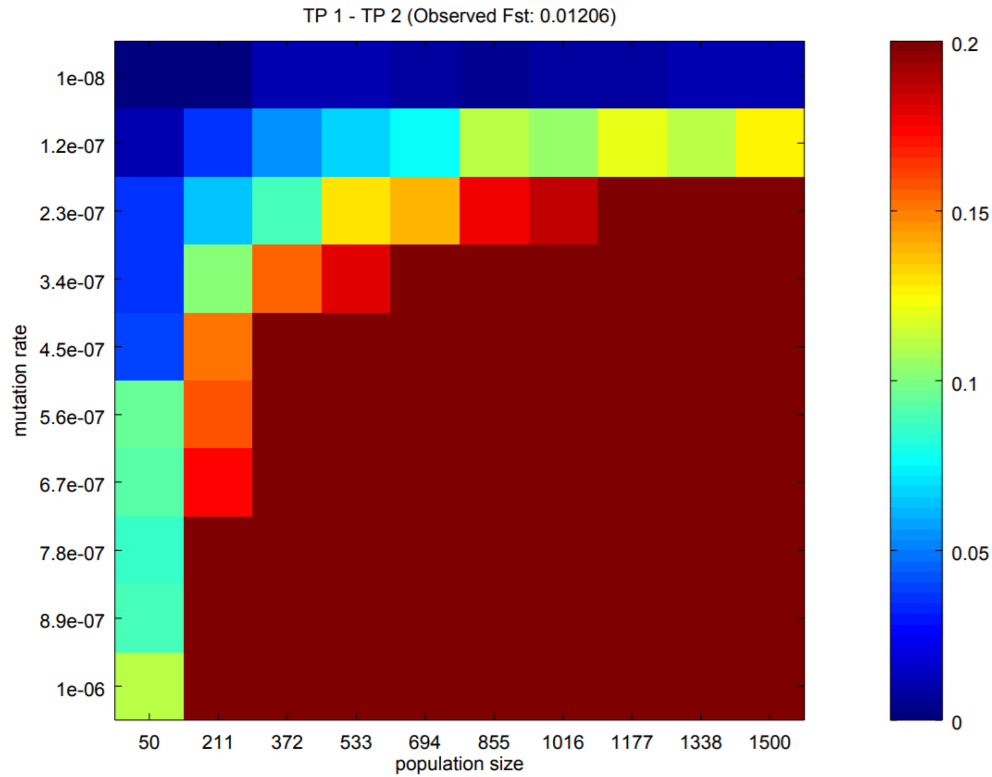


Figure 13. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and late Tepecik Çiftlik (TP 2) populations. Heatmap is colored according to the scale on the right of the figure. TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population, TP 2: Tepecik Çiftlik (6000 BCE – 5500 BCE) population

According to Figure 13 F_{ST} value observed between early Tepecik Çiftlik (TP 1) and late Tepecik Çiftlik (TP 2) populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 15 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Tepecik Çiftlik and late Tepecik Çiftlik populations can be rejected.

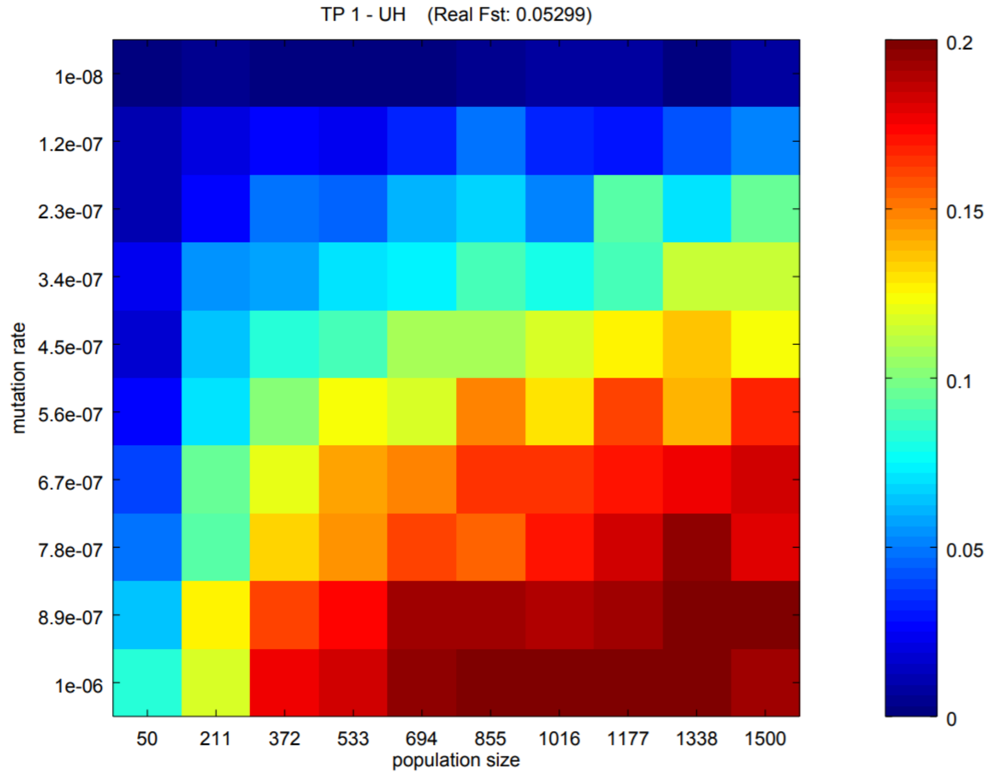


Figure 14. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and Ulucak Höyük (UH) Populations. Heatmap is colored according to the scale on the right of the figure. TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population, UH: Ulucak Höyük (7000 BCE – 5500 BCE) population.

According to Figure 14 F_{ST} value observed early Tepecik Çiftlik and Ulucak Höyük populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 28 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Tepecik Çiftlik and Ulucak Höyük populations can be rejected.

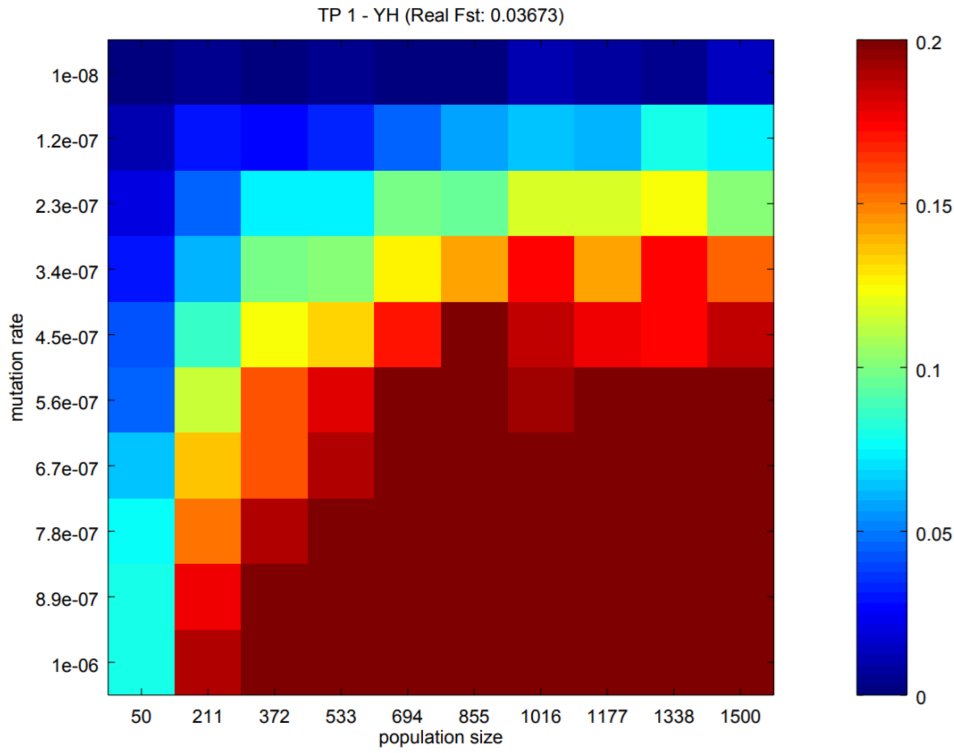


Figure 15. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and Yeşilova Höyük (YH) Populations. Heatmap is colored according to the scale on the right of the figure. TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population, YH: Yeşilova Höyük (6500 BCE – 5500 BCE) population.

According to Figure 15 F_{ST} value observed early Tepecik Çiftlik and Yeşilova Höyük populations is in the expected ($p > 0.05$) region for many of the pairs of. However, for 20 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Tepecik Çiftlik and Yeşilova Höyük populations can be rejected.

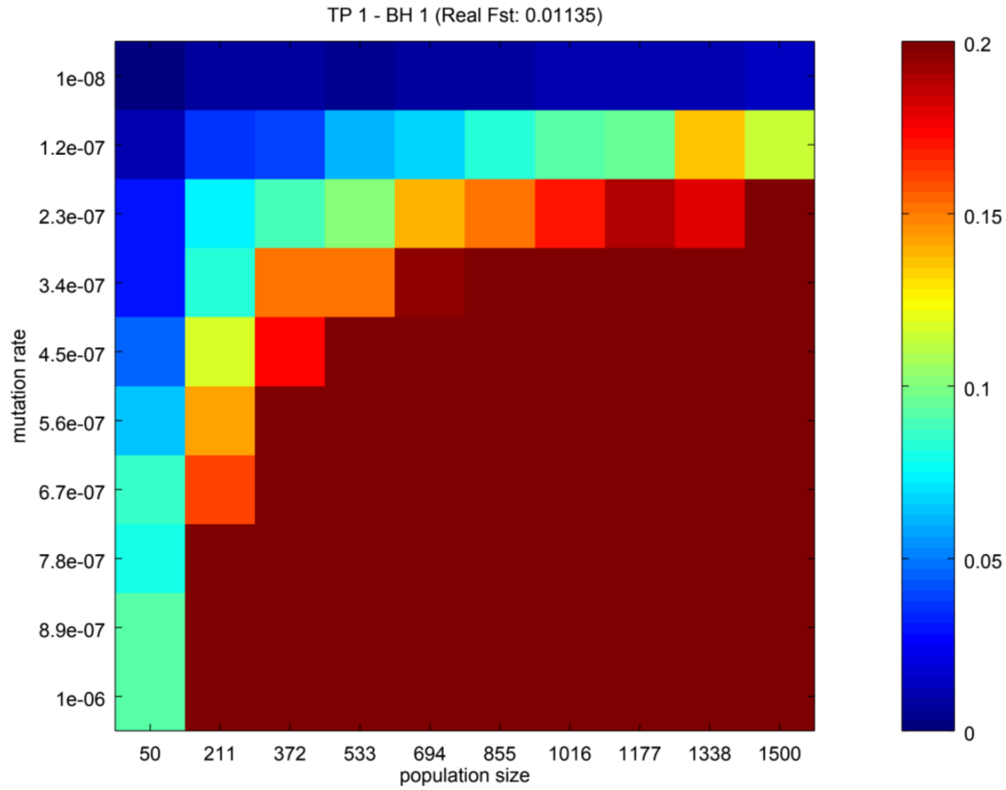


Figure 16. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and early Barcın Höyük (BH 1) populations. Heatmap is colored according to the scale on the right of the figure. TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population. BH 1: Barcın Höyük (6500 BCE – 6000 BCE) population

According to Figure 16 F_{ST} value observed early Tepecik Çiftlik and early Barcın Höyük populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 15 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Tepecik Çiftlik and early Barcın Höyük populations can be rejected.

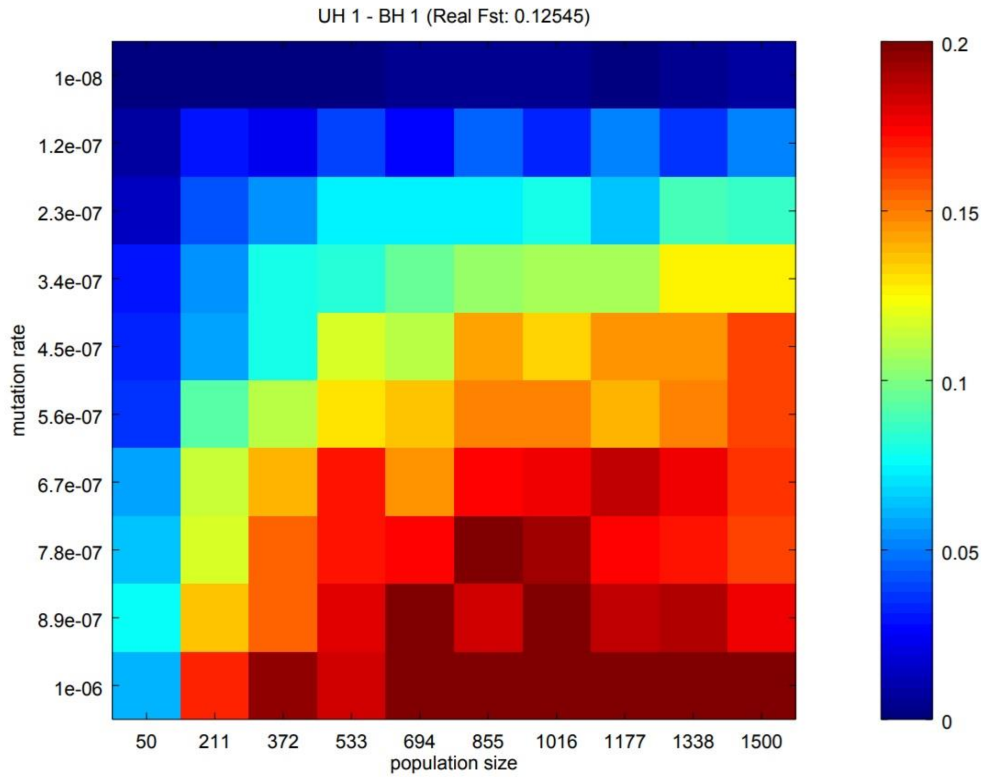


Figure 17. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Ulucak Höyük (UH 1) and early Barcın Höyük (BH 1) Populations. Heatmap is colored according to the scale on the right of the figure. UH 1: Ulucak Höyük (7000 BCE – 6500 BCE), BH 1: Barcın Höyük (6500 BCE – 6000 BCE) population.

According to Figure 17 F_{ST} value observed between early Ulucak Höyük and early Barcın Höyük Populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 23 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Ulucak Höyük and early Barcın Höyük Populations can be rejected.

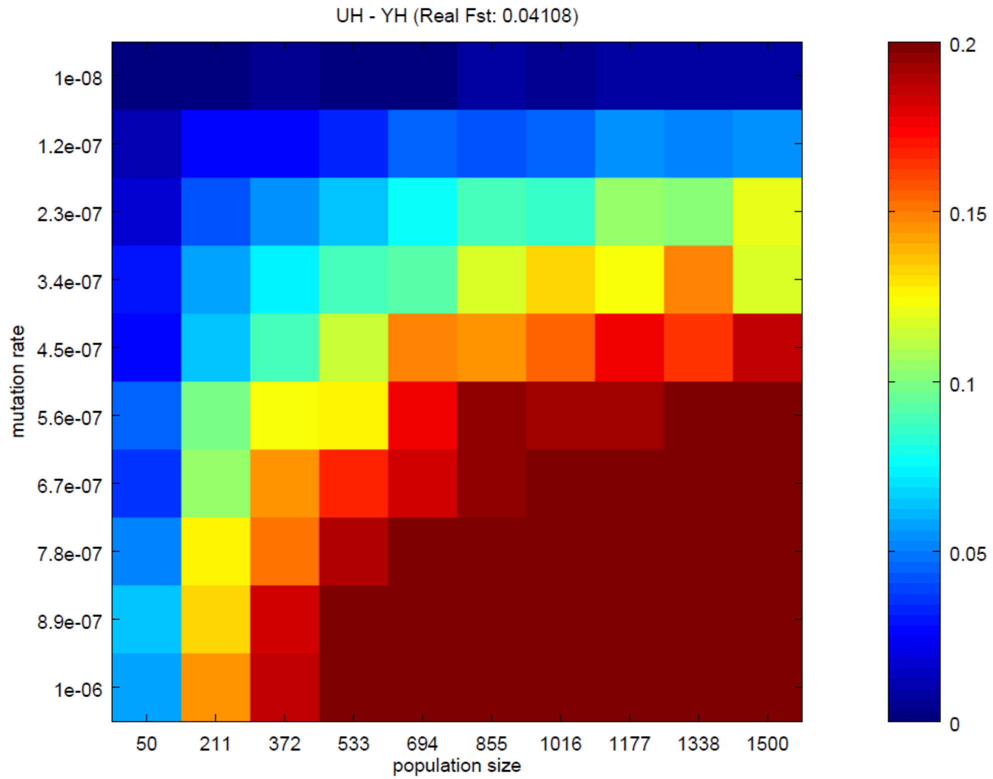


Figure 18. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Ulucak Höyük (UH) and Yeşilova Höyük (YH) Populations. Heatmap is colored according to the scale on the right of the figure. UH: Ulucak Höyük (7000 BCE – 5500 BCE, YH: Yeşilova Höyük (6500 BCE – 5500 BCE) population.

According to Figure 18 F_{ST} value observed between early Ulucak Höyük and Yeşilova Höyük Populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 23 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Ulucak Höyük and Yeşilova Höyük Populations can be rejected.

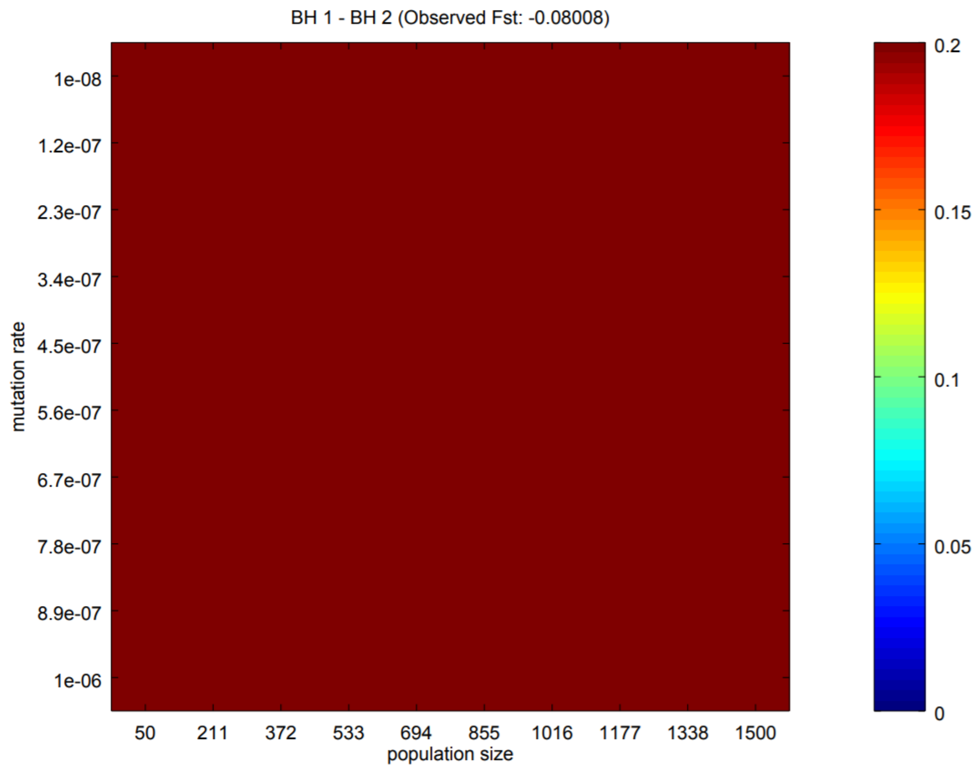


Figure 19. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early (BH 1) and (BH 2) late Barcın Höyük Populations. Heatmap is colored according to the scale on the right of the figure. BH 1: Barcın Höyük (6500 BCE – 6000 BCE) population, BH 2: Barcın Höyük (3800 BCE – 1800 BCE) population.

According to Figure 19 F_{ST} value observed between populations is in the expected ($p > 0.05$) region for all pairs of parameters. Therefore, continuity cannot be rejected between early and late Barcın Höyük populations.

3.5.2 Simulations with assumption of bottleneck and exponential growth

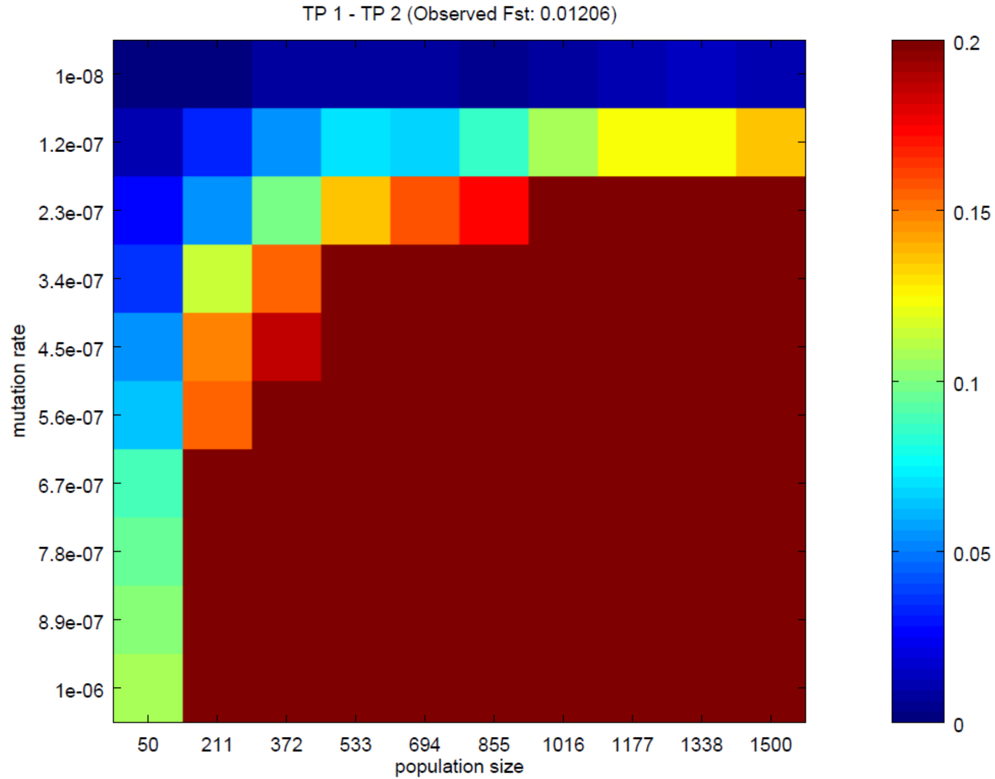


Figure 20. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 1) and late Tepecik Çiftlik (TP 2) populations. Heatmap is colored according to the scale on the right of the figure. TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population, TP 2: Tepecik Çiftlik (6000 BCE – 5500 BCE) population

According to Figure 20 F_{ST} value observed between early Tepecik Çiftlik (TP1) and late Tepecik Çiftlik (TP 2) populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 14 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Tepecik Çiftlik and late Tepecik Çiftlik populations can be rejected.

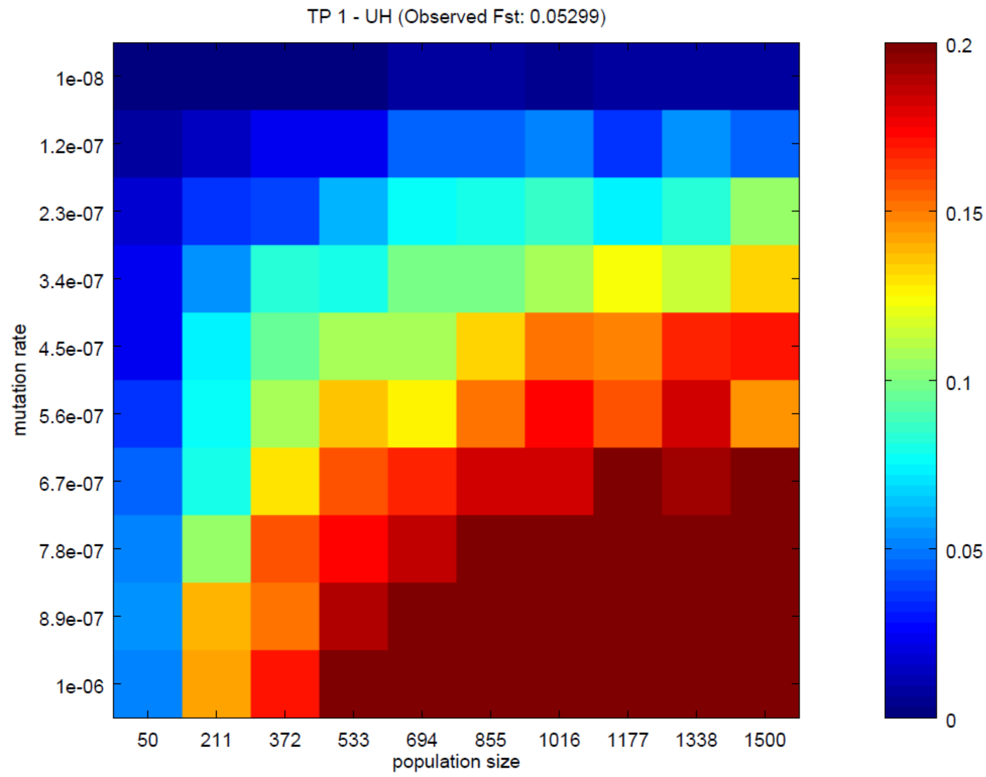


Figure 21. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik and Ulucak Höyük Populations. Heatmap is colored according to the scale on the right of the figure. TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population, UH: Ulucak Höyük (7000 BCE – 5500 BCE) population.

According to Figure 21 F_{ST} value observed early Tepecik Çiftlik (TP 2) and Ulucak Höyük (UH) populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 25 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Tepecik Çiftlik and Ulucak Höyük populations can be rejected.

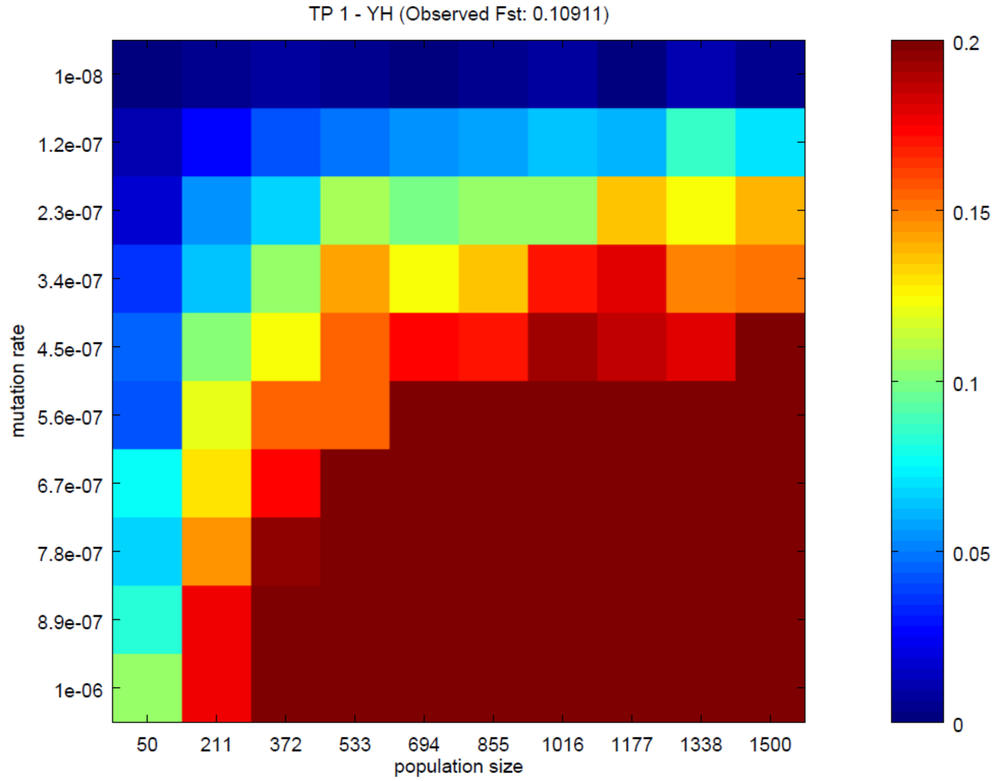


Figure 22. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik and Yeşilova Höyük Populations. Heatmap is colored according to the scale on the right of the figure. TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population, YH: Yeşilova Höyük (6500 BCE – 5500 BCE) population.

According to Figure 22 F_{ST} value observed early Tepecik Çiftlik (TP 2) and Yeşilova Höyük (YH) populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 18 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Tepecik Çiftlik and Yeşilova Höyük populations can be rejected.

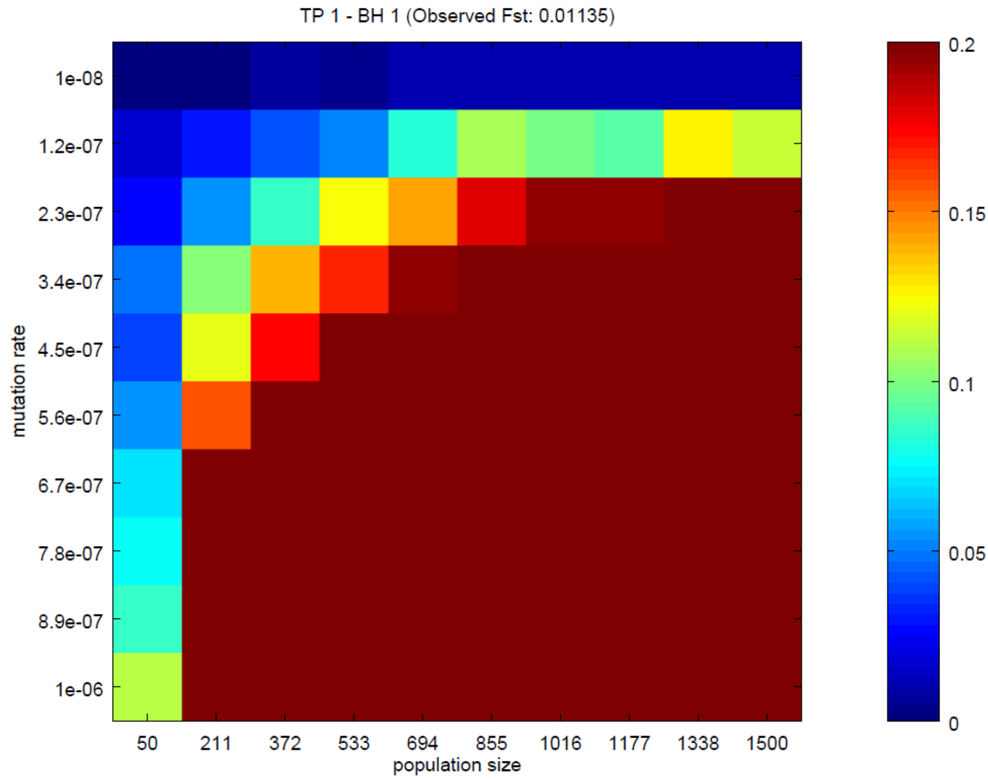


Figure 23. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Tepecik Çiftlik (TP 2) and early Barcın Höyük (BH 1) populations. Heatmap is colored according to the scale on the right of the figure. TP 1: Tepecik Çiftlik (7000 BCE – 6500 BCE) population. BH 1: Barcın Höyük (6500 BCE – 6000 BCE) population

According to Figure 23 F_{ST} value observed early Tepecik Çiftlik and early Barcın Höyük populations is in the expected ($p > 0.05$) region for many of the pairs of parameters. However, for 16 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Tepecik Çiftlik and early Barcın Höyük populations can be rejected.

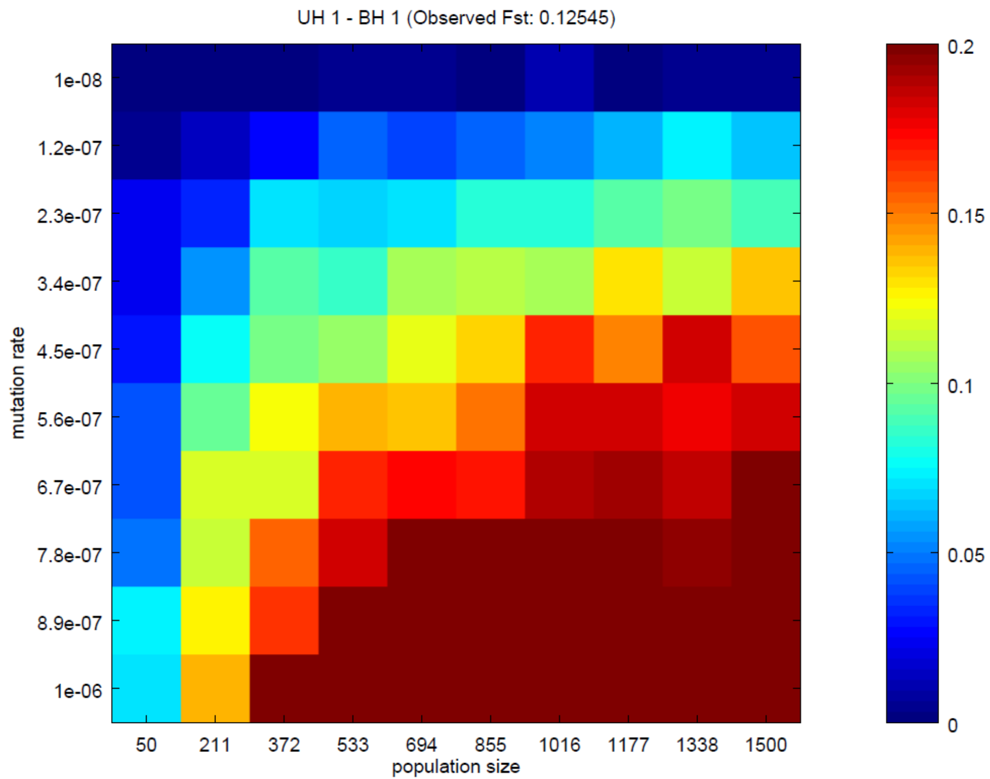


Figure 24. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Ulucak Höyük (UH 1) and early Barcın Höyük (BH 1) Populations. Heatmap is colored according to the scale on the right of the figure. UH 1: Ulucak Höyük (7000 BCE – 6500 BCE), BH 1: Barcın Höyük (6500 BCE – 6000 BCE) population.

According to Figure 24 F_{ST} value observed between early Ulucak Höyük and early Barcın Höyük Populations is in the expected ($p > 0.05$) region for many of the pairs of parameter. However, for 23 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Ulucak Höyük and early Barcın Höyük Populations can be rejected.

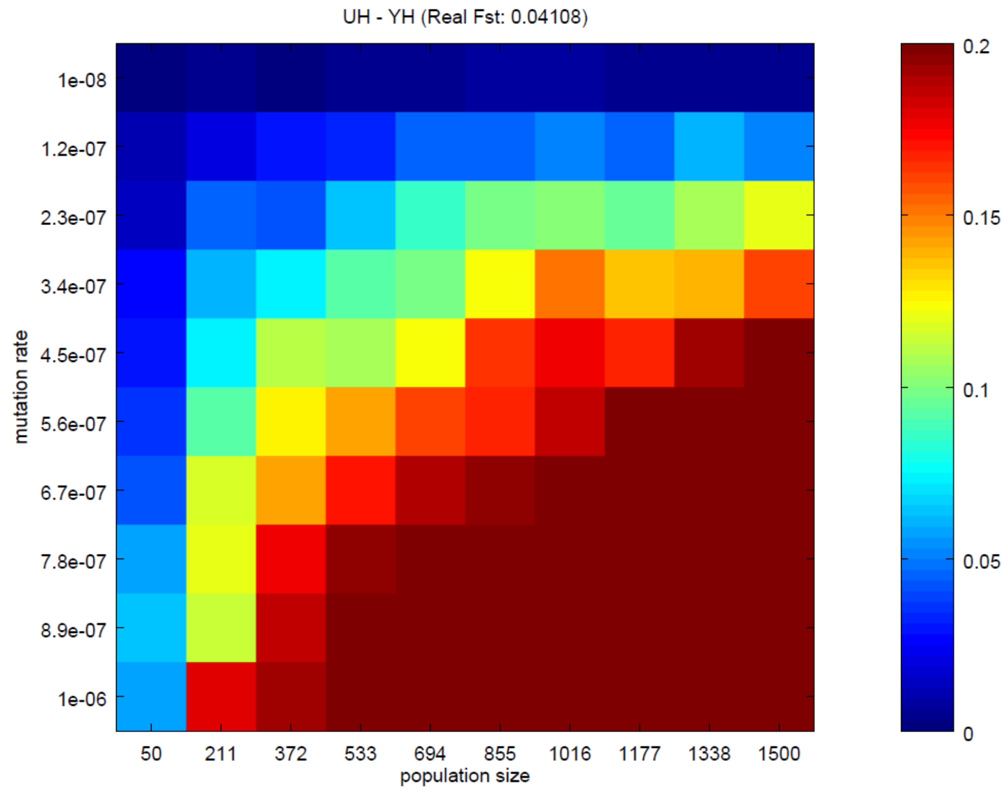


Figure 25. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early Ulucak Höyük (UH) and Yeşilova Höyük (YH) Populations. Heatmap is colored according to the scale on the right of the figure. UH: Ulucak Höyük (7000 BCE – 5500 BCE, YH: Yeşilova Höyük (6500 BCE – 5500 BCE) population.

According to Figure 25 F_{ST} value observed between early Ulucak Höyük and Yeşilova Höyük Populations is in the expected ($p > 0.05$) region for many of the pairs of parameter. However, for 24 of the parameter pairs probability of observing the real F_{ST} is unlikely ($p < 0.05$). Therefore, it can be argued that for low mutation rates and low effective population sizes, continuity between early Ulucak Höyük and Yeşilova Höyük Populations can be rejected.

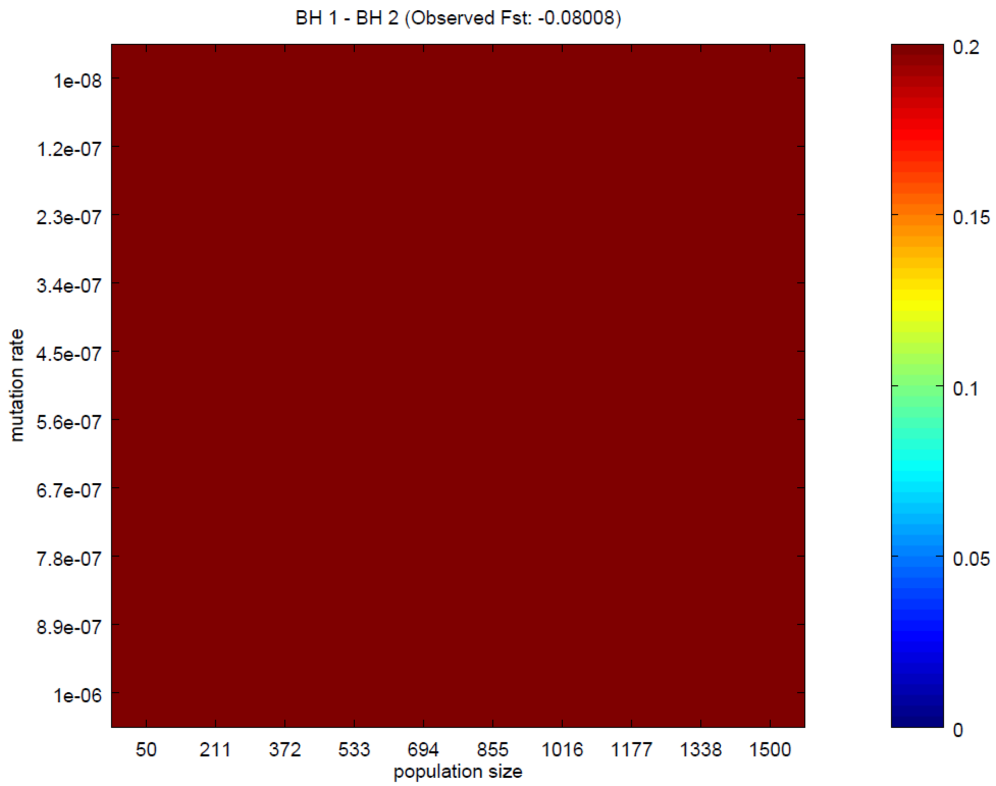


Figure 26. Heatmap showing the proportion of F_{ST} values that are greater than the real F_{ST} between early (BH 1) and late (BH 2) Barcın Höyük Populations. Heatmap is colored according to the scale on the right of the figure. BH 1: Barcın Höyük (6500 BCE – 6000 BCE) population, BH 2: Barcın Höyük (3800 BCE – 1800 BCE) population.

According to Figure 26 F_{ST} value observed between populations is in the expected ($p > 0.05$) region for all pairs of parameters. Therefore, continuity cannot be rejected between early and late Barcın Höyük populations.

Table 10. Frequency of simulated populations with indicated significance values over all mutation rate and population size values.

Pairwise Comparison	No growth			Bottleneck and exponential growth		
	p<0.05	p<0.1	p>0.2	p<0.05	p<0.1	p>0.2
TP1 - TP2	0.15	0.24	0.6	0.14	0.23	0.62
TP1 - BH1	0.15	0.25	0.56	0.16	0.26	0.58
BH1 - BH2	0	0	1	0	0	1
TP1 - UH	0.28	0.50	0.06	0.25	0.44	0.20
TP1 - YH	0.20	0.36	0.35	0.18	0.31	0.37
UH1 - BH1	0.23	0.44	0.09	0.23	0.42	0.21
UH – YH	0.23	0.41	0.26	0.24	0.39	0.28

Table 10 summarizes the results obtained from heat-map graphs. In order to calculate frequencies, the number of each parameter pair (mutation rate - effective population size) that yielded a significance value $p<0.05$, $p<0.1$ or $p>0.2$ was calculated.

CHAPTER 4

DISCUSSION

In the present study, 144 bp long fragment of mtDNA from ancient sheep samples of Barcın Höyük, Pınarbaşı Höyük, Boncuklu Höyük, Canhasan III Höyük was amplified. Obtained sequences were analyzed together with the data from Ulucak Höyük, Yeşilova Höyük, Aktopraklık Höyük, Çatal Höyük, Tepecik Çiftlik Höyük and Tepe Khalese (Iran).

4.1. Establishment of DNA extraction protocol

Previously, extraction method used in the Rohland et al.'s (2009) study was employed for DNA extraction from ancient sheep samples within our laboratory. Starting with the Barcın Höyük samples a new protocol used by Dabney et al. (2013) study is employed. This method is shown to be more effective in retrieving very short (<40 bp) DNA fragments. Given that the target fragment of this study is 144 bp long, it may seem futile to use a method that is more successful in obtaining short fragments. However, as a next step of this study, NGS libraries will be produced and SNP genotyping will be performed for some of the samples. Therefore, establishing a method which increases the efficiency of obtaining short fragments is required for the project (TÜBİTAK Grant no: 114Z356).

Adopting a new method requires some modifications in order to optimize the yield in a new laboratory. This optimization is needed due to the differences between the samples (age, tissue type, preservation status etc.) as well as equipment (brand and the material of equipment etc.) used by different laboratories. Optimization is performed by using the input from all members (Dr. Füsun Özer, Dr. Eren Yüncü, Nihan Dilşad Dağtaş) of our laboratory.

4.2. Spatial and temporal distribution and genetic diversities of mtDNA haplogroups in central and western Anatolia

Consistent with the modern sheep data (Demirci et al., 2013) HPG B was the most common haplogroup observed in ancient central and western Anatolian sheep. This pattern remains the same for each excavation site (except Yesilova in the west) studied in this research.

Through personal communication with zooarchaeologists of Tepecik Çiftlik, Can Hasan III, Boncuklu and Pınarbaşı excavations, it was confirmed that all samples that were older than 7000 BCE exhibited morphological characteristics of wild sheep. Whereas, in the present study, except one in Can Hasan III, all these individuals (n=17) possessed HPG B. If these individuals were assumed as wild or being at the early stages of herding as it was seen in Aşıklı (Stiner et al. 2014), then we may suggest that local wild sheep of central Anatolia was mainly composed by individuals possessing HPG B. Of course this proposition must be tested by further researches, mainly due to the lack of HPG B in modern wild Anatolian sheep *Ovis gmelinii anatolica*. (Demirci et al., 2013).

Right after the 7000 BC, the number of haplogroups observed in central and in western Anatolia increased dramatically. This increase was also reflected in the Shannon Diversity Index calculations such that the index jumps

to 0.945 from 0.223, and the difference in the indices was significant ($p<0.05$). Additionally, 7000 BC boundary coincided with the initial observation of morphologically domestic sheep in central and western Anatolia (Arbuckle 2014). Therefore, it can be argued that the increase in the number of observed haplogroups as well as the frequency of rare haplogroups (A, D, E, A/B), at least partly, was the result of introduction of domestic sheep into the region from elsewhere.

In modern domestic sheep, HPG A is mainly distributed in Asia (Singh et al, 2013). High proportion of HPG A sheep in Tepe Khalese (Abazari, 2017) shows that HPG A was already more frequent in the east of southeastern Anatolia back in between 7000 BCE and 6500.

In Arbuckle et al.'s (2014) study based on archaeozoological evidence it was suggested that there was a migration of domestic sheep to the Lakes district and western Anatolia through the southern Anatolian (maritime) route nearly 6800 BCE as summarized in Figure 27. Results of the present study are in concordance with Arbuckle et al.'s (2014) study. Perhaps, the very first observation of HPG A among Canhasan III samples dating (7400-7100 BCE) was one of the first early signatures of this migration. Later, with the new sheep migrations or trade in central Anatolia represented by Tepecik Ciftlik 7000-6500 BCE samples, HPG A frequency increased. A more striking observation is the haplogroup composition of Yeşilova Höyük Samples. At the beginning of the settlement, frequency of HPG A was 75 % (3/4). Low sample size for this time period restricts us from inferring strong conclusions but it was possible to explain this high number of HPG A individuals by referring to "maritime route" hypothesis (Zeder, 2008). Appearance of all four domestic animals (sheep, goat, pig and cattle) together with domestic plants on Cyprus around 8500 BCE indicates that during the initial diffusion of Neolithic Culture out of the Fertile Crescent, people were capable of building boats that could

carry even big animals to long distances (Vigne 2014). Taking advantage of this technology, people might have rapidly colonized favorable locations across the Mediterranean (Zilhao 2001). It is logical to assume that Aegean costs of the western Anatolia were also colonized by these seafaring colonists (Figure 27). In that case, sheep with HPG A may have been brought by these colonists to Yeşilova Höyük, supporting Arbuckle et al's (2014) study, again.

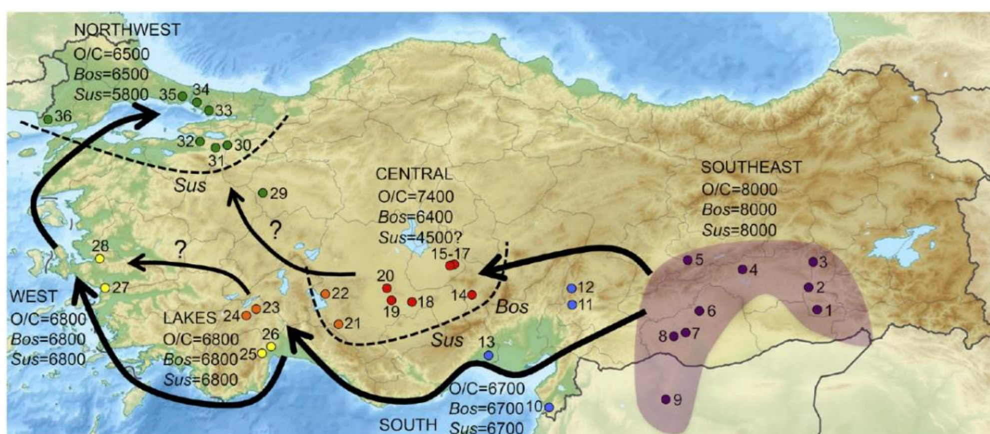


Figure 27. Map of Turkey showing possible routes for human mediated sheep dispersal with black arrows. Dots with numbers represent sites studied in Arbuckle et al. (2014). Dates are the approximate times (BCE) of initial appearance of domestic sheep/goat (O/C), cattle (Bos) and pig (Sus). Figure taken from Arbuckle et al. 2014.

In a recent study, Kılınç et al (2016) showed that genetic diversity in humans of Boncuklu Höyük dating back to nearly 8000 BCE was relatively low. Whereas 1500 year later, Tepecik Çiftlik and Barcın Höyük populations appeared to be more diverse than Boncuklu Höyük population. One of the suggested reasons for the observed increase in diversity was a migration from southern and eastern parts of Anatolia into Central Anatolia. Our results were in agreement with the observed increase in genetic diversity of central and western Anatolia human populations.

As well as HPG A, HPG E and HPG A/B (endemic to Ulucak) contributed to the HPG diversity increase of central and western Anatolian sheep gene pool. As have been noticed, between 7000 BCE and 6500 BCE the most frequent haplogroups after HPG B (72%) was HPG E (10%) and HPG A (10%). Lack of HPG E and HPG A/B from the Tepe Khalese samples supports the idea that the origin of HPG E may not be close to the east of Anatolia.

Regarding the HPG A/B 3 individuals with HPG identity in between HPG A and HPG B (denoted as HPG A/B) were observed among Ulucak Höyük samples. These individuals represented a haplogroup that was not observed in modern domestic sheep. Two of these individuals (UH30 and UH7) had an unusually high number of C > T transitions (5 and 4 respectively) in positions out of the HPG determining positions. High number of transitions in these individuals raised the possibility that the T observed at position 15459 could be the result of a post-mortem C > T transition. If this was the case, then these individuals should be considered as HPG B. On the other hand, 4 of these mentioned transitions were at the same positions of DNA sequences. Moreover, it was unlikely that these two bones belonged to the same individual because the archaeological context from which they were taken was different and there was a nearly 500-year time difference between their ages. Therefore, it was possible that these three individuals were representatives of the lineage that was

not present in modern domestic sheep. This hypothesis can be tested by amplifying a longer fragment of control region and checking more nucleotide positions for these individuals.

Increase in HPG diversity can also be seen if individuals of some other wild sheep gene pools harboring different HPG's were introgressed into the already existing gene pool of the sheep. Interestingly, Arbuckle suggested an "experimental" caprine exploitation strategy for Suberde (Arbuckle, 2008a) and Erbaba (Arbuckle, 2008b) excavation sites in Lakes Region. Archaeometric analysis of sheep remains from these sites indicated that herding of domestic sheep was performed together with hunting of wild sheep. Moreover, Ottoni el al. (2013) showed that the eastern Anatolian mtDNA haplotype has completely been replaced by the mtDNA haplotype of local wild boar after the introduction of domestic pig into western Anatolia. Therefore, as an extension of this "experimental" exploitation strategy, we suggest that females captured from the wild may have been introgressed into domestic sheep herds. Since, archaeological evidence suggests that the Neolithic culture has reached to Ulucak Höyük from the Lakes Region (Derin, 2005). This introgression may be the source of the observed new lineage in Ulucak Höyük. In order to test the hypothesis, samples from Bademağacı Höyük were requested from Dr. Bea De Cupere (archaeobotanist of Bademağacı Höyük). Unfortunately, none of the 16 provided samples yielded ancient DNA. Currently, some samples from Suberde Höyük will reach to our laboratory to be identified and used in analysis. These samples may help us to test our hypothesis related to the role of Lakes Region in sheep domestication.

HPG E is another candidate which could have stemmed from Lakes Region but as mentioned before, currently we lack data for testing hypothesis about the contribution of Lakes Region to domestic sheep. However, compared to HPG A/B, HPG E seemed to have a wider range of distribution among the

modern sheep: HPG E is observed in central and southern Anatolia as well as in Israel (Meadows et al. 2007) and China (Guo et al. 2005). Therefore, it might be migrated somewhere from southeast Anatolia/Levant/Caucuses from where we do not have data yet.

Yeşilova Höyük and Ulucak Höyük are separated from each other by approximately 12 km. Despite this very small geographic distance, it is intriguing to observe no HPG A in Ulucak Höyük and similarly no HPG E and HPG A/B in Yeşilova Höyük. As mentioned before, Ulucak Höyük has a cultural connection with Lakes Region. Difference in sheep haplogroup composition between these two very close sites may be considered as a sign of different domestication products carried by seafaring communities (represented by Yeşilova Höyük) and immigrants using the land (represented by Ulucak Höyük) as depicted in Figure 27.

Value of Shannon's Diversity Index drops to 0.606 after 6500 BCE and further decreases to 0.25 after 6000 BCE. Two explanations can be suggested for the observed decrease in the number and frequency of haplogroups. Firstly, rare haplogroups are expected to be lost due to genetic drift unless there is a continuous supply of rare haplogroups from the outside of the region. This effect should be stronger if the effective population size of domestic sheep is small. Another explanation can be the introgression of wild sheep into domestic flocks. As mentioned before, our data shows that the wild sheep of central Anatolia was mainly composed of individuals with B haplogroup. If ewes were taken from wild and incorporated into domestic flocks, the frequency of haplogroups other than B would decline.

HPG D is the rarest haplogroup among the modern sheep breeds. Only a few individuals from north Caucasus (Tapio et al. 2006), from Turkey (Meadows et al. 2007) and from Tibet (Liu et al. 2016) were assigned to HPG D. It was interesting to observe that two individuals among the 124 successfully

amplified samples had HPG D. One of these individuals were from Tepecik Çiftlik and the other was from Pınarbaşı Höyük, two sites that are both in the central Anatolia (Figure 3). Therefore, according to our data central Anatolia can be the region where HPG D individuals were originated or acquired possibly from Caucasus.

Absence of HPG C from the study populations indicated that the HPG C was either not present or present at very low frequencies during the initial dispersal of domestic sheep into central and western Anatolia. Individuals with HPG C identity were observed in almost all breeds of modern Anatolia studied today. These observations are in agreement with the relatively recent population expansion of HPG C (Tapio et al., 2006). In fact, in a previous study, Dağtaş (2013) found a significant increase between the HPG C frequencies of Oylum Höyük ancient sheep and its possible modern extent of Ivesi (Awasi) breed. A possible explanation of the lack or low frequency of HPG C in Neolithic sheep from Anatolia is that individuals with HPG C identity were brought to Anatolia after the first migration of domestic sheep. Several suggestions for this late introduction of HPG C were proposed by Dağtaş. One of them was related to the second wave of sheep dispersal from South-West Asia at around 3000 BCE as proposed by Chessa et al. (2009). Lack of HPG C individuals before Bronze Age, together with the first observation of them at Oylum Höyük (1800 – 330 BCE) samples dating to Late Bronze Age supports the idea that HPG C individuals could have been introduced into Anatolia via the second wave of sheep dispersal predicted by Chessa et al. (2009).

Demirci et al. (2013) observed that in wild sheep of Anatolia (*Ovis gmelini anatolica*) a haplotype which was not seen among the modern domestic sheep was observed. That haplotype in terms of our screening procedure i.e. searching the mutation points on 144bp long fragment of mtDNA was observed

as HPG C. Since HPG C was not observed in our samples, one of the haplotypes of *Ovis gmelini anatolica* might have arrived later than Bronze Age.

May be there were some other haplogroups lost in the time line of domestic sheep evolution and because of the limitations of our HPG screening method we might have missed them.

4.3 Continuity tests with coalescent simulations

These tests were mainly carried out to see how well random drift accommodates the differences between the successive populations of the same locality or between the different localities**.**

Continuity analyses showed that for low mutation rate and low effective population size values, Tepecik Çiftlik population dating to 7000 BCE – 6500 BCE (denoted as TP1) is not continuous with Barcın Höyük, Ulucak Höyük, Yeşilova Höyük and Tepecik population dating to 6000 BCE – 5500 BCE (denoted as TP2). In other words, it is likely that latter populations did not stem directly from the gene pool represented by Tepecik Çiftlik 1 population.

In previous sections (section 4.2), it was explained that Yeşilova Höyük may be a settlement founded by seafaring voyagers who followed the seaside and did not used the interior land routes. In that case, it was logical to expect a differentiation between Tepecik Çiftlik and Yeşilova Höyük.

Similarly, it was mentioned that Ulucak Höyük shows a cultural continuity with several Lakes Region settlements. If sheep was brought to Ulucak Höyük from Lakes Region, then divergence of Tepecik Çiftlik and Ulucak Höyük populations was again expected.

Barcın Höyük is located at the north-western Anatolia. It is possible that this region is affected both from the seafaring voyagers and from the inland immigrants (Figure 27) (Arbuckle et al 2014). Therefore, contribution of seafaring voyagers to the sheep gene pool may have led to divergence of Barcın Höyük Sheep from Tepecik Çiftlik 1 sheep.

Discontinuity between Tepecik Çiftlik 1 (7000 BC – 6500 BC) and Tepecik Çiftlik 2 (6000 BC – 5500 BC) populations is quite interesting and suggests that Tepecik sheep population had contribution from the outside. A possible source of this contribution may be the Lakes Region. Alternatively, sheep with south or north eastern origins may have been brought to region by migrants. Another possible reason of this discontinuity may be the hunting of wild sheep during the initial stages (7000 BC – 6500 BC) of site. Although our data suggests that the central Anatolia wild sheep population was composed of HPG B animals, this does not necessarily exclude the existence of other haplogroups in the central Anatolia region.

Two levels of Barcın Höyük seem to be continuous but the sample size of Barcın Höyük 2 population is very low (n=6). This low sample size may result in a lack of power for rejecting continuity. Unfortunately, no new sheep bones were found from the upper levels of Barcın Höyük hence, for now, it is not possible to increase the sample size of Barcın Höyük 2 population.

A summary for the presentation of continuity test results was given on Table 10. Overall trend observed within the results indicates a separation between western sites (Ulucak Höyük and Yeşilova Höyük) and central and northwestern sites (Tepecik Çiftlik – Barcın Höyük). TP1 – UH, TP1 – YH and UH1 – BH1 comparisons all yielded a relatively high frequency of $p<0.05$ and $p<0.1$. On the other hand, TP1 – TP2, TP1 – BH1 and BH1 – BH2 simulations yielded low frequency of $p<0.05$ and $p<0.1$. These results suggested that, Tepecik Çiftlik and Barcın Höyük sheep populations are more related within

themselves and with each other whereas Ulucak Höyük and Yeşilova Höyük populations are less likely to be continuous with Tepecik Çiftlik and Barcın Höyük populations. Ayanoğlu (2013) who studied retrovirus integrations in modern sheep showed that western (R1 and R0 were present) Anatolia breeds were different from the central and eastern (R0 was present) Anatolia breeds. Results of the current study supported this claim further suggesting that the observed differentiation between western and central Anatolian sheep breeds can be traced back to the initial dispersion of domestic sheep into Anatolia.

Overall, results indicate that interactions between the sheep of different settlements were complex and simple hypotheses with the assumptions of direct ancestry may not explain these relations. Moreover, possibility of introgression from the wild may further complicate these relations. More elaborate simulation methods that are applied to datasets with increased number of samples may uncover the ways in which different settlements were related to each other.

4.4 Temporal diversity changes of HPG B in central and western Anatolia

Domestication bottleneck can be defined as the reduced DNA polymorphism in domesticated animals comparing to their wild counterparts (Zeder 2006). Genome wide reduction in diversity is the result of small size of the initially domesticated population which cannot retain all diversity found in the ancestral wild population. It is also possible to observe a further dramatic decrease in loci that is target of the selection owing to the quick fixation of the selected allele resulting in a selective sweep.

Diversity indices summarized in Table 8. displayed high haplotype and nucleotide diversity for samples older than 7000 BCE. Despite the fact that

samples older than 7000 BCE has much smaller sample size than the younger groups and covers only a small region within the central Anatolia, both haplotype and nucleotide diversity were higher in these oldest group ($p<0.001$). This significantly high diversity value supports the proposition that these individuals represent wild sheep or sheep at the early stage of domestication. Thus presented results might be exhibiting the case where haplotype or nucleotide diversities may serve as a marker for the domestication status of sheep.

When we remember that related to the HPG diversity increase after 7000 BCE possibly by migration and or introgression(s) from new wild sheep population(s) we expect to observe haplotype or nucleotide diversity increase in HPG B. Apparently, bottleneck effect of domestication was quite severe and HPG B in domestics was quite uniform, not allowing to observe any increase in diversity in HPG B.

4.5 Conclusion

In the present study, a general view of sheep haplogroup distribution is obtained for Central and Western Anatolia for Neolithic Period. All haplogroups except HPG C is observed in our samples, indicating that the high diversity observed in modern sheep breeds of Turkey has been maintained since the initial domestication of sheep.

Comparative analysis of genetic diversity suggested the signatures of sheep migration from east into central and western Anatolia. Moreover, exceptionally high frequency of HPG A in Yeşilova samples suggested that at least part of this migration might have followed a maritime route through the southern coasts of Anatolia as was previously suggested by archeozoological analysis (Arbuckle et al 2014)

Continuity analysis based on the employed populations revealed that relationship between ancient breeds of Anatolia was more complex than simple direct ancestry. Most interestingly, as it was revealed by the temporal analysis carried out for HPG B, haplotype and nucleotide diversity distributions over the time and space can be used as a measure of domestication status of animals in their early stages of domestication process.

REFERENCES

Abazari, F. (2017), Phylogenetic Study of Mitochondrial HPG of South Zagros Caprine, M.Sc. Thesis, Iran

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

Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., … Bunce, M. (2012). The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings. Biological Sciences*, 279(1748), 4724–33. https://doi.org/10.1098/rspb.2012.1745

Arbuckle, B.S., and L. Atici (2013). Initial diversity in sheep and goat management in Neolithic southwestern Asia. *Levant* 45(2):219-235. DOI 10.1179/0075891413Z.00000000026.

Arbuckle, B. S. (2008a). Revisiting Neolithic Caprine Exploitation at Suberde, Turkey. *Journal of Field Archaeology*, 33(2), 219–236. https://doi.org/10.1179/009346908791071277

Arbuckle, B. S. (2008b). Caprine exploitation at Erbaba Hoyük: A Pottery Neolithic village in Central Anatolia. Archaeozoology of the near East VIII : *Proceedings of the Eighth International Symposium on the Archaeozoology of Southwestern Asia and Adjacent Areas : Aswa VIII*, Lyon, June 28-July 1, 2006, 345–365.

Arbuckle, B. S., Kansa, S. W., Kansa, E., Orton, D., Çakirlar, C., Gourichon, L., … Würtenberger, D. (2014). Data sharing reveals complexity in the westward spread of domestic animals across Neolithic Turkey. *PLoS ONE*, 9(6). https://doi.org/10.1371/journal.pone.0099845

Austin, J. J., Ross, A. J., Smith, A. B., Fortey, R. A. & Thomas, R. H. 1997a problems of reproducibility—does geologically ancient DNA survive in amber-preserved insects? *Proc. R. Soc. Lond. B* 264, 467–474. (doi:10.1098/rspb.1997.0067)

Ayanoğlu, İ. C. (2013) Revealing the History of Anatolian Sheep Domestication by Using Retrovirus Integrations, M.Sc. Thesis, Biology Department, Middle East Technical University, Ankara, Turkey

Barker, G. (2002). Examining the Farming/Language Dispersal Hypothesis. P. Bellwood & C. Renfrew (Eds). Cambridge: McDonald Institute for Archaeological Research.

Bos, K. I., Schuenemann, V. J., Golding, G. B., Burbano, H. A., Waglechner, N., Coombes, B. K., … Krause, J. (2011). A draft genome of Yersinia pestis from victims of the Black Death. *Nature*, 478(7370), 506–510. https://doi.org/10.1038/nature10549

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

Bruford, M. W., Bradley, D. G., & Luikart, G. (2003). DNA markers reveal the complexity of livestock domestication. *Nat Rev Genet*, 4(11), 900-910.

Cai, D. W., Han, L., Zhang, X. L., Zhou, H., & Zhu, H. (2007). DNA analysis of archaeological sheep remains from China. *Journal of Archaeological Science*, 34(9), 1347–1355. https://doi.org/10.1016/j.jas.2006.10.020

Cai, D., Tang, Z., Yu, H., Han, L., Ren, X., et al. (2011). Early history of Chinese domestic sheep indicated by ancient DNA analysis of Bronze Age individuals. *Journal of Archaeological Science*, 38, 896-902.

Cai, D., Sun, Y., Zhuowei, T., Songmei, H., Wenying, L., Zhao, X., … Hui, Z. (2014). The origins of Chinese domestic cattle as revealed by ancient DNA analysis. *Journal of Archaeological Science*, 41, 423–434. https://doi.org/10.1016/J.JAS.2013.09.003

Chessa, B., Pereira, F., Arnaud, F., Amorim, A., Goyache, F., Mainland, I., . . . Palmarini, M. (2009). Revealing the history of sheep domestication using retrovirus integrations. *Science,* 324(5926), 532-536.

Clutton-Brock, J. (1994). Animals and human society: Changing perspectives. A. Manning & J. A. Serpell (Eds.). London: Routledge.

Cooper, A. & Poinar, H. N. 2001 Ancient DNA: do it right or not at all. *Science* 18, 289.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., … Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences of the United States of America*, 110(39), 15758–63. https://doi.org/10.1073/pnas.1314445110

Dağtaş Kılıç, N. D. (2013) A Short Fragment of Ancient DNA and Its Use in Determination of Sheep Mitochondrial DNA haplogroups in Southeast Anatolia, M.Sc. Thesis, Biology Department, Middle East Technical University, Ankara, Turkey

Derin, Z. (2005). The Neolithic Architecture of Ulucak Höyük. How Did Farming Reach Europe? Retrieved from http://www.yesilova.ege.edu.tr/arsiv/ulucak mimarisi.PDF

Demirci, S. (2012). mtdna Based Genetic Diversity of Native Sheep Breeds and Anatolian Mouflon (Ovis gmelinii anatolica) in Turkey. MSc thesis. The Graduate School Of Informatics, Middle East Technical University, Ankara, Turkey.

Dymova, M. A., Zadorozhny, A. V., Mishukova, O. V., Khrapov, E. A., Druzhkova, A. S., Trifonov, V. A., … Filipenko, M. L. (2017). Mitochondrial DNA analysis of ancient sheep from Altai. *Animal Genetics*, 48(5), 615–618. https://doi.org/10.1111/age.12569

Edwards, C. J., Machugh, D. E., Dobney, K. M., Martin, L., Russell, N., Horwitz, L. K., … Bradley, D. G. (2004). Ancient DNA analysis of 101 cattle remains: limits and prospects. *Journal of Archaeological Science*. https://doi.org/10.1016/j.jas.2003.11.001

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

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C., and M. Foll (2013) Robust demographic inference from genomic and SNP data. *PLOS Genetics*, 9(10):e1003905.

Green, R. E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., et al. (2010). A draft sequence of the Neanderthal genome. *Science*, 328, 5979, 710-722, doi: 10.1126/science.1188021.

Guilaine J. (2003). The Neolithic of Cyprus. LeBrun A. (Ed). Athens: École Français d'Athèns.

Guo, J., Du, L. X., Ma, Y. H., Guan, W. J., Li, H. B., Zhao, Q. J., … Rao, S. Q. (2005). A novel maternal lineage revealed in sheep (Ovis aries). *Animal Genetics*, 36(4), 331–336. https://doi.org/10.1111/j.1365-2052.2005.01310.x

Haile JS, Froese DG, MacPhee RDE, Roberts RG, Arnold LJ, Reyes AV, Rasmussen M, Nielsen R, Brook BW, Robinson S (2009). Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *PNAS*, 106 pp. 22352–22357

Hiendleder, S., Mainz, K., Plante, Y., & Lewalski, H. (1998). Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources: no evidence for contributions from urial and argali sheep. *J Hered,* 89(2), 113-120.

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

Higgins, D., Rohrlach, A. B., Kaidonis, J., Townsend, G., & Austin, J. J. (2015). Differential Nuclear and Mitochondrial DNA Preservation in Post-Mortem Teeth with Implications for Forensic and Ancient DNA Studies. *PLOS ONE*, 10(5), e0126935. https://doi.org/10.1371/journal.pone.0126935

Higuchi R, Bowman B, Freiberger M et al (1984) DNA-sequences from the Quagga,
an extinct member of the horse family. *Nature* 312:282–284.

Hofreiter, M., Jaenicke, V., Serre, D., von Haeseler, A., & Pääbo, S. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research*, 29(23), 4793–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11726688

Horsburgh, K.A. and Rhines, A. (2010) Genetic characterization of an archaeological sheep assemblage from South Africa’s Western Cape. *Journal of Archaeological Science*, 37, 2906-2910.

Huerta-Sánchez, E., Jin, X., Asan, Bianba, Z., Peter, B. M., Vinckenbosch, N., … Nielsen, R. (2014). Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, 512(7513), 194–197.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647-1649.

Kılınç, G. M., Omrak, A., Özer, F., Günther, T., Büyükkarakaya, A. M., Bıçakçı, E., … Götherström, A. (2016). The Demographic Development of the First Farmers in Anatolia. *Current Biology*: CB, 26(19), 2659–2666. https://doi.org/10.1016/j.cub.2016.07.057

Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto Neto, L. R., San Cristobal, M., . . . International Sheep Genomics Consortium, Members. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol*, 10(2), e1001258.

Krause, J., Fu, Q., Good, J. M., Viola, B., Shunkov, M. V., Derevianko, A. P., & Pääbo, S. (2010). The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature*, 464(7290), 894–897. https://doi.org/10.1038/nature08976

Larson, G., Dobney, K., Albarella, U., Fang, M., Matisoo-Smith, E., Robins, J., … Cooper, A. (2005). Worldwide Phylogeography of Wild Boar Reveals Multiple Centers of Pig Domestication. *Science*, 307(5715), 1618–1621. https://doi.org/10.1126/science.1106927

Larson, G., Albarella, U., Dobney, K., Rowley-conwy, P., Tagliacozzo, A., Manaseryan, N., … Cooper, A. (2007). Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proceedings of the National Academy of Sciences*, 104.

Larson, G., & Burger, J. (2013). A population genetics view of animal domestication. *Trends in Genetics*, 29(4), 197–205. https://doi.org/10.1016/j.tig.2013.01.003

Leroy, G., Mary-Huard, T., Verrier, E., Danvy, S., Charvolin, E., & Danchin-Burge, C. (2013). Methods to estimate effective population size using pedigree data: Examples in dog, sheep, cattle and horse. *Genetics Selection Evolution*, 45, 1.

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

Liu, J., Ding, X., Zeng, Y., Yue, Y., Guo, X., Guo, T., … Yuan, C. (2016). Genetic Diversity and Phylogenetic Evolution of Tibetan Sheep Based on mtDNA D-Loop Sequences. *PLOS ONE*, 11(7), e0159308.

Luo, Y.Z., Cheng, S.R., Batsuuri, L., Badamdorj, D., Olivier, H., Han, J.L., 2005. Origin and genetic diversity of Mongolian and Chinese sheep using mitochondrial DNA D-loop sequences. *Acta Genetica Sinica* 32, 1256e 1265 (in Chinese)

Lv, F.-H., Peng, W.-F., Yang, J., Zhao, Y.-X., Li, W.-R., Liu, M.-J., … Li, M.-H. (2015). Mitogenomic Meta-Analysis Identifies Two Phases of Migration in the History of Eastern Eurasian Sheep. *Molecular Biology and Evolution*, 32(10), 2515–33. https://doi.org/10.1093/molbev/msv13

Meadow RH (1989). Osteological evidence for the process of animal domestication. The Walking Larder (pp. 80-90). In Clutton-Brock J (Ed.). London: Unwin Hyman.

Meadows, J. R., Cemal, I., Karaca, O., Gootwine, E., & Kijas, J. W. (2007). Five ovine mitochondrial lineages identified from sheep breeds of the near East. *Genetics*, 175(3), 1371-1379.

Meadows, J. R., Hiendleder, S., & Kijas, J. W. (2011). Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. *Heredity* (Edinb), 106(4), 700-706.

Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., … Paabo, S. (2012). A High-Coverage Genome Sequence from an Archaic Denisovan Individual. *Science*, 338(6104), 222–226. https://doi.org/10.1126/science.1224344

Miller W, Drautz DI, Ratan A, Pusey B, Qi J, Lesk AM, Tomsho LP, Packard MD, Zhao F, Sher A, et al. (2008). Sequencing the nuclear genome of the extinct woolly mammoth. *Nature*. 456(7220):387–390.

Nei, M. (1987). Molecular evolutionary genetics. Columbia University Press.

Niemi M, Bläuer A, Iso-Touru T, Nyström V, Harjula J, et al. (2013) Mitochondrial DNA and Y-chromosomal diversity in ancient populations of domestic sheep (Ovis aries) in Finland: comparison with contemporary sheep breeds. *Genet Sel Evol* 45: 2. doi: 10.1186/1297-9686-452.

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., … Willerslev, E. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, 499(7456), 74–78. https://doi.org/10.1038/nature12323

Orlando, L., Gilbert, M. T. P., & Willerslev, E. (2015). Reconstructing ancient genomes and epigenomes. *Nature Reviews Genetics*, 16(7), 395–408. https://doi.org/10.1038/nrg3935

Ottoni, C. (2012). Article Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics. *Molecular Biology and Evolution*, 1719, 136–138. https://doi.org/10.1093/molbev/mss261

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., … Hofreiter, M. (2004). Genetic Analyses from Ancient DNA. *Annual Review of Genetics*, 38(1), 645–679.

Pedrosa, S., Uzun, M., Arranz, J. J., Gutierrez-Gil, B., San Primitivo, F., & Bayon, Y. (2005). Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. *Proc Biol Sci*, 272(1577), 2211-2217.

Pereira, F., Davis, S. J., Pereira, L., McEvoy, B., Bradley, D. G., & Amorim, A. (2006). Genetic signatures of a Mediterranean influence in Iberian Peninsula sheep husbandry. *Mol Biol Evol*, 23(7), 1420-1426.

Peter, C., Bruford, M., Perez, T., Dalamitra, S., Hewitt, G., Erhardt, G., & Consortium, Econogene. (2007). Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. *Anim Genet*, 38(1), 37-44.

Peters J, von den Driesch A, Helmer D (2005). Theupper Euphrates-Tigris Basin, cradle of agro-pastoralism? The First Steps of Animal Domestication (pp. 96-124). In Vigne J-D, Peters J, Helmer D (Eds.). Oxford: Oxbow Books.

Price, T.D. (2000). Europe’s First farmers. Cambridge: Cambridge University Press.

Rannamäe, E., Lõugas, L., Niemi, M., Kantanen, J., Maldre, L., Kadõrova, N., & Saarma, U. (2016). Maternal and paternal genetic diversity of ancient sheep in Estonia from the Late Bronze Age to the post-medieval period and comparison with other regions in Eurasia. *Animal Genetics,* 47(2), 208–218. https://doi.org/10.1111/age.12407

Rohland, N., Siedel, H., & Hofreiter, M. (2009). A rapid column-based ancient DNA extraction method for increased sample throughput. *Molecular Ecology Resources*, 10(4), 677–683. https://doi.org/10.1111/j.1755-0998.2009.02824.x

Scheu, A., Powell, A., Bollongino, R., Vigne, J.-D., Tresset, A., Çakırlar, C., … Burger, J. (2015). The genetic prehistory of domesticated cattle from their origin to the spread across Europe. *BMC Genetics*, 16, 54. https://doi.org/10.1186/s12863-015-0203-2

Sherratt, A. (1983). The secondary exploitation of animals in the Old World. *World Archaeology*, *15*(1), 90–104. https://doi.org/10.1080/00438243.1983.997988

Singh, S., Kumar Jr, S., Kolte, A. P., & Kumar, S. (2013). Extensive Variation and Sub-Structuring in Lineage A mtDNA in Indian Sheep: Genetic Evidence for Domestication of Sheep in India. *PLoS ONE*, 8(11), e77858. https://doi.org/10.1371/journal.pone.0077858

Slon, V., Hopfe, C., Weiß, C. L., Mafessoni, F., Rasilla, M. D., Lalueza-Fox, C., . . . Meyer, M. (2017, May 12). Neandertal and Denisovan DNA from Pleistocene sediments. *Science,* 608(May), 605–608.

Stiner, M. C., Buitenhuis, H., Duru, G., Kuhn, S. L., Mentzer, S. M., Munro, N. D., … Ozbaşaran, M. (2014). A forager-herder trade-off, from broad-spectrum hunting to sheep management at Aşıklı Höyük, Turkey. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23), 8404–9. https://doi.org/10.1073/pnas.1322723111

Tapio, M., Marzanov, N., Ozerov, M., Ćinkulov, M., Gonzarenko, G., Kiselyova, T., … Kantanen, J. (2006). Sheep mitochondrial DNA variation in European, Caucasian, and Central Asian areas. *Molecular Biology and Evolution*, 23(9), 1776–1783. https://doi.org/10.1093/molbev/msl043

Uerpmann H-P. (1979). Problems of the Neolithization of the Mediterranean Region. Weisbaden: Ludwig Reichert.

Vigne J-D. (2014) The origins of mammals on the Mediterranean islands as an indicator of early voyaging. *Eurasian Prehistory*.;10(1-2):45–56.

Zeder, M. A., Emshwiller, E., Smith, B. D., & Bradley, D. G. (2006). Documenting domestication: the intersection of genetics and archaeology. *Trends Genet*, 22(3), 139-155.

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

Zilhao, J. (2001). Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe. *Proc Natl Acad Sci* U S A, 98(24), 14180-14185. doi: 10.1073/pnas.241522898.

APPENDIX A

SAMPLES USED IN THE STUDY

Table A. 1. Detailed information of the samples used in the study

Excavation Site	Sample Lab. ID	Sample Excavation ID / Labels	Dates	Number of Sequence Reads	HPG
Barcın Höyük	BH 2	26447	c. 6500-6400 BCE	2	B
Barcın Höyük	BH 3	26484	c. 6300-6100 BCE	3	B
Barcın Höyük	BH 4	26484	c. 6300-6200 BCE	-	
Barcın Höyük	BH 6	31083	c. 6300-6200 BCE	2	B
Barcın Höyük	BH 7	31183	c. 6300-6200 BCE	2	B
Barcın Höyük	BH 8	26486	c. 6300-6100 BCE	1	B
Barcın Höyük	BH 9	33417	c. 6500-6400 BCE	-	
Barcın Höyük	BH 11	31547	c. 6300-6200 BCE	3	A
Barcın Höyük	BH 12	26485	c. 6300-6200 BCE	3	B
Barcın Höyük	BH 14	30790	c. 6300-6200 BCE	3	B
Barcın Höyük	BH 15	30418	c. 6500-6400 BCE	2	B
Barcın Höyük	BH 16	26474	c. 6300-6200 BCE	3	B

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Barcın Höyük	BH 17	37302	c. 6300-6200 BCE	3	B
Barcın Höyük	BH 19	37301	c. 6500-6400 BCE	2	B
Barcın Höyük	BH 20	37303	c. 6500-6300 BCE	2	B
Barcın Höyük	BH 24	31257	c. 6500-6400 BCE	2	B
Barcın Höyük	BH 26	26488	c. 6300-6100 BCE	2	B
Barcın Höyük	BH 27	31319	c. 6500-6300 BCE	3	B
Barcın Höyük	BH 28	31157	c. 6300-6200 BCE	-	
Barcın Höyük	BH 29	32070	c. 6300-6200 BCE	-	
Barcın Höyük	BH 31	37305	c. 3800 BCE	2	B
Barcın Höyük	BH 32	37084	c. 6300-6200 BCE	2	B
Barcın Höyük	BH 33	36825	c. 6300-6200 BCE	-	
Barcın Höyük	BH 35	27306	c. 6500-6400 BCE	1	B
Barcın Höyük	BH 38	36841	c. 6300-6200 BCE	-	
Barcın Höyük	BH 41	37078	c. 6300-6200 BCE	3	B
Barcın Höyük	BH 44	36440	c. 6300-6200 BCE	-	
Barcın Höyük	BH 45	37307		-	
Barcın Höyük	BH 47	36969	c. 6600-6500 BCE	-	
Barcın Höyük	BH 50	37308		2	B
Barcın Höyük	BH 51	37309		-	
Barcın Höyük	BH 54	37045	c. 6600-6500 BCE	-	
Barcın Höyük	BH 55	37045	c. 6600-6500 BCE	-	

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Barcın Höyük	BH 56	37045	c. 6600-6500 BCE	-	
Barcın Höyük	BH 57	37260	c. 6300-6200 BCE	-	
Barcın Höyük	BH 58	37172	c. 6300-6200 BCE	-	
Barcın Höyük	BH 59	37252	c. 6300-6200 BCE	-	
Barcın Höyük	BH 60	37169	c. 6300-6200 BCE	-	
Barcın Höyük	BH 61	37329	c. 6500-6400 BCE	-	
Barcın Höyük	BH 62	37327		2	B
Barcın Höyük	BH 65	37324	c. 3800 BCE	4	B
Barcın Höyük	BH 66	37325	c. 6500-6400 BCE	-	
Barcın Höyük	BH 67	37321	c. 3800 BCE	-	
Barcın Höyük	BH 68	37322		-	
Barcın Höyük	BH 69	37330	c. 6500-6400 BCE	-	
Barcın Höyük	BH 70	37328		1	B
Barcın Höyük	BH 71	37331		-	
Barcın Höyük	BH 74	37315	c. 3800 BCE	1	B
Barcın Höyük	BH 75	37314	c. 3800 BCE	3	B
Barcın Höyük	BH 76	37312	c. 3800 BCE	2	B
Barcın Höyük	BH 77	37313	c. 3800 BCE	-	
Barcın Höyük	BH 79	37317	c. 2300 BCE	2	B
Barcın Höyük	BH 80	37318	c. 2300 BCE	2	B
Barcın Höyük	BH 81	37320	c. 2300 BCE	-	
Barcın Höyük	BH 82	37319	c. 2300 BCE	3	B

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Tepecik Çiftlik	TP 1	1/15J/52C	6000-5800 BCE	3	B
Tepecik Çiftlik	TP 1_depo	780/16K	7500-7000 BCE	1	B
Tepecik Çiftlik	TP 2_depo	779/16K	7500-7000 BCE	1	B
Tepecik Çiftlik	TP 3_depo	780/16K	7500-7000 BCE	2	B
Tepecik Çiftlik	TP 4	4/15K/100C	6000-5800 BCE	1	B
Tepecik Çiftlik	TP 4_depo	779/16K	7500-7000 BCE	3	B
Tepecik Çiftlik	TP 5	5/15J/57C	6000-5800 BCE	1	B
Tepecik Çiftlik	TP 5_depo	780/16K	7500-7000 BCE	3	B
Tepecik Çiftlik	TP 7	7/15K/100C	6000-5800 BCE	3	B
Tepecik Çiftlik	TP 7_depo	783/16L	6500-6400 BCE	3	B
Tepecik Çiftlik	TP 8	8/15K/100C	6000-5800 BCE	2	B
Tepecik Çiftlik	TP 8_depo	783/16L	6500-6400 BCE	3	B
Tepecik Çiftlik	TP 10	10/17K/101C	6000-5800 BCE	2	E
Tepecik Çiftlik	TP11_depo	-/16K	6500-5800 BCE	1	A
Tepecik Çiftlik	TP12_depo	-/16K	6500-5800 BCE	2	B
Tepecik Çiftlik	TP13_depo	-/16K	6500-5800 BCE	1	B
Tepecik Çiftlik	TP14_depo	-/16K	6500-5800 BCE	2	A
Tepecik Çiftlik	TP16_depo	-/16K	6500-5800 BCE	2	E
Tepecik Çiftlik	TP 17	17/15K/111C	6000-5800 BCE	2	B
Tepecik Çiftlik	TP 20	20/15K/109C	6000-5800 BCE	2	B

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Tepecik Çiftlik	TP 41	41/15K/122C	6000-5800 BCE	3	B
Tepecik Çiftlik	TP 43	43/15J/87C	6000-5800 BCE	3	B
Tepecik Çiftlik	TP 45	45/15K/119C	6000-5800 BCE	2	B
Tepecik Çiftlik	TP 46	46/15K/119C	6000-5800 BCE	2	B
Tepecik Çiftlik	TP 48	48/15K/65C	6000-5800 BCE	1	E
Tepecik Çiftlik	TP48_2014	1345/17K/280C	6850-6650 BCE	2	B
Tepecik Çiftlik	TP 49	49/15K/65C	6000-5800 BCE	2	B
Tepecik Çiftlik	TP 50	50/15K/89C	6000-5800 BCE	2	B
Tepecik Çiftlik	TP 52	52/16K/247C	6700-6500 BCE	2	B
Tepecik Çiftlik	TP 53	53/16K/245C	6700-6500 BCE	3	B
Tepecik Çiftlik	TP 54	54/16K/253C	6700-6500 BCE	3	B
Tepecik Çiftlik	TP 55	55/16K/263C	6700 BCE	3	E
Tepecik Çiftlik	TP 56	56/16K/265C	6700 BCE	2	E
Tepecik Çiftlik	TP 57	57/16K/265C	6700 BCE	3	B
Tepecik Çiftlik	TP57_2014	1353/16K/173C	6850-6650 BCE	2	B
Tepecik Çiftlik	TP 58	58/16K/265C	6700 BCE	2	B
Tepecik Çiftlik	TP 60	60/16K/267C	6700 BCE	3	B
Tepecik Çiftlik	TP60_2014	1356/16K/250C	6850-6650 BCE	2	B
Tepecik Çiftlik	TP 61	61/16K/262C	6700 BCE	3	B
Tepecik Çiftlik	TP 62	62/16K/262C	6700 BCE	2	B

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Tepecik Çiftlik	TP68_2014	1362/16K/276C	6850 BCE	2	B
Tepecik Çiftlik	TP72_2014	1365/16K/267C	6850 BCE	3	B
Tepecik Çiftlik	TP78_2014	1370/17K/222C	6850-6650 BCE	2	A
Tepecik Çiftlik	TP82_2014	1372/17K/229C	6850-6650 BCE	2	B
Tepecik Çiftlik	TP83_2014	1373/17K/5-6 212C	6850-6650 BCE	2	A
Tepecik Çiftlik	TP87_2014	1377/17K/229C	6850-6650 BCE	2	B
Tepecik Çiftlik	TP93_2014	1383/17K/5-6 210C	6850-6650 BCE	2	B
Tepecik Çiftlik	TP94_2014	1384/17K/6- 226C	6850-6650 BCE	2	B
Ulucak Höyük	UH1	Ulucak 26, IV	6000-5700/5600 BCE	-	
Ulucak Höyük	UH2	Ulucak 21, IV	6000-5700/5600 BCE	-	
Ulucak Höyük	UH3	Ulucak 13, VI	7000-6500 BCE	-	
Ulucak Höyük	UH4	Ulucak 7, VI	7000-6500 BCE	3	B
Ulucak Höyük	UH5	Ulucak 14, VI	7000-6500 BCE	-	
Ulucak Höyük	UH6	Ulucak 16, VI	7000-6500 BCE	3	B
Ulucak Höyük	UH7	Ulucak 4, VI, DP4	7000-6500 BCE	3	A / B
Ulucak Höyük	UH8	Ulucak 15, VI, DP4	7000-6500 BCE	-	
Ulucak Höyük	UH9	Ulucak 10, VI, Astv	7000-6500 BCE	1	B
Ulucak Höyük	UH10	Ulucak 11, VI, MC	7000-6500 BCE	-	
Ulucak Höyük	UH11	Ulucak 12, VI, DP4	7000-6500 BCE	-	

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Ulucak Höyük	UH12	Ulucak 8, VI, Astv	7000-6500 BCE	3	B
Ulucak Höyük	UH13	Ulucak 9, VI, Astv	7000-6500 BCE	-	
Ulucak Höyük	UH14	Ulucak 6, VI, Ctc	7000-6500 BCE	-	
Ulucak Höyük	UH15	Ulucak 1, VI, MC	7000-6500 BCE	-	
Ulucak Höyük	UH16	Ulucak 2, VI, Astr	7000-6500 BCE	2	B
Ulucak Höyük	UH17	Ulucak 24, IV, Astr	6000-5700/5600 BCE	-	
Ulucak Höyük	UH18	Ulucak 18, IV, Astr	6000-5700/5600 BCE	-	
Ulucak Höyük	UH19	Ulucak 23, IV, MC	6000-5700/5600 BCE	-	
Ulucak Höyük	UH20	Ulucak 25, IV, MC	6000-5700/5600 BCE	-	
Ulucak Höyük	UH21	Ulucak 27, IV, MC	6000-5700/5600 BCE	-	
Ulucak Höyük	UH22	Ulucak 20, IV, MD	6000-5700/5600 BCE	-	
Ulucak Höyük	UH23	Ulucak 19, IV, HM	6000-5700/5600 BCE	1	B
Ulucak Höyük	UH24	Ulucak 22, IV, MC	6000-5700/5600 BCE	-	
Ulucak Höyük	UH25	Ulucak 17, IV, MD	6000-5700/5600 BCE		
Ulucak Höyük	UH26	Ulucak 3, VI, M3	7000-6500 BCE	2	A / B
Ulucak Höyük	UH27	Ulucak 5, VI, MD	7000-6500 BCE	2	E

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Ulucak Höyük	UH28	L12, ctd, 141, 215.11, 215.08, KOG	6500-6000 BCE	-	
Ulucak Höyük	UH29	L126, 157, 215.12, 214.96, KZR	6500-6000 BCE	-	
Ulucak Höyük	UH30	L126, 125, 215.36, 215.15, KKL	6500-6000 BCE	2	A / B
Ulucak Höyük	UH31	L12C, 139, KNU	6500-6000 BCE	3	B
Ulucak Höyük	UH32	L12D, 168, 215.13, 215.07, LHU	6500-6000 BCE	2	B
Ulucak Höyük	UH33	L126, 157, 215.12, 215.07, KYZ	6500-6000 BCE	-	
Ulucak Höyük	UH34	L12, ctd, Birim 141, 215.11, 215.08, KOG	6500-6000 BCE	3	B
Ulucak Höyük	UH35	L12, atc, Birim 146, 215.04, 211.06, KPS	6500-6000 BCE	-	
Ulucak Höyük	UH36	L12D, 163, 215.16, 215.07, LEB	6500-6000 BCE	1	B
Ulucak Höyük	UH37	L12, ctd, 94, 215.02, 214.50, LHP	6500-6000 BCE	1	B
Ulucak Höyük	UH38	L12C, 94, 215.02,	6500-6000 BCE	-	

Table A. 1. Detailed information of the samples used in the study (Cont'd)

		215.00, LDV			
Ulucak Höyük	UH39	L12C, 94, 215.02, 215.00, LDV	6500-6000 BCE	-	
Ulucak Höyük	UH40	L12C, ctd, 94, 215.02, 215.50, LHP	6500-6000 BCE	-	
Yeşilova Höyük	YH1	Kamu, 6/6C, Guneydogu kesim III, 2(3?), 16.25-16.10, 21.10.05, AME	6250-6060 BCE	3	A
Yeşilova Höyük	YH2	Kamu, 6/6C, Guneydogu kesim III, 2(3?), 16.25-16.10, 21.10.05, AME	6250-6060 BCE	1	B
Yeşilova Höyük	YH3	Kamu, 6/6C, Guneydogu kesim III, 2(3?), 16.25-16.10, 21.10.05, AME	6250-6060 BCE	-	
Yeşilova Höyük	YH4	Kamu, 6/6C, Guneydogu kesim III, 2(3?), 16.25-16.10,	6250-6060 BCE	-	

Table A. 1. Detailed information of the samples used in the study (Cont'd)

		21.10.05, AME			
Yeşilova Höyük	YH5	Kamu, 6/6C, Guneydogu kesim III, 2(3?), 16.25-16.10, 21.10.05, AME	6250-6060 BCE	2	A
Yeşilova Höyük	YH6	ATZ, III1, Ovis	6000-5800 BCE	-	
Yeşilova Höyük	YH7	ATZ, III3, Ovis/Capra	6000-5800 BCE	-	
Yeşilova Höyük	YH8	ATZ, III1, Ovis/Capra	6000-5800 BCE	1	B
Yeşilova Höyük	YH9	ATA, 45, Ovis/Capra	6000-5800 BCE	-	
Yeşilova Höyük	YH10	ATA, 45, Ovis/Capra	6000-5800 BCE	-	
Yeşilova Höyük	YH11	ATA, 45, Ovis/Capra	6000-5800 BCE	-	
Yeşilova Höyük	YH12	ATA, 45, Ovis/Capra	6000-5800 BCE	-	
Yeşilova Höyük	YH13	AZZ, III2, Ovis/Capra	6000-5800 BCE	2	B
Yeşilova Höyük	YH14	AZZ, III2, Ovis/Capra	6000-5800 BCE	-	
Yeşilova Höyük	YH15	AEC, III3	6250-6060 BCE	2	A
Yeşilova Höyük	YH16	AEC, III3	6250-6060 BCE	-	
Yeşilova Höyük	YH17	AEC, III3	6250-6060 BCE	2	B
Yeşilova Höyük	YH18	AEC, III3	6250-6060 BCE	1	B
Yeşilova Höyük	YH19	AEJ, III3	6250-6060 BCE	-	
Yeşilova Höyük	YH20	AEJ, III3	6250-6060 BCE	-	

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Yeşilova Höyük	YH21	AEJ, III3	6250-6060 BCE	-	
Yeşilova Höyük	YH22	AYA, III1, Ovis/Capra	6000-5800 BCE	-	
Yeşilova Höyük	YH23	AYA, III1, Ovis/Capra	6000-5800 BCE	-	
Yeşilova Höyük	YH24	BOG, Ovis/Capra	6250-6060 BCE	-	
Pınarbaşı Höyük	PB2	193	10th/9th M cal BCE	3	B
Pınarbaşı Höyük	PB6	188	7th M cal BCE	2	B
Pınarbaşı Höyük	PB8	10207	7th M cal BCE	2	B
Pınarbaşı Höyük	PB9	10072	7th M cal BCE	-	
Pınarbaşı Höyük	PB10	10208	7th M cal BCE	2	D
Pınarbaşı Höyük	PB12	10025	7th M cal BCE	2	B
Pınarbaşı Höyük	PB13	10015	10th/9th M cal BCE	-	
Pınarbaşı Höyük	PB14	9355	10th/9th M cal BCE	-	
Pınarbaşı Höyük	PB15	9354	10th/9th M cal BCE	-	
Pınarbaşı Höyük	PB24	8680	Epipaleo-lithic	-	
Pınarbaşı Höyük	PB25	8688	Epipaleo-lithic	2	B
Pınarbaşı Höyük	PB26	8034	Epipaleo-lithic	-	
Pınarbaşı Höyük	PB28	8194	Epipaleo-lithic	2	B
Pınarbaşı Höyük	PB29	10397	10th/9th M cal BCE	-	

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Pınarbaşı Höyük	PB31	PB5071,8264	Epipaleo-lithic	-	
Pınarbaşı Höyük	PB32	#4094	Epipaleo-lithic	-	
Pınarbaşı Höyük	PB33	#4094	Epipaleo-lithic	2	B
Boncuklu Höyük	BK1	9318	8400-7800 cal BCE	1	B
Boncuklu Höyük	BK2	2540	8400-7800 cal BCE	-	
Boncuklu Höyük	BK3	2945, TO CHECK	8400-7800 cal BCE	2	B
Boncuklu Höyük	BK4	4386	8400-7800 cal BCE	-	
Boncuklu Höyük	BK5	4370	8400-7800 cal BCE	-	
Can Hasan III	CH1	BONE ID 142	7400-7100 cal BCE	2	B
Can Hasan III	CH2	BONE ID 137	7400-7100 cal BCE	2	B
Can Hasan III	CH3	BONE ID 087	7400-7100 cal BCE	2	B
Can Hasan III	CH4	BONE ID 150	7400-7100 cal BCE	1	A
Can Hasan III	CH5	BONE ID 014	7400-7100 cal BCE	1	B
Çatalhöyük	CT1	1023.F201	6700-6500 BCE	1	B
Çatalhöyük	CT4	1889.F183	7300-6800 BCE	1	B
Çatalhöyük	CT5	1889.F184	7300-6800 BCE	-	
Çatalhöyük	CT8	4824.F15	7300-6800 BCE	-	
Çatalhöyük	CT10	5290.F2608	7300-6800 BCE	-	
Çatalhöyük	CT15	5328.F139	7300-6800 BCE	1	B
Çatalhöyük	CT19	11370.F42	6400-6000 BCE	1	B

Table A. 1. Detailed information of the samples used in the study (Cont'd)

Çatalhöyük	CT20	11370.F44	6400-6000 BCE	2	B
Çatalhöyük	CT26	16262.F64	6400-6000 BCE	-	
Çatalhöyük	CT27	16262.F102	6400-6000 BCE	-	
Çatalhöyük	CT28	17047.F480	6400-6000 BCE	-	
Aktopraklık Höyük	AKT 1	-	MÖ 6320	-	
Aktopraklık Höyük	AKT 2	-	MÖ 6320	-	
Aktopraklık Höyük	AKT 3	-	MÖ 6320	-	
Aktopraklık Höyük	AKT 4	-	MÖ 6320	-	
Aktopraklık Höyük	AKT 5	-	MÖ 6320	1	B
Aktopraklık Höyük	AKT 6	-	MÖ 6320	-	
Aktopraklık Höyük	AKT 7	-	MÖ 6320	1	B
Aktopraklık Höyük	AKT 8	-	MÖ 6320	-	

APPENDIX B

SIMULATION RESULTS WITH GENERATION TIME OF 4 YEARS

Table B.1. Frequency of simulated populations with indicated significance values over all mutation rate and population size values. Generation time was assumed as 4 years.

Pairwise Comparison	No growth			Bottleneck and exponential growth		
	$p<0.05$	$p<0.1$	$p>0.2$	$p<0.05$	$p<0.1$	$p>0.2$
TP1 - TP2	0.17	0.30	0.51	0.17	0.32	0.53
TP1 - BH1	0.17	0.33	0.46	0.17	0.3	0.49
BH1 - BH2	0	0	1	0	0	1
TP1 - UH	0.37	0.63	0	0.32	0.56	0.02
TP1 - YH	0.27	0.44	0.06	0.25	0.43	0.13
UH1 - BH1	0.32	0.6	0	0.31	0.52	0
UH – YH	0.29	0.49	0.13	0.29	0.48	0.15