ANCIENT DNA ISOLATION AND MITOCHONDRIAL DNA ANALYSIS OF HUMAN SAMPLES FROM ÇEMİALO SIRTI, BATMAN IN SOUTHEAST ANATOLIA

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

BY
REYHAN YAKA

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

JUNE 2015

Approval of the thesis:

ANCIENT DNA ISOLATION AND MITOCHONDRIAL DNA ANALYSIS OF HUMAN SAMPLES FROM ÇEMİALO SIRTI, BATMAN IN SOUTHEAST ANATOLIA

submitted by **REYHAN YAKA** in partial fulfillment of the requirements for the degree of **Master of Science in Biology 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**

Prof. Dr. İnci Togan _____
Supervisor, **Biology Dept., METU**

Examining Committee Members:

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

Prof. Dr. İnci Togan _____
Biology Dept., METU

Assoc. Prof. Dr. Ercan Arıcan _____
Molecular Biology and Genetics Dept., Istanbul University

Assoc. Prof. Dr. Aslı Erim Özdoğan _____
Archaeology Dept., Çanakkale University

Assist. Prof. Dr. Yasemin Yılmaz _____
Archaeology Dept., Düzce University

Date: 23.06.2015

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: Reyhan Yaka

Signature :

ABSTRACT

ANCIENT DNA ISOLATION AND MITOCHONDRIAL DNA ANALYSIS OF HUMAN SAMPLES FROM ÇEMİALO SIRTI, BATMAN IN SOUTHEAST ANATOLIA

Yaka, Reyhan

M.Sc., Department of Biology

Supervisor: Prof. Dr. İnci Togan

June 2015, 182 pages

The main purpose of the study was to obtain aDNA sequences of ancient human remains in the dedicated ancient DNA (aDNA) laboratory which was established at Middle East Technical University, in 2012. For this purpose, human samples approximately dating between 600-500 BC from Çemialo Sırtı excavation site in Batman in southeast Anatolia, were employed. aDNA extraction was performed using bone and teeth samples from 9 human skeletal remains. Then the mitochondrial DNA (mtDNA) Hypervariable region I and Hypervariable region II (HVRI-HVRII) of 7 samples could be successfully amplified and their sequences were obtained. Success rates in extractions (9/9) and amplifications (7/9) were 100% and 77.8%, respectively. Postmortem nucleotide changes (misincorporations) were also detected supporting the authenticity of the sequences. aDNA sequences were used to determine mtDNA haplogroups (HPGs) of the samples. The observed percentage of each haplogroup was: HPG H=28.57%, HPG HV=14.29%, HPG M=14.29%, HPG R=28.57%, HPG U2=14.29% for the 7 Çemialo Sırtı individuals.

When the mtDNA based haplogroup and sequence data of the present study were evaluated together with the results of some other ancient and modern populations by means of Principle Component Analysis (PCA), it was observed that Çemialo Sırtı population is relatively similar to the neareastern Neolithic population (7300 BC) from northern Syria. Continuity between the Çemialo Sırtı population (n=7) and the same Neolithic population from northern Syria (n=10) was tested by calculating the F_{ST} value between these two populations, and comparing that with the F_{ST} values generated by coalescent simulations. During the simulations, random drift was assumed to be the force of genetic differentiation between the two populations. Under the exponential growth and no growth models, continuity from Neolithic population to Çemialo Sırtı population could not be rejected. Since the sample size was small in Çemialo Sırtı population, comparative results must be taken with caution.

Keywords: Ancient DNA, human, mtDNA HVRI-HVRII, Çemialo Sırtı, Anatolia

ÖZ

GÜNEYDOĞU ANADOLU'DAKİ BATMAN, ÇEMİALO SIRTI KAZI YERİNDEN ÇIKARILMIŞ İNSAN ÖRNEKLERİNDEN ANTİK DNA İZOLASYONU VE MİTOKONDRİYAL DNA ANALİZİ

Yaka, Reyhan
Yüksek Lisans, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. İnci Togan

Haziran 2015, 182 sayfa

Sunulan çalışmanın ana amacı eski dönemlerde yaşamış insanların kalıntılarından antik DNA (aDNA) dizi eldesinin Orta Doğu Teknik Üniversitesi'nde 2012 yılında kurulmuş olan adanmış antik DNA laboratuvarında gerçekleştirilmesidir. Bu amaçla, Güneydoğu Anadolu'da bulunan, Batman ilindeki Çemialo Sırtı kazı yerinden yaklaşık olarak M.Ö. 600-500 yılları arasına tarihlendirilen insana ait örnekler çalışılmıştır. Antik DNA izolasyonu, 9 insan iskeletinden alınan kemik ve diş örnekleri kullanılarak yapılmıştır. Daha sonra, 7 birey için mitokondriyal DNA'nın (mtDNA) Yüksek Çeşitlilik Gösteren I. ve II. bölgesi (HVRI-HVRII) başarılı bir şekilde yükseltgenebilmiş ve dizileri elde edilmiştir. Bu örnekler için DNA izolasyonu (9/9) ve yükseltgenmesindeki başarı oranı (7/9) sırasıyla %100 ve %77.8'dir. Ayrıca, canlı öldükten sonra gerçekleşen nükleotid değişimlerinin (hatalı baz yerleştirme) gözlenmiş olması dizilerin otantikliğini desteklemektedir. Elde edilen aDNA dizileri bireylerin mitokondriyal haplogruplarının (HPG) belirlenmesinde kullanılmıştır. Yedi Çemialo Sırtı bireyi için her bir haplogrubun

görülme yüzdesi: HPG H=%28.57, HPG HV=%14.29, HPG M=%14.29, HPG R=%28.57, HPG U2=%14.29 şeklindedir.

Sunulan çalışmanın mtDNA bazındaki haplogrup ve dizi verileri, bazı antik ve günümüz populasyonlarının sonuçlarıyla birlikte Temel Bileşenler Analizi ile değerlendirildiğinde; Çemialo Sırtı populasyonunun Kuzey Suriye'deki Yakın Doğu Neolitik populasyonuna göreceli olarak benzer olduğu gözlemlenmiştir. Çemialo Sırtı populasyonu (n=7) ve Kuzey Suriye'deki Neolitik populasyon (n=10) arasındaki süreklilik, iki populasyon arasındaki gözlenmiş F_{ST} değeri ile simulayonlarla oluşturulan iki populasyon arasındaki F_{ST} değerlerinin karşılaştırılmalarıyla test edilmiştir. Simulasyonlarda iki populasyonun genetik farklılaşmasının rastlantısal sürüklenme ile oluştuğu varsayılmıştır. Bir populasyondan diğerine üssel büyüme ve büyüme olmayan olmak üzere iki model altında, Neolitik populasyondan Çemialo Sırtı populasyonuna olan süreklilik reddedilememiştir. Çemialo Sırtı populasyonuna ait birey sayısının az olmasından dolayı karşılaştırmalı sonuçlara dikkatle yaklaşılması gerekmektedir.

Anahtar Kelimeler: Antik DNA, insan, mtDNA HVRI-HVRII, Çemialo Sırtı, Anadolu

To My Family,

ACKNOWLEDGEMENTS

First I would deeply like to express appreciation to my supervisor Professor İnci Togan not only for her scientific suggestions and academic guidance but also for her support, attention, advises and encouragements for the completion of this thesis. She was there for me all the time whenever I needed scientific suggestions and supports. I am extremely thankful to her for teaching the academic life style as well as the life itself and that learning made me feel stronger. It is a great honor to be a member of her team.

I am deeply grateful to Dr. Claudio Ottoni for teaching me the ancient DNA isolation procedure and analysis that I used in my study. Special thanks for his supports, helpfulness and scientific suggestions not only during my internship in his laboratory but also after it when I came back.

I would like to many thanks to Assoc. Prof. Mehmet Somel who made me to be involved in his project and researches. Also most importantly he contributed to this study with his valuable comments and encouragements all the time whenever I needed suggestions and supports. I am indebted to him for teaching me the scientific life and being a scientist.

I am extremely thankful to Dr. Ceren Caner Berkman not only for her important contributions, advises and encouragements but also for her friendly attitude and always being nice and bighearted. She never denied her scientific and moral supports to me since the very beginning of this study. I would express my deep debt of gratitude.

Special thanks go to Dr. Can Acan, Eren Yüncü and Nihan Dilşad Dağtaş Kılıç for their infinite patience, scientific supports, friendships, advises and encouragements whenever I was in trouble and dealing with my analysis. I am extremely thankful to all of them also for their moral support whenever I needed. I would like to warmly thank Dilan Saatoğlu, once more Eren Yüncü and Nihan Dilşad Dağtaş Kılıç for their friendships, joyful moments and for being that bighearted whenever I felt dispirited, almost going to give up. My present and former labmates and friends: Dr. Füsun Gezgin Özer, Eren Yüncü, Dr. Can Acan, Nihan Dilşad Dağtaş Kılıç, Dilan Saatoğlu and Onur Özer never denied their technical and companionship to me since the very beginning of this study. I express them many thanks.

I am also thankful to Assoc. Prof. Aslı Erim Özdoğan, Assist. Prof. Yasemin Yılmaz and Sidar Gündüzalp who have contributed to this study by sharing and evaluating the valuable archaeological materials. Another special thanks go to Sidar Gündüzalp who was a phone call away whenever I needed information about the archaeological site and materials. I would also like to thank Sidar Gündüzalp, Aliye Gündüzalp and Esra Çiftçi for their attention and contributions to the sample collection for this study.

I would also like to special thank my lovely friend, Pınar Şimşek who was always with me with her beautiful heart, friendship and never-ending support. I offer warmly thanks to Shahrzad Nikghadam and Deniz Çakal for always being that nice and bighearted, their friendships, moral supports and joyful moments. They were always with me whenever I needed scientific and technical helps. I truly appreciate all of those mentioned friends of mine. I would deeply like to express my appreciation and thanks to my little lovely friend and puppy Kaymak for teaching me the most important thing in life. I must express my special appreciation and thanks to bighearted people; Can Acar and Mehmet Kural for lending their computers during my analysis; Melike Dönertaş and Poorya Parvizi for their endless helps about R program analysis. I would like to thank my friends Cemile Kılıç Bektaş, Seray Akça, Onur Bulut, Hande Öztürk that they never denied whenever I asked for help.

Finally, but most entirely, I would like to express my deepest gratitude and appreciation to my family, to my mother Zehra Yaka, my father Kazım Yaka and my lovely sister Seher Yaka, for their never-ending love, technical and moral supports, infinite patience, care and generosity. They were always there with me whenever I needed helps about everything and felt hopeless, almost going to give up. I am extremely thankful to my family for providing me everything to be successful and for their many sacrifices all the time. It is not possible to thank them enough, I am fabulously glad and happy that they are my family and always with me!

This study was supported by the Scientific and Technical Research Council of Turkey (TÜBİTAK) as a part of the project "Bir Ön Çalışma Olarak Çemialo Sırtı Yerleşim Yerinden Elde Edilmiş 2500-3000 Yıllık İnsan Diş Örneklerinden Otantik Antik DNA Eldesi ve mtDNA Haplogrup Tayini" under the grant number 114Z159 and by Middle East Technical University as a part of the project "Oylum Höyük arkaeolojik kazı yerinden steril koşullarda çıkartılmış insan diş ve kemiklerinden MtDNA ekstraksiyonu ve buna bağlı haplogrup tayini" under the grant number BAP-08-11-2013-027.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

AD: After Date

aDNA: Ancient DNA

BCE: Before Common Era

BC: Before the Common

BLAST: Basic Local Alignment Search Tool

bp: base pair

np: nucleotide position

BP: Before Present

CRS: Cambridge Reference Sequences

DNA: Deoxyribonucleic acid

dNTP: Deoxyribonucleotide triphosphates

UV: Ultraviolet

RNAse: Ribonuclease

dH2O: Distilled water

nm: nanometer

HYA: Synthetic Hydroxyapatite powder)

mg: miligram

EDTA: Ethylenediaminetetraacetic acid

M: Molar

pH: Power of Hydrogen

EB: Elution Buffer

UDG: Uracil-DNA-Glycosylase

HPG: Haplogroup

HVRI: Hypervariable region I

HVRII: Hypervariable region II

PC1: First principle component

PC2: Second principle component

PC: Principle component

PCA: Principal Component Analysis

Ne: Effective population size

F_{ST}: Fixation index

min: minute

ml: milliliter

mm: millimeter

mtDNA: mitochondrial DNA

Myr: Million years

PCR: Polymerase Chain Reaction

rpm: revolutions per minute

s: second

µl: microliter

cal: calibration

CHAPTER 1

INTRODUCTION

In the present study, a newly established dedicated ancient DNA (aDNA) laboratory was used to isolate ancient DNA from humans. Furthermore, two segments of mitochondrial DNA (mtDNA) of ancient human samples were amplified and sequenced. The bones and teeth samples of 9 individuals which were unearthed from the Çemialo Sırtı excavation site near Batman, in the southeast Anatolia, were used for ancient DNA extraction. From their mtDNA, Hypervariable region I and Hypervariable region II (HVRI, HVRII) sequences were amplified and sequenced by the Sanger sequencing method. The mtDNA HVRI-HVRII of 7 individuals were successfully sequenced and evaluated comparatively together with some modern and ancient human populations.

This study is the first study where the questions about the human aDNA are addressed in the dedicated aDNA laboratory of Middle East Technical University (METU), Ankara, Turkey.

1.1 Ancient DNA Studies and Related Problems in Brief

aDNA is a type of DNA extracted from archaeological findings, museum specimens and fossil remains such as teeth, bones, hair, tissue, seeds and other ancient sources that may contain DNA (Pääbo et al., 2004). The first ancient DNA was successfully obtained from a piece of dried muscle of the Quagga (a subspecies of zebra); aDNA was extracted and a short fragment of its mtDNA, 229 base pairs in length, was cloned from one individual for almost 30 years ago in the study of Higuchi et al. (1984). Afterwards, aDNA study techniques have been further developed for the next

30 years and they were employed to answer various questions in relation to extinct species and populations (Gilbert et al., 2005; Pääbo, 2014). For instance, aDNA analysis is used to retrieve information about the species' phylogenies and history of the populations, phylogeography of the species, diet and behaviour of, for example, humans, evolution of the microbes, origins of domestication (Pääbo et al., 2004). In the last decade, with the development of high-throughput and next generation sequencing technologies, including the shotgun approach, many questions about the genomes of ancient species and their populations, could be answered (Pääbo et al., 2004). However, the recent advanced technologies are expensive for the time being, therefore, in the present study parts of HVI and HVII regions (HVRI-HVRII) of human mitochondrial DNA (mtDNA) were sequenced by the Sanger sequencing method. mtDNA is one of the most widely used molecular marker in the studies of humans and further information about this widely used genetic marker is going to be given in below in section 1.2 (Mitochondrial DNA as a Molecular Marker).

In the first aDNA study on humans 2400-year-old Egyptian mummy which belongs to a child was examined. The 3.4 kilo bases (kb) region of the mummy DNA was sequenced using the plasmid vector cloning method (Pääbo, 1985). However, later it was revealed that this aDNA sequence was a contamination from a modern human DNA (Del Pozzo et al., 1989; Knapp et al., 2015). A well learned lesson, after this observation, was that the aDNA are prone to the contamination from modern human DNA. Thus, since that time, researchers studying aDNA took extreme precautions to avoid the contaminations from the modern human DNA (Pääbo et al., 2004; Gilbert et al., 2005; Knapp et al., 2015). In relation to aDNA studies it was also understood that aDNA is highly fragmented (Pääbo et al., 2004; Sawyer et al., 2012; Knapp et al., 2015) and size of the fragments decrease in time (Knapp et al., 2015). Furthermore, after the death, due to the modifications of some of its bases in some copies of the aDNA incorrect bases are incorporated during the amplification step of aDNA. Thus, amplified aDNA harbours non-authentic (misincorporated) bases in aDNA sequences (Hansen et al., 2001; Hofreiter et al., 2001; Pääbo et al., 2004;

Olivieri et al., 2010). For each of these problems of aDNA studies, guidelines or procedures are developed to avoid or to control them.

However, contamination could not be excluded completely from the aDNA studies (Gilbert et al., 2005). The guidelines for the aDNA studies, in order to avoid the exogenous contamination, and to obtain the authentic sequences are listed below (Pääbo et al., 2004; Gilbert et al., 2005):

1. In the aDNA laboratory the sample preparation, aDNA extraction and PCR preparation steps should be in three different cabinets or in the three separated rooms of the aDNA laboratory.
2. There should be negative controls in the each step for the grinding of ancient samples, aDNA extraction and PCR amplification to check the possible contaminations at each step during the experiments.
3. Repeats of the extraction and PCR amplification should be carried out multiple times and the successfully amplified sequences should be compatible with each other for the reproducibility.
4. Especially for the ancient human studies the genetic profile of the research team members should be compared with the genetic profiles of the ancient samples. Ancient profiles must be different from each of the modern profiles.
5. The pre-PCR and PCR amplification steps should be carried out in a dedicated aDNA laboratory which must have the positive pressure facility.
6. In the aDNA laboratory all the benches and plastic tools should be cleaned using the 5% sodium hypochlorite solution and should be irradiated with UV (254 nm) light regularly.
7. The reagents which are used for all of the procedures in the aDNA laboratory should be nuclease-free, DNA-free and if possible, UV (254 nm) irradiated at least for 20 minutes.
8. After putting on full body cover, gloves, face masks, over boots, bonnets; should be entered to the aDNA laboratory and gloves should be changed often during the experiments.

9. To check the reproducibility and accuracy of results, the experiments should be repeated using a second sample of the same individual when the second sample of the same individual is available.
10. The multiple amplifications should be carried out to detect the nucleotide misincorporations which occur in some copies of aDNA molecules. The details of the nucleotide misincorporations are given in below in section 1.4 (Molecular Damage, Nucleotide Misincorporations and Contaminations of the aDNA).
11. The aDNA molecules are fragmented when compared to modern DNA. Thus, the aDNA sequences should be obtained on the basis of small fragments.

1.2 Mitochondrial DNA as a Molecular Marker

Human mitochondria exist in high numbers in cells. Mitochondria have double stranded, circular DNAs called as mitochondrial DNA (mtDNA). mtDNA have the same sequence in all mitochondria of all cells in an individual. Therefore, in an individual, identical mtDNA copies exist in many cells (such as bones, teeth, hair and skin). Furthermore, mtDNA has relatively high mutation rate, compared to other parts of the genome. Especially, the two regions of mtDNA known as Hypervariable I (mutation rate is nearly 3.6 $x10^{-6}$ per base pair (Richards et al., 2000)) and hypervariable II have highest mutation rate in the genome. This property of mtDNA provides high resolution in differentiating individuals. Furthermore, because mtDNA is a haploid molecule, it is more affected by drift and can be fixed in small populations. Thus, mtDNA can be used to discriminate the populations. mtDNA is maternally inherited. Since, it does not exhibit recombination, in the statistical analysis it is easy to handle. Because of all these properties, since the very early days of DNA studies, mtDNA is commonly used to trace the maternal lineages, for instance in humans (Jobling et al., 2004). The preference of mtDNA in the DNA studies in turn resulted the accumulation of ample amount of data (in particular HVRI based data) to be employed for comparative studies. However, it is just one gene in the whole genome; therefore results based on mtDNA should be considered with caution. mtDNA sequences are aligned and evaluated in accordance of a

standard reference sequence: Cambridge Reference Sequence (CRS, Anderson et al., 1981). CRS is a human mtDNA sequence which was published by Anderson et al. (1981) and used as a reference sequence for the comparison of mtDNA sequences. The CRS belongs to European haplogroup H2a2a which is included in macrohaplogroup H (Anderson et al., 1981).

For the mtDNA as well as the sequences of individuals mtDNA haplogroups (groups formed on the basis of similar sequences) of the individuals are, also, commonly used and reported. More specifically: A haplogroup is a group of similar haplotypes (mtDNA sequence variants) which share the same variation as mutation motifs. Haplogroups and the mutation motifs used to identify them are given in Materials and Methods Chapter of the thesis in section 2.2.7.2 (Haplogroup Determination). These groups were composed of individuals who shared a common ancestor at some times in the past based on mtDNA. The major mtDNA haplogroups, represented as the major branches on the mtDNA phylogenetic tree was given in Figure 1.1. Each major mtDNA haplogroup was indicated with the capital letters from A to Z according to their order of observations (van Oven et al., 2009).

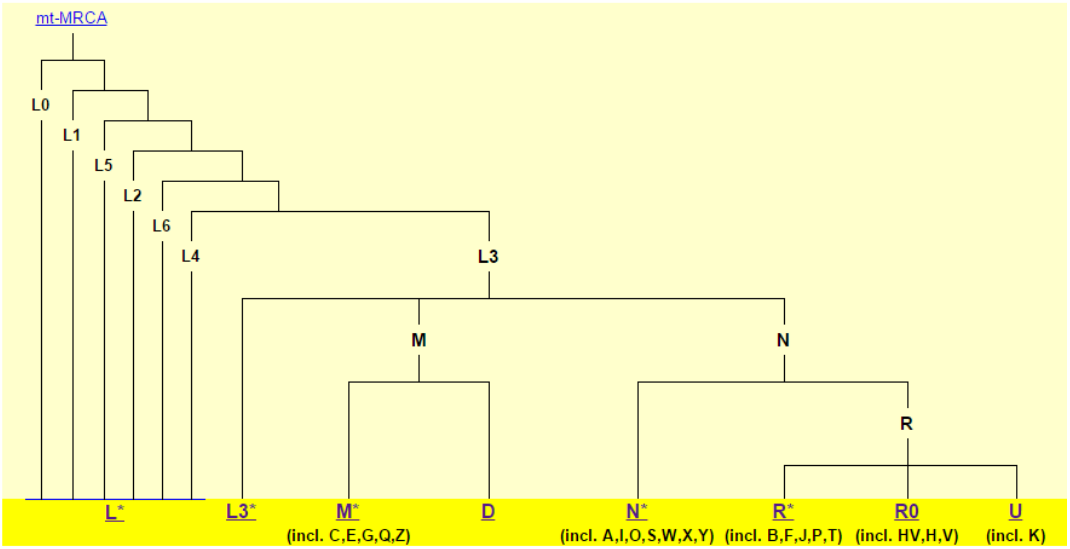


Figure 1.1 A screenshot from http://www.phylotree.org/ as of February, 2014. It shows the phylogenetic tree with respect to human mtDNA major haplogroups." İncl": including.

As of 2014, approximately 35 major mtDNA haplogroups were classified on the mtDNA phylotree (http://www.phylotree.org/). Some specific mtDNA major haplogroups could be frequently observed in particular geographical regions and /or in populations. For instance, haplogroups H, HV, I, J, K, M, N, R, T, U and Z are found mostly in the West Eurasian populations, and thus accepted as the typical West Eurasian haplogroups. Whereas, A, B, C, D and F haplogroups are frequently observed in the Central and East Asian populations (Cherni et al., 2009; van Oven et al., 2009; Schönberg et al., 2011). The haplogroup X is accepted as Caucasian haplogroup and it is also found frequently in the Native Americans (Schönberg et al., 2011; Sykes 2001; Malhi and Smith, 2002). The ancestor of all major haplogroups is accepted as haplogroup L which splits into six sub-groups (Figure 1.1). The other mtDNA major haplogroups branched from L3 haplogroup (Figure 1.1). L3 is known as the origin of all human mtDNA haplogroups, those emerged at the outside of Africa (Metspalu et al., 2004).

There are, already, some studies covering mtDNA haplogroups of ancient human populations. It is emerging that for example, the LBK (Linear Pottery culture, 5500-

4775 cal BC) is characterized with a special group of haplogroups N1a, T2, K, J, HV, V, W, X, thus, collectively they are called as a mitochondrial "Neolithic package" (Brandt et al., 2013). The haplogroup H was common in Neolithic and modern populations but since it is also observed in pre-agricultural groups in Iberia, it was not in the Neolithic package (Hervella et al., 2012; Szécsényi-Nagy et al., 2015). Furthermore the European hunter-gatherers were defined by mostly U lineages such as U, U4, U5, and U8 with based on mtDNA data (Bramanti et al., 2009; Fu et al., 2013). It is also known that haplogroup U lineages (U2, U4, U5) was especially found in Palaeolithic, Mesolithic and Bronze Age samples from Russia (Keyser et al., 2009; Krause et al., 2010; Der Sarkissian et al., 2013; Brandt et al., 2013). Brandt et al.'s (2013) study suggested that the most frequent haplogroups of Eurasian populations are N1a, I, I1, W, X, HV, HV0/V, H, H5, T1, T2, J, U, U2, U3, U4, U5a, U5b,U8, K; while the haplogroup L is African and haplogroups A, B, C, D, E, F, G, Q, Y, and Z are defined as Asian haplogroups.

1.3 aDNA Studies Based on Human mtDNA HVI and HVII Regions

In this section, some ancient human mtDNA studies which are related to population genetics studies including the kinship analysis from Near Eastern and European regions are presented.

In a study by Ottoni et al. (2011), 85 samples from the Sagalassos (mid-Byzantine period: 11th–13th century AD) excavation site in southwest Anatolia were analysed on the basis of sequences of mtDNA HVI-HVII regions and aDNA sequences were obtained from 53 (53/85= 62.4%) individuals. mtDNA haplogroups were also determined and compared with those of present-day populations' from Asia, Balkans, East Mediterranean, Near/Middle East, North Africa, North Europe and South Europe. It was suggested that, individuals of the mid-Byzantine period (11th–13th century AD) population from southwest Anatolia are closely related to Balkan (Northern Greece, Macedonia, Bosnia and Bulgaria), Italian and Iran populations based on mtDNA HVI-HVII regions.

In another study from France (Lacan et al., 2011), DNA was extracted from 29 out of 53 samples (29/53= 54.7%), their mtDNA as well as Y chromosome sequences were amplified. The samples were approximately 5,000 year-old and excavated from the cave of Treilles in France. In the study between the samples from the same grave, the close kinship was observed and additionally Y-chromosome comparisons indicated that South France might have more Neolithic migration than the Middle Europe.

In another study, Matney et al. (2012) examined 26 samples from 13 individuals excavated from Titriş Höyük in southeast Anatolia. mtDNA HVI region of 13 individuals which are dated to Early Bronze Age were analysed and compared with modern populations'. It was observed that there is a little variation among the mtDNA sequences of 12 individuals, 10 of them had CRS. The result was interpreted as either the population went through a bottleneck or those individuals were from closely related families.

Malmström et al.'s (2012) study presented a kinship analysis of three individuals excavated from Varnhem grave in Vastergotland, Sweden. The three samples, dated to 13th century AD, were analysed based on mtDNA, Y-chromosome and nuclear SNP analyses. Then, the study revealed that all of the three individuals have different mtDNA haplogroups, while the Y-chromosome analysis showed a kinship between the two males.

In the Fernandez et al.'s (2014) study mtDNA profiles of 15 out of 63 individuals from the Pre-Pottery Neolithic B (PPNB) sites of Tell Halula, Tell Ramad and Dja'de El Mughara which were dated back to 8,700–6,600 BC were analysed. This mtDNA data was, for the first time, presenting the original near eastern Neolithic groups. In this study, mtDNA haplogroups of 15 samples could be determined and compared with those of the modern populations retrieved from the databases. Moreover, on the basis of mtDNA haplogroups, and comparative analysis between modern populations, haplogroup compositions at the time of Neolithic spread was inferred. As a result, it was suggested that K and N derived mitochondrial DNA

haplogroups can be the typical markers of the Neolithic expansion and this genetic signature could have reached both to the Iberian coasts and the Central European plains. It was also suggested that due to the genetic similarity between the PPNB samples and the modern populations of Cyprus and Crete, the Neolithic culture might be introduced into Europe for the first time through the seafaring colonization (Fernandez et al., 2014).

In the Szécsényi-Nagy et al.'s (2015) study, in total 84 for the mtDNA and 9 samples for the Y chromosomal haplogroups were determined. Samples were from southeast Europe and they represented STA (Early Neolithic Starčevo culture, 6000–5400 BC) and LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC) cultures. The STA culture is believed to play a role for the Neolithization of southeast Europe by expanding from modern day Serbia to the western Carpathian Basin, covering the present-day northern Croatia and southwest Hungary regions. The LBKT is an earliest Linearbandkeramik culture (LBK) emerged in western Hungary, around Transdanubia, and caused the beginning of farming life in Central Europe. It is known that these two cultures coexisted together for approximately 100–150 years in that region.

In the Szécsényi-Nagy et al.'s (2015) study, mtDNA HVI region of 84 samples and the HVII region of 25 samples were sequenced. mtDNA haplogroup composition and haplotype diversity of the 84 samples indicated that those early farming populations from the Carpathian Basin and the populations of the Central European LBK shared a maternal genetic ancestry. Szécsényi-Nagy et al. (2015) also found that there is an affinity between the early farmers and modern populations from the Near East and Caucasus. On the other hand, the differences were detected between the haplogroup variations of Y- chromosomal data and mtDNA based data, most likely, due to the patrilocality (when women move to their husband’s birth place after they get married) in the early farming populations.

In another study, Brandt et al. (2013), determined the mtDNA haplogroups of 364 individuals from different Neolithic cultures in Central Europe, such as LBK (Linear Pottery culture, 5500-4775 cal BC), RSC (Rössen culture, 4625-4250 cal BC), SCG (Schöningen group, 4100-3950 cal BC), BAC (Baalberge culture, 3950-3400 cal BC), SMC (Salzmünde culture, 3400-3100/3025 cal BC), BEC (Bernburg culture, 3100-2650 cal BC), CWC (Corded Ware culture, 2800-2200/2050 cal BC), BBC (Bell Beaker culture, 2500-2200/2050 cal BC) and UC (Unetice culture, 2200-1550 cal BC). The dates for these populations changes between the Early Neolithic and Early Bronze Age. These 364 samples were collected from 25 sites in central European regions and their mtDNA HVI-HVII regions were sequenced. Then the mtDNA haplogroups of samples were determined. These haplogroups were used in comparative analysis to detect the genetic similarities between those samples and the samples from Mesolithic, Neolithic, and Bronze Age populations from western Eurasia (n=198). Likewise, similarities were searched to 67,996 sequences of the modern Eurasian populations from databases. In the Brandt et al.'s (2013) study it was suggested that Late Neolithic cultures had a contribution to the present-day Central European gene pool based on the mtDNA haplogroup compositions.

In the study of Rivollat et al. (2015) 102 samples from the Gurgy 'Les Noisats' population (5000-4000 cal. BC) an Early/Middle Neolithic region in the southern of the Paris Basin were analysed based on their mtDNA profiles. For the 39 samples mtDNA HVI region sequences and for the 55 samples mtDNA haplogroups could be obtained. In this study, for the region under consideration they tried to understand if the spread of Neolithic culture was from the Danube valley or from the Mediterranean Sea. Also the haplogroup frequencies and shared haplotypes data of Gurgy samples were analyzed and compared with ancient and modern European, Near Eastern and other ancient populations which were retrieved from the databases. As a result, Rivollat et al. (2015) suggested that Gurgy populations had the equal genetic contributions from the farmers arriving both from the Danube and Mediterranean Sea.

1.4 Molecular Damage, Nucleotide Misincorporations and Contaminations of the aDNA

In the living cells the DNA damages are repaired by enzymatic pathways (Lindahl, 1993). After death this repair mechanisms do not work (Pääbo et al., 2004). Molecular damages occur in the DNA molecules of fossil and archaeological remains. The most common DNA damages were given in Table 1.1. DNA might be degraded by some lysosomal enzymes (Eglinton et al., 1991). The fragmentation of aDNA molecules into the short length fragments between 100 to 500 bp in length (Hofreiter et al., 2001; Pääbo, 1989) is seen. Thus, aDNA sequences are relatively shorter than modern DNA (Pääbo et al., 2004).

Table 1.1 The frequently occurred damages on the aDNA and their effects with the possible avoidances from these effects (Pääbo et al., 2004).

Damage type	Effects on DNA	Possible avoidance
Breaking of the strands	Reduction of sizes and DNA amounts	PCRs of the overlapping short fragments
Damage on the bases and deoxiyribose residues	Fragmentation of bases and sugar	PCRs of the overlapping short fragments
Amino group losing: Conversion; from cytosine to uracil or from 5-methyl-cytocine to thymine	Coding changes (Misincorporation)	Multiple repeated PCRs and sequences independently or UDG (uracil-DNA-glycosylase) treatment

As a type of aDNA damage; loss of amino groups on the bases which is called as the hydrolytic loss of amino groups from bases, causes the so called nucleotide misincorparations (corporation of incorrect nucleotides to sequences during the PCR amplification of aDNA). The most common hydrolytic loss of amino groups are from

the cytosine to uracil and from the 5-methyl-cytosine to thymine (Friedberg et al., 1995; Pääbo et al., 2004). After the cytosine deamination the nucleotide thymine (T) rather than cytosine (C) is incorporated during the amplification of aDNA region by PCR. Similarly, sometimes, adenine (A) rather than guanine (G) is incorporated during the PCR amplification (Hansen et al., 2001) of the aDNA molecule. Moreover the uracil that was formed as a result of the cytosine deamination in aDNA, could be removed using an enzyme which is called as uracil-DNA-glycosylase (UDG) (Pääbo, 1989).

There are two types of the nucleotide transitions (Hansen et al., 2001); Type1 is the changes from A to G (A→G) or T to C (T→C) and known as the PCR artifacts, while Type2 is the changes from G to A (G→A) or C to T (C→T) and indicates the authenticity of ancient samples (Olivieri et al., 2010). Moreover transversions (changes from a pyrimidine nucleotide to a purine nucleotide or vice versa) and indels in aDNA sequences can also be observed.

1.5 A Short History of Ancient Anatolia and Çemialo Sırtı Excavation Site

During the Last Glacial Maximum (LGM, 18.000-16.000 years ago) when the northern Eurasia was covered with ice sheet the populations were driven to Iberia, Italy and Anatolia where conditions were favourable for human life (Hewitt, 2004; Underhill et al., 2001; Cinnioglu et al., 2004). After this cold period, first of all, human populations which were in Anatolia started to migrate to the northwest and northeast of Turkey. Then, Turkey being at the cross roads, hosted many civilizations starting from ancient times.

Most importantly, central and southeast Anatolia witnessed a quite important transition stage of the humanity: Transition from hunter gatherer life style in the Palaeolithic age to farming life style in the Neolithic age. Some of the known Palaeolithic and Neolithic sites are given in the Database of Archaeological Sites in

Turkey (http://taygis.tayproject.org/TAYGIS_ENG/TAYGISeng.html). As it can be seen from the database the number (4273) of sites is immense.

Many populations; the Hattians, Hurries, Hittites, Phrygians, Lydians, Urartians, Persians, Meds, Romans, Sassanids, Byzantines, Seljuk Turks, and Ottomans had chosen the Anatolia as their homeland (Akurgal, 2003).

At the beginning of the 13th Century BC many migrations took place from Balkans to Anatolia such as the migrations of Phrygians and Ionian Greeks (Akurgal, 2003; Tambets et al., 2000). Then the Western part of Anatolia was occupied by Lydians between 900-547 BC and the Eastern by Medians between 678-549 BC (Diakonoff, 1985; Akurgal, 2003; Tambetset al., 2000; http://www.iranchamber.com/history/median/median.php). Then the Persian Achaemenid kingdom (550–330 BC) dominated the Anatolia (Tambets et al., 2000; David, 2005; Tavernier, 2007; Sampson, 2008). Between 550-330 BC most parts of Anatolia were ruled by the Persian Empire until the armies of Alexander the Great conquered Anatolia from the Persians in 334–333 BC (Green, 1990; http://www.timemaps.com/history-anatolia). After that the Hellenistic Period began and it lasted until31 BC and ended with the conquest of the last Hellenistic kingdom by Rome (Green, 1990; McLean, 2002).

The control of Anatolia was further followed by the Romans (27 BC-395 AD) and the Byzantines (330-1453 AD) (Tambets et al., 2000). On the other hand, Central Asia started to be populated by Turkic speaking nomadic groups in the 5th-7th Century AD (Roux, 1997) and a distinct nomadic group who were known as Göktürks (T'ukü-e) emerged in Mongolia region (Roux, 1997). It was a Turkic speaking nomadic group who used the word "Türk" for the first time in the history, as a political name (Roux, 1997).

The Pechenegs, Uz and Kipchaks who were the Turkic speaking groups controlled the Northern Black Sea region, started to migrate to Eastern Europe and to the

Balkans between the 9th and 11th Century AD (Roux, 1997; Salman, 2004).These Turkic speaking groups migrated not only to Europe but also to the Near East and Anatolia (Rossabi, 1994).

The last well-known cultural influence of Central Asian migrations, officially, took place around the 11th Century AD with the migrations of Turkic speaking nomadic groups called as the Oghuz Turks into the area resulted in cultural influence and the language replacement in the region. (Lewis, 1995; Long, 1991; Vryonis, 1971).

In the present study, ancient populations (within the limits of their few representatives) from the Near East will be compared with the ancient populations from Europe, with the modern populations of the same region as well as with some of the Central Asian, and European populations on the basis of their mtDNA sequences or haplogroups. The gene pools and mtDNA compositions as a part of gene pools of human populations are expected to change over the time due to drift or migrations. For instance, the modern Turkish population might have had mtDNA contribution from Central Asia through the Turkish speaking nomads arriving to Anatolia nearly after 1000 AD. It is hoped that the direction of deviation from the modern Turkish population, for instance, from Çemialo Sırtı population can be interpreted in the line of known episodes of the history.

1.5.1 Çemialo Sırtı

Figure 1.2 Map showing the location of Çemialo Sırtı archaeological site in Batman where the samples of the present study were unearthed. The picture of the excavation site provides a general view of the site. The picture was taken by Sidar Gündüzalp.

Çemialo Sırtı is in the southeast Anatolia, 1 km southeast of the Yazıhan Village of Gedikli District and on the west of Garzan stream in the border of Batman province. The Çemialo Sırtı is located on the hillside of Late Pleistocene low conglomerates which are about 10-20 meters from the base of Garzan Valley (Figure 1.2) placed at

the altitude of 503-512 m high. Çemialo Sırtı is one of the sites around Ilısu Dam region which has the great probability of being flooded by the Ilısu Dam directly.

Excavations have been started in 2009 by Assoc. Prof. Dr. Aslı Erim Özdoğan from Çanakkale University, Çanakkale, Turkey. The oldest history of the site was dated back to approximately 2000 BC. Between the excavation seasons of 2009-2013, three different periods were uncovered at this site. The first period belonged to Artuqid period covering the dates between 1102-1231 AD, 2nd period covered the Iron Age which was around 1000 BC which also covers the Late Iron Age approximately 600-500 BC and the 3rd period was the Middle Bronze Age dated back to between 1500-2000 BC. Moreover, the preliminary surveys suggested that there were remains of Early Bronze and Chalcolithic Ages, however, they were perhaps intertwined.

Most of the burial places, from where the samples of the present study were collected, belong to one of the settlements in the site which was approximately dated back to 600-500 BC. Three samples from the site, very soon, are going to be sent to C14 aging for the accurate timing of the period. Nevertheless, before the given time period the region was dominated by Medians between 678-550 BC (Diakonoff, 1985). Their area of distribution on the map is given in Figure 1.3. The Median kingdom was taken under the control in 550 BC by Cyrus the Great, who found the Persian Achaemenid Empire (David, 2005; Tavernier, 2007; Sampson, 2008). The given time period is when the Persian Achaemenid Empire was ruling the region (David, 2005; Tavernier, 2007). In Figure 1.4 distribution of Achaemenid Empire was shown. In the site, buried individuals exhibited two different positions: one is in hocker (position of the unborn fetus) position and placed in jars, another one is in semi-hocker or dorsal (laid flat) position in stone cist graves. One of the interesting findings at this site is that; a silver Tetra Drachma was found which is dated to Hellenistic Period between 330-270 BC. In the Figure 1.5 the conquests of Alexander the Great in 323 BC as the Hellenistic settlements in the region is depicted.

Figure 1.3 The map showing the Median Empire (678–550 BC) before the given time period for the region. The figure was taken from Ian Mladjov's Resources from the University of Michigan's History department (http://www.iranchamber.com/history/median/median.php).

Figure 1.4 The map showing the Persian Achaemenid Empire (550–330 BC) in the given time period for the region. The figure was taken from Ian Mladjov's Resources from the University of Michigan's History department (http://sitemaker.umich.edu/mladjov/files/persia500tbc.jpg).

Figure 1.5 The map showing the Hellenistic settlements in the region (323 BC) during the time period of Alexander the Great. The figure was taken from Ian Mladjov's Resources from the University of Michigan's History department (http://sitemaker.umich.edu/mladjov/files/alexandros323nbc.jpg).

1.6 Objectives of the study

The main objectives of the present study are:

1. To sequence ancient human mtDNA HVI and HVII regions (in total 576 bp long region) on the basis of 5 partially overlapping fragments and to determine the mtDNA haplogroups of the individuals, in the dedicated aDNA laboratory of Middle East Technical University in Ankara.

2. To reveal preliminary information about the mtDNA haplogroup composition of the Çemialo Sırtı population, on the basis of individuals who lived, presumably, between 600-500 BC in Çemialo Sırtı.

3. To contribute to the understanding of history of population in Anatolia by using comparative analysis between Çemialo Sırtı population, various modern populations (such as modern Turkish, Cypriot and Iranian) and ancient populations from the region (Ancient Sagalassos population, Neolithic population from Syria, other Neolithic populations, from south eastern (STA, LBKT) and central European regions (LBK, RSC, SCG, BAC, SMC, BEC, CWC, BBC, UC) and the Gurgy population from Paris basin.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

2.1.1 Sample Collection

In ancient DNA studies, archaeological samples, especially human remains, are susceptible to contamination from modern DNA. On the other hand, some sample types are less risky for the exogenous contamination and can contain sufficient amount of DNA. Thus, choice of suitable sample types for aDNA studies are important and there are some criteria for this selection. For example, suitable samples should be particularly hard, well-preserved and compact; such as long bones, ribs, humerus, femur and some parts of the skull as petrous bone. Furthermore, teeth samples (without fissures, caries, or wear) are another type of preferred samples, and also less risky for the contamination than other sample types.
All of the samples used in the present study were collected from Çemialo Sırtı excavation site by archaeological team members Dr. Yasemin Yılmaz, Sidar Gündüzalp, Aliye Gündüzalp and Esra Çiftçi. The samples were provided by anthropologists Dr. Yasemin Yılmaz from Düzce University Department of Archaeology.

The handling and storage methods described in the Ottoni et al.'s (2011) study were followed with slight changes during the onsite sample collection, to avoid the contamination. The procedure for sample collection is given below:

1. Samples might be less preserved at sites located in warmer climates. Thus, the samples should be kept and stored in a cool, shady and dry place immediately after they are unearthed and each sample should be placed directly in a single-use sterile plastic bag or box with its sediments and without applying any archaeological procedure. In other words, after collecting the samples, washing, brushing, use of consolidants or chemicals procedures routinely performed by archaeologists, should not be carried out for the samples which will be used for aDNA isolation.

2. Ancient human remains can be easily affected by modern human contamination. Thus, to keep the contamination risk at the possible lowest level, excavators follow a number of steps when unearthing and collecting the ancient samples. When the human skeleton is unearthed, samples from a part of skeleton preferentially should be handled by only one anthropologist. The anthropologist should wear sterile and single-use face mask, bonnet, a pair of gloves, over sleeve and (if possible) clean room suit during the handling process to avoid contamination not only from modern DNA but also from environmental DNA. Gloves should be changed between handling of samples from each individual. Tools which are used to collect samples should be decontaminated with bleach between two different individuals to avoid cross contamination.

3. There are additional measures taken by the archaeological team members of the Çemialo Sırtı to avoid the contamination during the sample collection. The skeletons are usually unearthed by tracking bones of the individual which could be in a position either dorsal (laid flat) or hocker (mother's posture unborn fetus). Thus the places of the teeth or phalanx samples or the aDNA extraction, are already known prior to the opening of the part where the samples will be collected. The archaeological team members, who will take the teeth or phalanx samples wearing the clothes which are for contamination avoidance, firstly open the unexcavated part of the grave single-handedly and then collect the samples as explained above. The sample is picked up right before finishing the unearthing procedure on the day of

excavation. (This information was provided by Sidar Gündüzalp). The samples are kept in the fridge at +4°C until they are brought to the aDNA laboratory.

2.1.2 Storage of Samples

Following sample collection, when the samples arrived at the aDNA laboratory, each sample was put into a new single-use sterile plastic bag and labelled with all of the information relevant to the sample, written on the old plastic bag/box that were from the archaeological site. Moreover the old bag/box of each sample was also kept in case of any information needed. All of the samples were stored in the fridge at +4°C during the pre and post experimental periods. After the samples were performed when the samples remained, each sample is was weighed and labelled.

2.1.3 Type of the Samples

In the present study the teeth and phalanx (finger bone) samples were used. However, the teeth samples are known as more conserved remains, thus, may contain higher amount of DNA. In the present study, 8 teeth and 1 phalanx samples from Çemialo Sırtı excavation site were used for aDNA extraction and following procedures such as PCR amplification and sequencing. The samples used for aDNA extraction and their information is given in Table 2.1. Also the pictures of the samples taken in the excavation site and aDNA laboratory are shown in Figures 2.1-2.2, respectively.

Table 2.1 The approximate archaeological dates, types and excavation seasons of the Çemialo Sırtı samples (n=9) which were used for the further analysis in the present study.

Approximate archaeological date (BC)	**Sample**	**Sample type**	**Excavation season of sample**
600-500	SK24	Tooth	2013
600-500	SK26	Tooth	2013
600-500	SK28	Tooth	2013
600-500	SK29	Tooth	2013
600-500	SK30	Tooth	2013
600-500	SK7	Tooth	2013
600-500	SK17	Tooth	2013
600-500	SK21	Phalanx	2013
600-500	SK35	Tooth	2014

SK24 SK26 SK28 SK29 SK30

Figure 2.1 The skeletons from which the samples SK24, SK26, SK28, SK29 and SK30 were collected at the Çemialo Sırtı excavation site.

Figure 2.2 Pictures of the samples in the aDNA laboratory after the surface cleaning step which is performed in the pre-extraction room.

2.2 Methods

2.2.1 Avoiding Contamination in the aDNA Laboratory

All of the pre-PCR experiments, as the pre-extraction, extraction and PCR preparation procedures, except post-PCR steps, were carried out in a dedicated aDNA laboratory at the Middle East Technical University. This dedicated aDNA laboratory is in a physically separated building away from the post-PCR and modern DNA laboratories. The modern DNA has been neither introduced to the aDNA laboratory nor studied in it. Furthermore the entrance to the aDNA laboratory was confined to only four researches (RY, NDD, EY and FO). All of the researchers had access to the aDNA laboratory after wearing sterile, single-use overall suits, over-shoes boots, double-pairs of gloves, surgical face masks, bonnets inside the clean-entrance room of the aDNA laboratory. Moreover some irreversible steps were followed by researchers to avoid the contamination. For example; researchers were not allowed to enter in aDNA laboratory if the PCR products were handled on the same day in the post-PCR laboratory (modern DNA lab). The changing step of old plastic bags of samples with new ones in the entrance room of the aDNA laboratory (as explained in section 2.1.2 (Storage of Samples)) were applied without touching samples neither with gloves. Furthermore the gloves were changed after mistakenly touching any of samples.

All of the laboratory instruments used during the experiments in the aDNA laboratory were cleaned with 5% bleach and RNAse AWAY (decontamination reagent) and also regularly irradiated to UV light at 254nm wavelength in the cross-linker for one hour. The solutions that are not sensitive to UV irradiation such as nuclease-free H_2O, buffer, $MgCl_2$, BSA were also UV-irradiated for 10 minutes (but not, for example, proteinase K, which is sensitive to UV). Moreover, before entering to the clean room each object including plastic items was UV-irradiated and washed with bleach. However, if it is a metal item or reagent tube, it was cleaned with RNAse AWAY in the entrance room. In both pre-extraction (grinding) and

extraction (clean room) rooms, all the benches were also regularly wiped with 5% bleach solution after each experiment prior to sterilization with overnight UV irradiation.

Surface cleaning of the samples, DNA extraction and PCR preparation steps were carried out inside different UV-irradiated biosafety cabinets. Surface cleaning and grinding steps were performed in a separated pre-extraction (grinding) room. During the experiments three kinds of negative controls were used; one HYA (synthetic Hydroxyapatite powder) control for each grinding step, two blanks (negative controls) for each extraction and PCR amplification steps. The first blank is placed at the beginning and the second blank is placed at the end of the sample set and the same order was always followed during the other steps. This order was used to decrease cross-contamination risk and to detect exogenous modern DNA contamination.

At least two DNA extractions (if there is enough amount of sample for two extractions) and three PCR amplifications from each extraction were carried out for each sample. Following this, two strands of aDNA were sequenced in forward and reverse directions to check the reproducibility and the accuracy of the results. If there are more samples from same individual, independent extractions were performed using the different samples from the same individual. Moreover, DNA-free, nuclease-free reagents and UV-irradiated materials were used for all of the steps during the experimental procedures.

To prevent post-PCR contamination in the aDNA laboratory from the amplified PCR products, PCR and post-PCR steps were performed in the post-PCR laboratory, which is at a physically separated building.

Furthermore, the buccal swab samples of the aDNA laboratory researchers and archaeological team members who handled the samples at the excavation site were taken. Afterwards, DNA from the buccal swab samples was extracted, the mtDNA

control region (HVI and HVII regions, as the same region with ancient samples) of the buccal swab samples were amplified and sequenced after informed permission to detect contamination from modern DNA.

Figure 2.3 Working in the aDNA laboratory

2.2.2 Sample Preparation

Prior to aDNA extraction, the sample preparation step, was performed inside a Class 2A Biosafety cabinet (MetiSafe) in the pre-extraction room of the aDNA laboratory. The cabinet was UV-irradiated (254 nm) for minimum of 20 minutes (min) to prevent contamination from previous use. Inside the cabinet, the outer surface of tooth/bone was removed on a different piece of aluminium foil for each sample using a sterile blade or drill (Proxxon Micromot 40/E) with a sterile single-use cutting discs (Proxxon, catalog numbers: NO28812 and NO28830). The drill was used at the minimum speed to avoid overheating which may cause further DNA damage. When the sample was very hard the drill was used for surface cleaning. Otherwise surface

cleaning was performed using sterile blades. Furthermore for teeth samples, surfaces were gently wiped with 5% bleach solution and rinsed with UV-irradiated nuclease-free water. During the surface cleaning, in between two samples and also at the end of surface cleaning work in a given day, the materials were cleaned with RNAse AWAY solution and distilled water (dH_2O), respectively, while the work area (the inside of cabinet) was cleaned with first Alconox solution and then RNAse away solution.

Each side of all the samples were UV-irradiated inside the cross-linker for 50 min (under 254 nm wavelength, approximately at 5 cm distance). Samples were ground to fine powder in the 6750 Freezer Mill and 200-300 mg of powder was transferred into 2 mL eppendorf tubes, as one tube for each extraction and labelled. The remaining powder of each sample were also stored at 4°C until the second extraction. If the sample is big, a piece of the sample was ground each time; otherwise the entire sample was used in one grinding step. Furthermore, 200 mg of HYA was used as a negative control for grinding step to check cross contamination among samples and modern DNA contamination from the researcher.

2.2.3 aDNA Extraction

DNA extractions of the ancient samples were carried out using the silica-based spin column method following the protocol in Ottoni et al.'s (2011) study with slight modifications. The protocol is given below;

The Lysis Buffer which is a mixture of 1.6 mL EDTA (0.5M, pH8, Sigma) and 8 µL proteinase K (20 mg/mL, Thermo Scientific),was prepared inside the laminar flow cabinet. Then inside the cabinet, 1.608 mL of Lysis Buffer was added to each tube containing 200-300 mg sample powder and to two tubes of extraction blanks (negative controls). Following that the tubes were vortexed and placed into the rotatory wheel which is placed in the incubator (Fine PCR Combi-SV12). Then they were incubated for 24 hours at 56°C and then 24 hours at 37°C. After the incubation,

the samples were centrifuged for 5 minutes at 13,000 rpm in the microcentrifuge (Thermo Scientific MicroCL17) and the supernatant was added to the tubes containing 6 mL binding buffer (QG of QIAquick Gel Extraction Kit, Qiagen and Isoamyl Alcohol). The tubes with pellet were stored at -20°C. After vortexing, the samples were transferred into the spin columns which were already attached to the extender tubes of the manifold (Qiagen vacuum pump and manifold) and the sample solutions were passed through the spin columns by vacuuming. Then the washing step was performed by applying 4 mL Washing buffer (QG and PE of QIAquick Gel Extraction Kit, Qiagen) to each extender tube. After that, the columns were dried for 5 minutes at room temperature inside the work area and then the columns were placed into new 1.5 mL Low-Bind eppendorf tubes. Following that the Elution Buffer (EB of QIAquick Gel Extraction Kit, Qiagen) was added onto the centre of each filter in three times intervals (55 uL + 55 uL + 26uL). Then the tubes were incubated for 5 minutes at 56 °C and centrifuged for 2 minutes at 11,800 rpm in the microcentrifuge and this procedure was repeated for each of three times intervals. If the DNA extracts will be used soon the tubes were stored at +4°C, if not, stored at -20°C for the long term storage.

2.2.4 aDNA Amplification

The human mtDNA Control region (non-coding region) contains the first and second Hypervariable regions (HVRI-HVRII or HVI-HVII regions) between the nucleotide positions (np) 16000-16569 and 1-400 for the HVI-HVII regions, respectively, according to the Cambridge Reference Sequence (CRS; Anderson et al., 1981). In the present study, 359 bp of the HVRI between np 16008–16366 and 217 bp of the HVRII between np 49–265 were studied and amplified with 5 primer pairs as 5 overlapping fragments (A-B-C-D-E) for HVI region and 2 primer pairs as 2 overlapping fragments (F-G) for HVII region. In total PCR amplifications of the mtDNA HVI-HVII regions were performed jointly as in the form of 7 fragments using 7 primer pairs. Sequences of the primers and length of the fragments were

taken from Ottoni et al.'s (2011) study. All of the fragments with their length and sequences of the primer pairs were given in the Table 2.2.

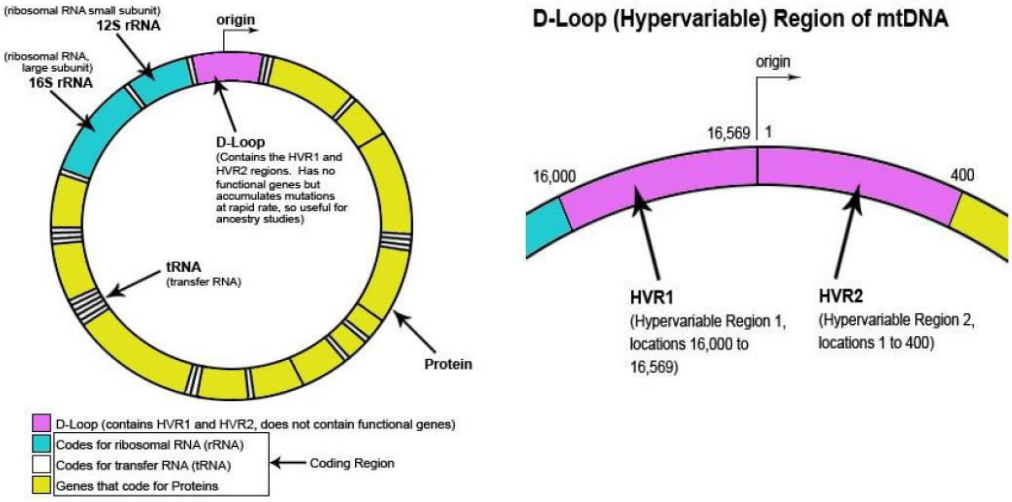


Figure 2.4 The complete human mitochondrial genome including coding and noncoding (Control region or HVRI-HVRII) regions. In the second part of the figure Hypervariable region (D-loop) of human mtDNA was shown in detail. The figure was taken from Genebase Tutorials (http://www.genebase.com/learning/article/17).

Table 2.2 Seven overlapping fragments of mtDNA HVRI-HVRII (HVI-HVII regions) and the forward and reverse primer sequences with their lengths (base pair, bp) (F16008- R16108, F16099-R16201, F16157-R16228, F16208-R16306, F16285-R16366 for HVRI, F48-R155, F120-R285 for HVRII) (Ottoni et al., 2011).

mtDNA regions	Primers	Primer sequences	Length of fragments (bp)	Fragments
HVRI	F16008 R16108	5′- CCCAAAGCTAAGATTCTAAT -3′ 5′- GTACAATATTCATGGTGGCT -3′	140	A
HVRI	F16099 R16201	5′- CCGCTATGTATTTCGTAC -3′ 5′- TTGATTGCTGTACTTGCTTG -3′	139	B
HVRI	F16157 R16228	5′- ATACTTGACCACCT -3′ 5′- TTGCAGTTGATGTGTGATAG -3′	109	C
HVRI	F16208 R16306	5′- CCCCATGCTTACAAGCAAGT -3′ 5′- GTAAATGGCTTTATGTGCTATG -3′	133	D
HVRI	F16285 R16366	5′- CCCACTAGGATACCAACAAA -3′ 5′- TGAGGGGGGTCATCCATG -3′	118	E
HVRII	F48 R155	5’- CTCACGGGAGCTCTCCATGC -3’ 5’- TGAACGTAGGTGCGATAAATA -3’	147	F
HVRII	F120 R285	5’- CGCAGTATCTGTCTTTGATTCC -3’ 5’- GTTATGATGTCTGTGTGGAA -3’	166	G

PCR amplification of the mtDNA HVI-HVII regions were performed in 50 µL (5 µL template DNA with 45 µL PCR mix) reaction volumes. The dUTP mix which contains uracil deoxynucleotide instead of thymine deoxynucleotide, was used in the PCR mixture to be reacted with UDG (Uracil-DNA-Glycosylase) enzyme. The details of the reagents with their volumes given in Table 2.3.

Table 2.3 PCR reagents with their volumes for amplification.

Nuclease free H_2O (Molecular biology grade)	31.5µL
Buffer (10x,UV) (AB)	5.0 µL
$MgCl_2$ (25 mM, UV) (AB)	5.0 µL
dUTPs mix (Bioline) (10 mM/each)	1.0 µL
Primers (FWD+REV) (10 µM/each)	1.0 µL
BSA (UV) (50 mg/mL)	0.5 µL
UNG (1U/µL)	0.5 µL
AmpliTaq Gold 360 (5 U/µL)	0.5 µL
DNA template	5.0 µL

PCR thermocycler (ABI) was used for amplifications with the reaction conditions described in Table 2.4.

Table 2.4 Thermocycling conditions of the PCR amplification.

UNG activation	37 °C	15 min	
Initial denaturation UNG inactivation	94 °C	10 min	
Denaturation	94 °C	45 sn	60 cycles
Primer Annealing	54-53 °C	1.15 min	
Extension	72 °C	1 min	
Final extension	72 °C	5 min	
Base temperature	4 °C	∞	

2.2.5 Gel Electrophoresis for the Amplified aDNA Fragments

PCR products were visualized on the 3% agarose gel which was prepared with0.5x TBE buffer and 4 µL Ethidium bromide. Five µL of each PCR product was mixed

with 3μL 2x Blue loading dye and loaded into the each well of the gel. Then the agarose gel containing PCR products was run at 95 V approximately for 60 min and visualized under the UV light using Vilber Lourmat CN-3000WL. Moreover 50 bp DNA ladder (Thermo Fisher Scientific, Waltham, MA, USA) was used as the reference to check if the sizes of the PCR products were as; 140 bp for fragment A, 139 bp for fragment B, 109 bp for fragment C, 133 bp for fragment D, 118 bp for fragment D, 147 bp for fragment F and 166 bp for the fragment G. The DNA ladder is a reference DNA which is composed of the DNA fragments with different known sizes (50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000 bps).

2.2.6 Purification of PCR Products and Sequencing of aDNA Fragments

After checking the PCR products on the gel, successfully amplified PCR products were purified using ExoSap-IT (USB Affimetrix) enzyme following the manufacturer's protocol to remove any excess dNTPs, salts, primers and enzymes. The procedure is as explained: 2μL of ExoSap-IT was mixed with the 5 μL of PCR product. The ExoSap-IT is an enzyme which treats the PCR products ranging in size from less than 100 bp to over 20 kb, with absolutely no sample loss while removing unused primers and nucleotides. This purification method is commonly used for the purification of the human mtDNA sequences. The thermocycling for the purification was performed using a PCR thermocycler (Biometra-Thermocycler T1 Thermoblock) with the cycle conditions of ExoSap-IT reaction which were given in Table 2.5.

Table 2.5 Cycle conditions of the ExoSap-IT purification method.

37 °C	15 min
80 °C	15 min
4 °C	∞

2.2.7 Molecular Analyses of Ancient Samples

2.2.7.1 Alignment of the fragment sequences

In the present study, forward and reverse sequences of the overlapping fragments of human mtDNA HVI-HVII regions were obtained and aligned with CRS (Anderson et al., 1981) using the BioEdit 5.0.9 software (http://mbio.ncsu.edu/bioedit/bioedit.html). Both of the forward and reverse sequences of the same fragment were used to obtain one correct fragment sequences of a sample by comparing the two sequences which are forward and reverse. When the beginning of the forward sequence did not have successful chromatogram peaks or sequences, the beginning of the reverse sequence was used for the assemblage of aligned fragment sequence or vice versa for the end of reverse sequence. Once more, the forward and reverse sequences were alternately used to read the sequence of each fragment correctly. Then all the repeated sequences (i.e, the same aDNA region is sequenced multiple times) of each fragment were aligned and all of the overlapping fragments of each region were assembled to obtain partial HVI-HVII regions of each sample. Furthermore, using BLAST tool of GenBank database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) the sequences were checked if the sequences best match with humans'.

2.2.7.2 Haplogroup Determination

Mitochondrial DNA Haplogroups (HPGs) of Çemialo Sırtı individuals were determined according to the nucleotide differences at certain nucleotide positions. These nucleotide differences and nucleotide positions were detected by aligning of the all fragment sequences of each individual with CRS as reference sequence. Then, haplogroups were assigned considering the mutation motifs at the PhyloTreemt website which is regularly updated human mtDNA haplogroup tree (http:// phylotree.org, 19 February 2014). The mutation motifs for the HVI region (HVRI), which were used to determine macrohaplogroups of the Çemialo Sırtı individuals,

are listed in Table 2.6. This table of the HVRI mutation motifs were also used to assign the macrohaplogroups of the modern and other ancient population data retrieved from literature and databases, and used for comparative analysis in the present study. Moreover, CRS was taken as reference sequence and does not contain any of the mutation motifs.

Furthermore, to detect the subhaplogroups of the ancient samples, the HVRII mutation motifs were also used along with the HVRI mutation motifs. The mutation motifs of the HVRII observed among the Çemialo Sırtı individuals are listed in Table 2.7.

Table 2.6 The list of mutation motifs on the sequence of mtDNA HVRI used to determine haplogroups of the population according to human mtDNA phylotree (http://phylotree.org/). Numbers are the nucleotide position and the capital letters indicate the transition mutations from the first to second letter. The little letters indicate the transversion mutations from the first capital letter to second little letter. "+" : The mutation motifs that must exist in each haplogroup, "/ /" : Each different mutation motif could be possible to be grouped in the same haplogroup," ( )" : At least one of them, " ! ": Back mutation (the reversions to an ancestral state) from the first letter to second letter at that nucleotide position, "*" : The major group of that haplogroup.

Haplogroups	HVRI Mutation Motifs
A	C16290T + G16319A + (T16362C, C16242T, T16311C!, C16223T, C16292a, C16344T)
B	T16189C + (C16111T, T16140C,C 16147T, T16217C, T16243C, T16249C, C16261T, C16184a, A16235G)
C	T16298C + C16327T + (C16239T, C16261T, T16288C, T16093C, C16223T, G16129A!, C16344T, T16357C)
D	T16362C + (T16093C+C16184T, C16223T, C16234g, C16294T , T16311C!, G16319A, G16129A!, C16245T, T16189C!, T16325C)
F	T16304C + (T16172C, G16129A!, T16189C!, C16232a, T16249C, T16311C!)
G	16362C + C16223T + (C16114a, A16227G, G16274A, C16278T!, C16214T, A16265c, A16293c)
H	A16293G / C16259T / A16051G + C16259T / C16287T / T16304C+C16291T / C16114T / G16153A+T16362C / A16162G /A16302G / C16294T / C16291T / C16111a / T16189C! + T16356C / A16227G / A16265c / A16265c / G16274A / T16325C / T16140C / T16124C / A16051G / T16288C / A16293G + C16278T!+ T16311C! / C16354T / C16148T / C16355T / T16093C / C16239T / C16239g / C16168T / C16176T / T16362C +A16482G / A16240G / T16172C / C16218T / T16189C! + T16356C+ T16362C / A16299G / C16270T / C16256T / T16209C / C16184T / A16300G / G16319A / C16261T / C16188T / C16278T! / C16266T / T16092C / G16390A / G16145A / C16234T / C16192 / T16304C+C16294T / T16304C+C16192T / A16037G / C16344T / C16344T+C16114T / G16244A / A16227G+G16145A / C16320T / C16320T+C16294T / A16070G / C16256T + T16352C / T16352C / C16328a / T16263C / A16080G / C16169T / C16270T / T16126C / T16312C / T16368C / C16291T + G16390A / T16325C+ C16360T

Table 2.6 (continued)

HV	C16221T / T16298C / C16067T+G16274A / C16067T / C16067T+C16355T / T16217C+C16214T +A16335G / C16069T / G16129A! + T16221C! / C16221T+C16291T / T16298C +T16311C! / T16298C +T16126C / T16298C+G16274A / T16356C / C16327a / A16113c / T16311C! +T16172C+A16113c / C16292T / T16189C!+C16067T / C16356T / T16220C + C16292T
I	G16391A + G16129A+ (T16172C, C16148T, T16311C!, C16223T, G16319A, T16075C, T16086C, C16169T)
J	T16126C + C16069T
K	T16224C + T16311C!
L0	T16172C+ (C16188g, G16319A, G16188A, T16278C, G16188A, C16223T)
L3	T16278C + C16311T+ (T16124C, C16223T)
L*	T16172C + C16223T + T16311C! / C16214T+G16274A+C16223T+T16311C! / A16215G+C16223T+T16311C! / A16207t / A16207t+C16223T+G16230A / G16145A+C16223T
M	G16129A! + C16291T + C16223T+16298C / T16126C+T16154C +C16223T+T16224C / G16129A!+T16189C!+C16223T+ T16249C+T16311C!+ (T16359C, C16185T, A16399G) / T16325C+C16223T / T16086C+C16223T+T16362C / G16129A! + C16223T+C16256T+T16362C / C16111T+C16223T / G16129A!+C16223T+T16311C!+ (T16362C+T16093C) / T16298C +G16319A+C16184T+C16223T / C16295T+C16223T / G16129A!+C16223T+T16297C+ (T16298C, A16309G) / C16327T+C16223T / T16172C+C16192T+C16223T / A16293c +C16223T
N	C16147g +T16172C+ G16145A +C16223T+C16248T +C16355T / C16176g +G16390A+G16145A+C16223T / T16086C +T16172C +C16187T!+ T16189C! +T16217C+ C16223T / C16257a +C16261T+C16223T+(G16129A!+C16111T)
R0	T16362C+ (T16126C, C16301T, C16355T, C16114T)
R	C16071T / T16304C / T16304C+ (G16145A, T16172C, C16266T, T16356C, T16325C) / C16266T+ T16311C! / T16304C+ C16192T + (A16309G, G16390A)

Table 2.6 (continued)

T	T16126C+C16294T+ (C16296T, A16163G, G16153A, C16186T, C16292T, T16243C)
U1	T16249C+(T16189C!, C16111T, C16327T, G16129A!, G16145A, T16362C, T16311C!, C16242T, A16166d)
U2	A16051G + (A16206c, A16230G!, T16311C!, G16129c, T16189C!, T16362C, C16234TT, C16168T, G16145A, C16256T, C16179T, C16266T)
U3	A16343G / A16343G + (G16390A, C16168T, T16362C)
U5	C16192T / C16270T + (C16256T, A16399G, C16291T, G16129A!, T16249C, T16189C!, T16304C, C16114a, C16294T, C16114T)
U*	C16355T / T16172C / T16356C + (C16134T, C16179T) / A16309G + A16318t / T16342C / T16189C! +C16234T / C16278T! + (T16172C, T16311C!, T16136C) / A16146G / A16146G+T16342C
V	T16298C + (A16183G, A16216G, C16290T, A16240G, A16203G, A16162G, G16153A, A16219G, C16301T, C16261T + T16311C!, G16153A)
W	C16292T + C16223T+(C16192T, T16325C, C16286T)
X	T16189C! + C16278T! + C16223T + (C16104T, T16278C!, G16213A, C16248T, G16255A,T16126C, C16189a)
Y	T16126C + T16223C + T16231C + (C16266T, T16189C, T16311C!)
Z	T16298C + C16185T + C16260T + C16223T! + (G16129A!, T16224C, A16302G, C16193d)
CRS	

In the present study, the mtDNA HVRI data which was retrieved from modern and ancient populations, categorized in 30 groups using mutation motifs of the HVRI region based on PhyloTreemt (http://phylotree.org/). These 30 macrohaplogroups were defined to be able to compare the haplogroup compositions of modern and other ancient populations with the Çemialo Sırtı individuals.

Table 2.7 The list of the mutation motifs for the sequences of mtDNA HVRII region were used to determine the subhaplogroups of the Çemialo Sırtı individuals.

Macrohaplogroups	Subhaplogroups	HVRII Mutation Motifs
H	H1a	A73G - A263G
H	H1z1	A263G
HV	HV	A263G
M	M1a1	A73G - A263G
R	R2	A73G - T152C - A263G
R	R6	T195C - A263G
U2	U2b1	A73G - T146C - A263G

2.2.8 Methods Used to Compare Populations on the Basis of Their mtDNA

2.2.8.1 Principle Component Analysis (PCA)

PCA (Principle Component Analysis) is a way to generate a graphical representation of the relationship between populations. It is used to visualize the genetic similarities of many populations in two or three dimensional space using the both sequences and frequency data (for instance frequencies of mtDNA haplogroups). In the PCA, individual axes, known as principle components (PCs), are extracted sequentially. In the PCA plots, the first PC explains the highest variation of all the data that can be accounted for by the compound axis (formed by the combination of the variables) while the second PC explains the next highest variation. The other PCs continue in the same way (Berkman, 2006).Thus, the most of the variation was explained in the

first two or three axes (Jobling et al., 2004). The schematic representation of PCA and graphics of five populations was shown in Figure 2.4.

In the present study two PCAs were carried out. The first PCA was performed based on HVRI sequences obtained from Çemialo Sırtı individuals and retrieved from 19 modern and ancient populations using the GenAlEx package program (http://biology-assets.anu.edu.au/GenAlEx/Welcome.html). The overlapping region of mtDNA HVI region sequences for all of the populations was determined as 240 bp in length between 16126-16366 np on the human mtDNA reference sequence, CRS. Sequences of this common region from all of the populations were used in the first PCA. The multiple sequence alignments were performed using ClustalW software (http://www.ebi.ac.uk/Tools/msa/).

The second PCA was performed using R program (http://www.r-project.org/; v. 3.1.2) and the "prcomp" function for categorical PCA and plotted on a two dimensional space using "biplot" function to display the first and second principal components. Both functions are included in the R "statistics" package. In the second PCA haplogroup frequencies data of each population based on the 30 haplogroups were used. The 30 haplogroups which were recategorized in the present study according to the common HVRI mutation motifs of the total data given in section 2.2.7.2 (Haplogroup Determination) of this chapter.

Total mtDNA HVRI data was consisting of HVRI sequences and haplogroup frequencies retrieved from modern and ancient populations. This total data were used in further analysis (for example, in PCA and fastsimcoal2 analysis) in the present study and given in Tables 2.8-2.9.

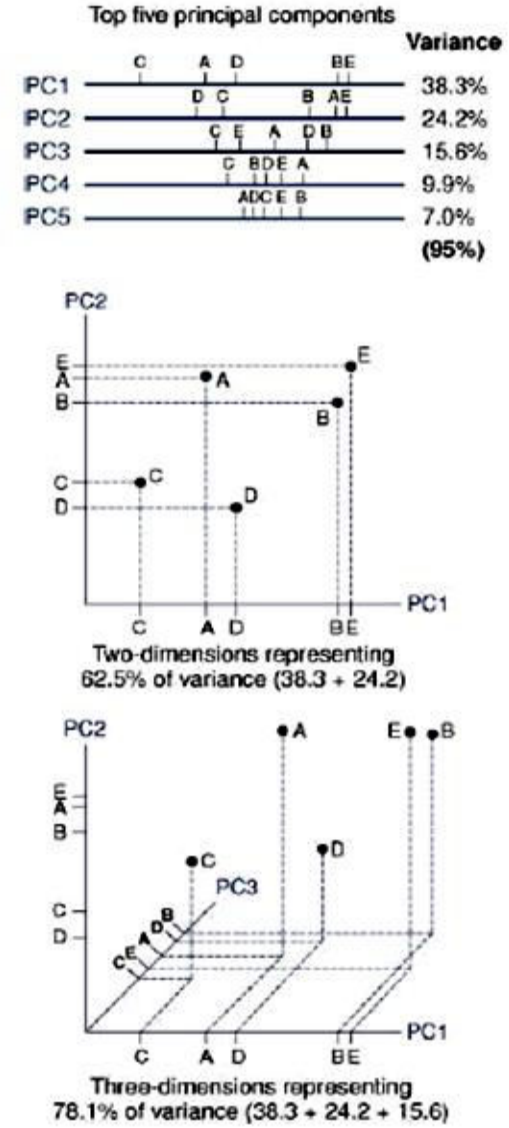


Figure 2.5 Schematic representation of PCA of five population in two and three dimensions (Jobling et al., 2004). The figure was taken from the Berkman's (2006) study.

Table 2.8 List of retrieved data from modern populations with their regions, sample sizes and related references. N: Sample size of each population, N_T: Type of the samples as the mtDNA HVRI sequence or haplogroup frequency data.

Region	Population	N	N_T	References
Balkans	Albania	34	Haplogroup frequencies	Belledi et al. 2000
	Bulgaria	668	Haplogroup frequencies	Karachanak et al.,. 2011
	Greece	320	HVRI sequences	Irwin et al., 2008
		243	Haplogroup frequencies	
	Macedonia	160	Haplogroup frequencies	Zimmerman et al., 2007
	Romania	91	Haplogroup frequencies	Bosch et al. 2006
Central Asia	Kazakhstan	44	HVRI sequences	Comas et al., 2004, Yao et al., 2000
		43	Haplogroup frequencies	
	Turkmenistan	20	HVRI sequences	Comas et al., 2004
		53	Haplogroup frequencies	Quintana-Murci et al., 2004
	Uzbekistan	20	HVRI sequences	Comas et al., 2004
		56	Haplogroup frequencies	Comas et al., 2004, Quintana-Murci et al., 2004

Table 2.8 (continued)

Turkey	Turkey	93	HVRI sequences	Alkan et al., 2014, Schoenberg et al., 2011 Roostalu et al., 2007, Comas et al., 1996
		130	Haplogroup frequencies	Alkan et al., 2014, Schoenberg et al., 2011 Roostalu et al., 2007, Quintana-Murci et al., 2004, Comas et al., 1996
Cyprus	Cyprus (Greek)	91	HVRI sequences	Irwin et al., 2008
		74	Haplogroup frequencies	
Southern Caucasus	Armenia	30	HVRI sequences Haplogroup frequencies	Schoenberg et al., 2011
	Azerbaijan	69	HVRI sequences	Schoenberg et al., 2011
		63	Haplogroup frequencies	Schoenberg et al., 2011
	Georgia	83	HVRI sequences	Nasidze I. 2001, Schoenberg et al., 2011
		24	Haplogroup frequencies	Schoenberg et al., 2011
Near East	Iran	50	HVRI sequences	Schoenberg et al., 2011, Comas et al., 2004
		421	Haplogroup frequencies	Schoenberg et al., 2011, Comas et al., 2004 Quintana-Murci et al., 2004
	Iraq	156	Haplogroup frequencies	Al-Zahery et al., 2011
	Syria	48	HVRI sequences	Vernesi et al., 2001
		46	Haplogroup frequencies	

Table 2.9 List of retrieved data from ancient populations with their regions, sample sizes and related references. N:Sample size of each population, $N_{T:}$ Type of the samples as the mtDNA HVRI sequence or haplogroup frequency data.

Region	Population	N	N_T	References
Southeastern Europe	STA (Early Neolithic Starčevo culture, 6000–5400 BC)	43	Haplogroup frequencies	Szécsényi-Nagy et al., 2015
	LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC)	39	HVRI sequences Haplogroup frequencies	Szécsényi-Nagy et al., 2015
Central Europe	LBK (Linear Pottery culture, 5500-4775 cal BC)	54	HVRI sequences	Brandt et al., 2013
		108	Haplogroup frequencies	
	RSC (Rössen culture, 4625-4250 cal BC)	7	HVRI sequences	Brandt et al., 2013
		11	Haplogroup frequencies	
	SCG (Schöningen group, 4100-3950 cal BC),	33	Haplogroup frequencies	Brandt et al., 2013
	BAC (Baalberge culture, 3950-3400 cal BC),	19	Haplogroup frequencies	Brandt et al., 2013
	SMC (Salzmünde culture, 3400-3100/3025 cal BC)	28	Haplogroup frequencies	Brandt et al., 2013
	BEC (Bernburg culture, 3100-2650 cal BC),	17	Haplogroup frequencies	Brandt et al., 2013
	CWC (Corded Ware culture, 2800-2200/2050 cal BC),	32	HVRI sequences	Brandt et al., 2013
		42	Haplogroup frequencies	
	BBC (Bell Beaker culture, 2500-2200/2050 cal BC)	20	HVRI sequences	Brandt et al., 2013
		32	Haplogroup frequencies	
	UC (Unetice culture, 2200-1550 cal BC)	88	HVRI sequences	Brandt et al., 2013
		88	Haplogroup frequencies	

Table 2.9 (continued)

Southern Paris	Gurgy population (5000-4000 cal. BC)	31	Haplogroup frequencies	Rivollat et al., 2015
Near East (Syria)	Neolithic population (8000-7300 BC)	10	HVRI sequences	Fernandez et al., 2014
		12	Haplogroup frequencies	
Southwestern Turkey	Sagalassos population (11th–13th century AD)	53	HVRI sequences Haplogroup frequencies	Ottoni et al., 2011

2.2.8.2 Fastsimcoal2 Simulations

The continuity between the populations was tested by calculating the F_{ST} value between the two populations, and comparing that with the F_{ST} values generated by coalescent simulations. Simulated DNA samples for the two populations (ancient and more ancient) are generated through fastsimcoal2 (Excoffier et al. 2013). We assumed that we sampled these populations serially through time under three population growth scenarios: no growth, instantaneous growth, and exponential growth. Instantaneous growth rate is calculated as the ratio of the ancient and more ancient population sizes, whereas the exponential growth rate is calculated as the natural logarithm of that ratio. We assumed that the generation time for humans is 25 years, and the mutation rate is $3.6x10^{-6}$ per site per generation for the human mtDNA HVRI (Richards et al., 2000). We carried out 1000 simulations for each combination of population sizes, and calculated the proportion of simulated F_{ST} values that are greater than the observed F_{ST}. F_{ST} calculations were carried out by Arlequin version 3.5.2 (Excoffier and Lischer 2010). In the present study all the fastsimcoal2 simulations were run by Dr. Ayşegül Birand.

In the present study for the fastsimcoal2 simulation, the mtDNA HVRI sequences of 7 Çemialo Sırtı individuals and 10 Neareastern Neolithic samples from Syria were used. First, all of the sequences were aligned using ClustalW multiple sequences alignment software. Then, the overlapping region of the mtDNA HVRI sequences was detected as 240 bp in length (between 16126-16366 np on the CRS (Anderson et al., 1981)) based on the sequences of 17 ancient individuals. The archaeological date of the 10 Neolithic and 7 Çemialo Sırtı samples were assumed as 7300 BC and 500 BC, respectively. Furthermore effective population sizes (Ne) were assumed as a range from 250-5000 for the near eastern Neolithic population, while Ne was within the range of 5000-50000 for the Çemialo Sırtı population in the simulations under three growth models (exponential growth, instantaneous growth and no growth).

2.2.9 Calculations of the Nucleotide Misincorporations as per Nucleotide Transition Rates

The nucleotide misincorporations which are the nucleotide changes in the DNA molecule after the death, present as nucleotide transitions, transversions or indels on the sequences. The misincorporated nucleotides were identified by observing mismatches between the repeated sequences of the same fragments for each Çemialo Sırtı sample. Misincorporations which were observed among Çemialo Sırtı samples, are given in Section 3.2 (Sequence Alignment of the Fragments) of the Results Chapter in the present study.

For the seven Çemialo Sırtı samples, the rates of transitions per nucleotide sequenced per sample or per fragment were calculated as: The number of transitions observed for one sample divided by the number of nucleotides sequenced for that particular sample or for each fragment.

2.2.10 Testing for the Possible Contamination by the Research Team Members

The buccal swab samples of aDNA laboratory researchers and archaeological team members who worked in the field and handled the samples, were obtained with informed consents. Afterwards mtDNA HVI-HVII regions of the samples were sequenced and their haplogroups were determined. The sequences of researchers were compared by aligning with each fragment sequence of each ancient sample to detect possible contamination from modern human DNA. Moreover to avoid the contamination from modern samples to ancient ones, sampling, extraction, PCR amplification and purification steps of the buccal swab samples were carried out in the modern DNA laboratory in a period when not working regularly at the aDNA laboratory to avoid contamination from buccal swab samples.

2.2.10.1 Buccal swab samples

Buccal swab samples were obtained from the aDNA laboratory researchers (Inci Togan, Nihan Dilşad Dağtaş, Eren Yüncü, Füsun Özer and Reyhan Yaka) and archaeological team members (Yasemin Yılmaz, Sidar Gündüzalp, Aliye Gündüzalp and Esra Çiftçi) using a sterile single-use buccal swab brush. Buccal swabs were immediately placed into 1.5 mL eppendorf tubes containing 300 mL Lysis Buffer of the 5 Prime DNA Extraction Kit.

2.2.10.2 DNA Extraction of the Buccal Swab Samples

DNA from the buccal swab samples were extracted using 5 Prime (Archive Pure DNA Tissue Kit) Extraction Kit following the manufacturer's instructions. For the extraction step two extraction blanks, one at the beginning and the other one at the end of the sample set, were used as negative controls. The same order of samples was followed in every steps such as DNA extraction and PCR amplification.

2.2.10.3 PCR Amplification of the Buccal Swab Samples

DNA of the buccal swab samples were not expected to be fragmented as aDNA. Thus, the PCR amplifications of the samples were performed as modern DNA amplification procedure using two primer pairs to amplify HVI and HVII regions, each of the regions as one complete sequence instead of several small fragments which was in the case of ancient samples. So the sequences of HVI-HVII regions were not amplified fragment by fragment.

For these buccal swab samples, PCR amplification of human mtDNA HVI-HVII regions which are the same regions with the sequences of ancient samples were carried out. The HVI-HVII regions were amplified as 2 sequences (sequences of HVI and HVII regions) using the 2 different primer pairs (4 different primers) obtained from Ottoni et al.'s (2011) study. To amplify the 359 bp part of HVI region

(between 16008–16366 np on the CRS), forward primer of fragment A and reverse primer of fragment E were used; while the forward primer of fragment F and reverse primer of fragment G were used to amplify the 217 bp part of HVII region (between np 49–265 on the CRS). All of the primer sequences used to amplify buccal swab samples with their mtDNA regions, length of the amplified regions and fragments are given in Table 2.10.

Table 2.10 The mtDNA HVI-HVII regions, fragments of the each primer sequence, the length of sequences (bp) including the forward and reverse sequences of the primers (F16008- R16366 for HVRI, F48-R285 for HVRII, Ottoni et al., 2011) used for the amplification of the buccal swab samples.

mtDNA regions	**Primers**	**Primer sequences**	**Length of sequences (bp)**	**Fragments**
HVRI	F16008	5′- CCCAAAGCTAAGATTCTAAT -3′	359	A
HVRI	R16366	5′- TGAGGGGGGTCATCCATG -3′	359	E
HVRII	F48	5′- CTCACGGGAGCTCTCCATGC -3′	217	F
HVRII	R285	5′- GTTATGATGTCTGTGTGGAA -3′	217	G

PCR amplifications of the mtDNA HVI-HVII regions of the buccal swab samples were performed in 50 µL (5 µL template DNA with 45 µL PCR mix) reaction volumes. The PCR blanks (one at the beginning and another one at the end of the sample set) were also included as negative controls. The details of the PCR reagents are given in Table 2.11.

Table 2.11 PCR reagents with their volumes for the amplification of the buccal swab samples.

Nuclease free H_2O (Molecular biology grade)	32.0 µL
Buffer (10x,UV) (Thermo Scientific)	5.0 µL
$MgCl_2$ (25 mM, UV) (Thermo Scientific)	5.0 µL
dNTPs mix (Bioline) (10 mM/each)	1.0 µL
Primers (FWD+REV) (10 µM/each)	1.0 µL
BSA (UV) (50 mg/mL)	0.5 µL
Taq polymerase (5 U/µL)	0.5 µL
DNA template	5.0 µL

Buccal swab DNA PCR conditions were slightly modified from the PCR conditions of Ottoni et al.'s (2011) study. PCR conditions for the amplification of buccal swab samples are given in Table 2.12. Different thermocycler was used for the amplifications of buccal swab samples to prevent cross-contamination with aDNA samples.

Table 2.12 PCR conditions of the DNA from buccal swab samples

Initial denaturation	94 °C	10 min	
Denaturation	94 °C	45 s	35 cycles
Primer Annealing	56 °C	1 min	
Extension	72 °C	1 min	
Final extension	72 °C	5 min	
Base temperature	4 °C	∞	

2.2.10.4 Gel Electrophoresis of the Buccal Swab Samples

PCR products were visualized on %1 agarose gel prepared with 0.5x TBE buffer and 4 µL Ethidium bromide. Five µL of each PCR product was mixed with 3 µL 1x Blue loading dye and then loaded into the each well of the agarose gel. The gel electrophoresis was run at 110 V approximately for 50 min.PCR products were visualized under the UV light using Vilber Lourmat CN-3000WL. Moreover, the 50

bp DNA ladder (Thermo Fisher Scientific, Waltham, MA, USA) was used as the reference to check if the sizes of the PCR products were 359 bp for HVI and 217 bp for HVII regions. The DNA ladder is a reference DNA which is composed of the DNA fragments with different known sizes (50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000 bps).

2.2.10.5 Purification and Sequencing of the Buccal Swab Samples

After checking the PCR products on the gel, successfully amplified PCR products were purified using ExoSap-IT (USB Affimetrix) enzyme following the manufacturer's protocol to remove any excess dNTPs, salts, primers and enzymes. The procedure is as explained: 2μL of ExoSap-IT was mixed with 5 μL of PCR products. The thermocycling of this purification step was performed using a different PCR thermocycler than the one used for ancient samples. The cycle conditions of ExoSap-IT reaction are given in Table 2.13.

Table 2.13 Cycle conditions of the ExoSap-IT purification method.

37 °C	15 min
80 °C	15 min
4 °C	∞

2.2.10.6 Molecular Analyses of the Buccal Swab Samples

2.2.10.6.1 Alignment of the sequences

The forward and reverse sequences of the human mtDNA HVI-HVII regions were aligned with CRS (Anderson et al., 1981) using the BioEdit 5.0.9 software (http://mbio.ncsu.edu/bioedit/bioedit.html). Both of the forward and reverse sequences of one region (HVI or HVII region) were used for the alignment of sequences to obtain one correct sequences for each of the HVI and HVII regions.

When the beginning of the forward sequence did not have successful chromatogram peaks or sequence, the beginning of the reverse sequence was used for the assemblage of aligned sequence or vice versa for the end of the reverse sequence. Once more, the forward and reverse sequences were alternately used to read the sequence of each region correctly. The sequence of each region (HVI-HVII regions) was aligned with reference sequence, CRS. Furthermore, using BLAST tool of GenBank database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) the sequences were checked if they best match to humans'.

2.2.10.6.2 Haplogroup Determination of the Buccal Swab Samples

The mtDNA Haplogroups of the aDNA laboratory researchers and archaeological team members were determined with the same method which was used for the ancient samples. The haplogroup determination method was given in Section 2.2.7.2 (Haplogroup Determination) of this chapter.

CHAPTER 3

RESULTS

For the present study the samples were collected from the excavation site by Dr. Yasemin Yılmaz, Sidar Gündüzalp, Aliye Gündüzalp and Esra Çiftçi during 2013 and 2014 archaeological seasons. The first set of samples was consisted of 7 teeth samples of 5 individuals unearthed in 2013 excavation season. The second set was consisted of 11 samples of which 9 of them were collected in 2013 season and the last 2 in 2014 season. All of the 16 samples were brought to aDNA laboratory, stored and aDNA extractions were carried out here. Results of aDNA extractions, PCR amplifications and mtDNA HVRI-HVRII sequences of all the samples from the first sample set and four samples from the second sample set, in total 9, were presented in this chapter of the present study.

3.1 Optimization of the Extraction and PCR Amplification Steps

In the present study aDNA extraction and amplification protocols established by Ottoni et al. (2011) should be modified (reoptimized) in our laboratory to obtain efficient results. The modified conditions in relation to the extraction of aDNA process are summarized in Table3.1. These modifications are in relation to some parts of the aDNA extraction and amplification procedure given in sections 2.2.1.3 (aDNA Extraction) and 2.2.1.4 (aDNA Amplification) of Materials and Methods Chapter.

Table 3.1 The modified conditions in relation to the extraction of aDNA.

Modified steps	In the Ottoni et al.'s (2011) procedure	In the modified procedure
Speed of the first centrifugation after the incubation step	12,800 rpm	13,000 rpm
Speed of Elution Buffer centrifugation	10,000 rpm	11,800 rpm
Adding of Elution Buffer	52 uL + 52 uL (Two repeats)	55 uL + 55 uL + 26uL (Three repeats)

The conditions of the optimization process for the PCR amplification protocol were determined after two types of PCR amplification experiments.

First PCR amplification was carried out by applying Gradient PCR on modern DNA samples collected from the researcher and other colleagues whose DNA might dominate the aDNA samples. Gradient PCR is usually performed to optimize a PCR, to figure out what annealing temperatures work best. For this PCR, the same PCR reagents of the aDNA amplification protocol were used, but, instead of aDNA, the DNA from modern samples were added. The master mix of the gradient PCR was prepared in the aDNA laboratory following the same procedure employed for aDNA amplification. Then, the modern DNA samples were added in the laboratory which is used for modern samples. Gradient PCR was carried out to see the optimum annealing temperature for the primer pairs of fragment C and G which did not exhibit clearly visible bands in comparison to other five fragments. The temperature range of the gradient PCR was chosen between 53°C and 57°C for the annealing step of amplification cycle. Results of the Gradient PCR amplification is shown on the 3% agarose gel in Figure 3.1.

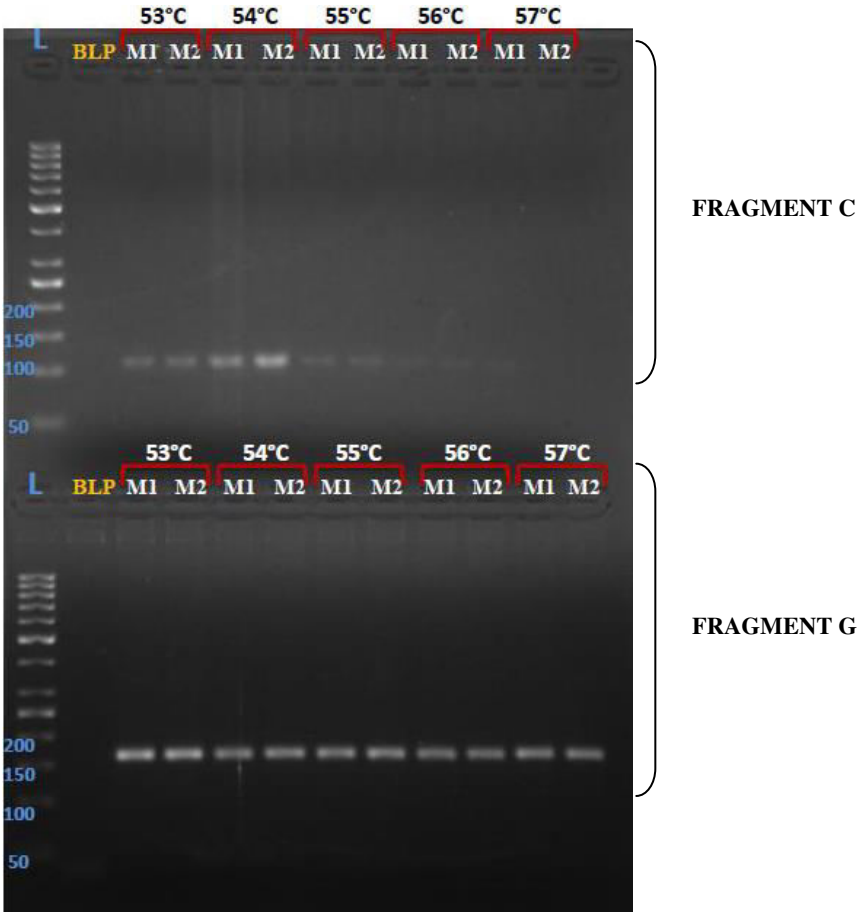


Figure 3.1 The gradient PCR products for the fragments C and G of two different modern DNA samples (M1and M2) at 5 different temperatures (53°C, 54°C, 55°C, 56°C, 57°C) on the 3% agarose gel. L: DNA size ladder between 50-1000 bp, BLP: PCR blank, Samples: M1 and M2.

As it can be seen from Figure 3.1, for the fragment C, bands at 54°C were the brightest. Whereas for the fragment G bands were bright enough at all temperatures, but the brightest bands were at 53°C. Although, modern DNA samples were used still the bands of the fragment C were not as bright as the bands of fragment G, possibly, due to the shorter length of the C fragment sequences. Both of the PCR blanks (BLP, PCR negative controls), in Figure 3.1 were clean indicating that there was no contamination , such as cross contamination, while performing the Gradient PCR. The second PCR amplification was carried out for each of the fragments of mtDNA HVI and HVII regions of the same modern DNA samples used in Figure 3.1. For the five fragments of HVRI, 54°C annealing temperature which was the

optimum temperature for fragment C, was tested to check if it was also suitable for the other four fragments of HVRI as the annealing temperature during the PCR. Likewise, for the HVRII fragment F, 53°C (the optimum temperature for fragment G) was tested as the annealing temperature for fragment G. Results of these PCR amplifications were shown on the 3% agarose gel in Figure 3.2.

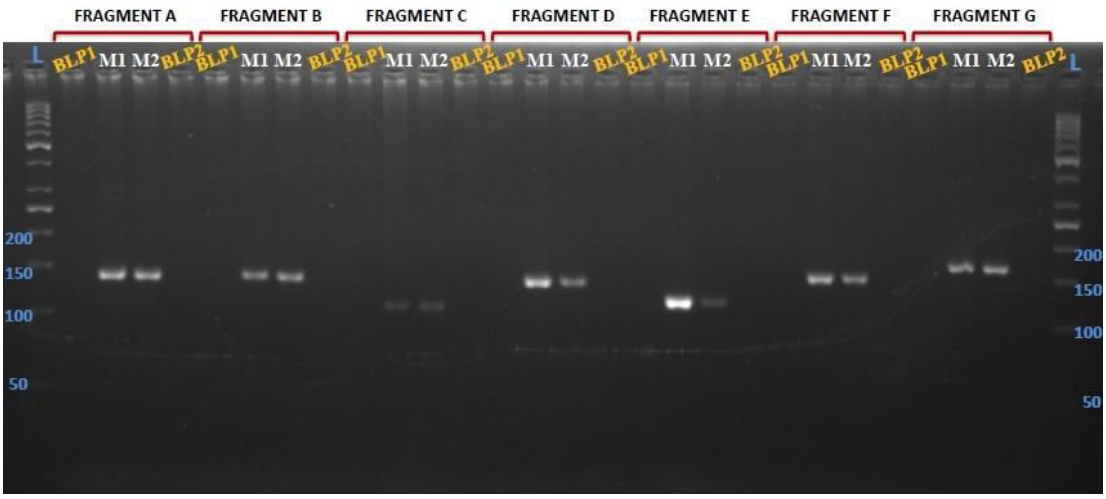


Figure 3.2 The 7 overlapping short fragments of the mtDNA HVI-HVII regions of two modern (M1 and M2) DNA samples with their two PCR blanks on 3% agarose gel. L: DNA size ladder between 50-1000 bp, BLP1: PCR blank1, BLP2: PCR blank2, Samples: M1 and M2.

In Figure 3.2 it can be seen that all of the fragments can be visualized on the agarose gel after their amplification by PCR. The lengths of the bands A, B, C, D, E, F and G are approximately equal to 140 bp, 139 bp, 109 bp, 133 bp, 118 bp, 147 bp and 166 bp, respectively, and these lengths are in good fit with the expected lengths of the PCR amplification products.

Results, in relation to the optimization of PCR amplifications are summarized in Table 3.2.

Table 3.2 The modified conditions for the PCR amplification of the 7 fragments of HVI and HVII regions of mtDNA.

Modified steps	**In the Ottoni et al.'s (2011) procedure**	**In the modified procedure**	**Regions (HVRI-HVRII) / Fragments (A-B-C-D-E-F-G)**
Number of cycles	45	60	HVRI-HVRII / A-B-C-D-E-F-G
Primer Annealing temperature (°C)	56 °C	54°C	HVRI / C
	56 °C	53 °C	HVRII / G
Primer annealing time	1 min	1.15 min	HVRI-HVRII / A-B-C-D-E-F-G

Then, the modified PCR amplification conditions were also tested on the fragments of aDNA samples. Observing that the modified conditions were also revealing better results for aDNA, they were employed for all of the samples in the present study.

3.2. aDNA Extraction and Amplification

aDNA extractions from each of the samples were carried out at least two times and PCR amplifications from each extractions were performed at least three times to obtain at least two repeated sequences of each fragment. The latter was necessary to check the reproducibility and the accuracy of the sequences. Numbers of aDNA extractions and PCR amplifications used for each fragment of the each sample including the identification number of samples from each individual are presented in Table 3.3.

Table 3.3 Çemialo Sırtı samples, their identification numbers, types of samples, number of times aDNA were extracted, number of PCRs performed, number of successful PCRs obtained and number of successful sequences obtained for each fragment of the mtDNA HVI and HVII regions. I and II indicates the first and second samples of the same individual, respectively.

Sample / Number of samples	Sample Type	Number of Extractions	Number of PCRs	Successful PCRs	Number of Successful Sequences for the Fragments							Total
					A	B	C	D	E	F	G	
SK24 / I	Tooth	3	12	4	1	1	1	2	1	1	1	8
SK24 / II	Tooth	2	6	4	1	1	1	1	1	2	1	8
SK26 / I	Tooth	3	12	7	1	2	1	1	1	1	1	8
SK26 / II	Tooth	2	6	4	2	2	1	1	1	1	1	9
SK28 / I	Tooth	3	12	8	6	4	2	5	2	2	2	23
SK29 / I	Tooth	3	12	8	1	1	0	3	1	2	1	9
SK30/ I	Tooth	3	12	8	2	0	1	1	2	2	0	8
SK7 / I	Tooth	2	6	4	2	2	2	2	3	2	2	15
SK17 / I	Tooth	2	6	5	3	3	3	3	3	2	2	19
SK21 / I	Phalanx	2	6	5	2	3	3	2	3	2	2	17
SK35 / I	Tooth	2	6	6	4	2	4	3	4	2	3	22

After the amplification step, prior to sequencing, all of the PCR products were checked on 3% agarose gel to visualize the successfully amplified sequences. If their lengths were roughly 140 (A), 139 (B), 109 (C), 133 (D), 118 (E), 147 (F) and 166 (G) bp, respectively, then they were sent to sequencing. An example of the agarose gel displaying successfully amplified fragments of the HVRI and HVRII regions of the two samples (SK17, SK35) and three types of blanks (5 blanks per fragment when sample sizes are 2) used at each step are shown in Figure 3.3 and Figure 3.4, respectively.

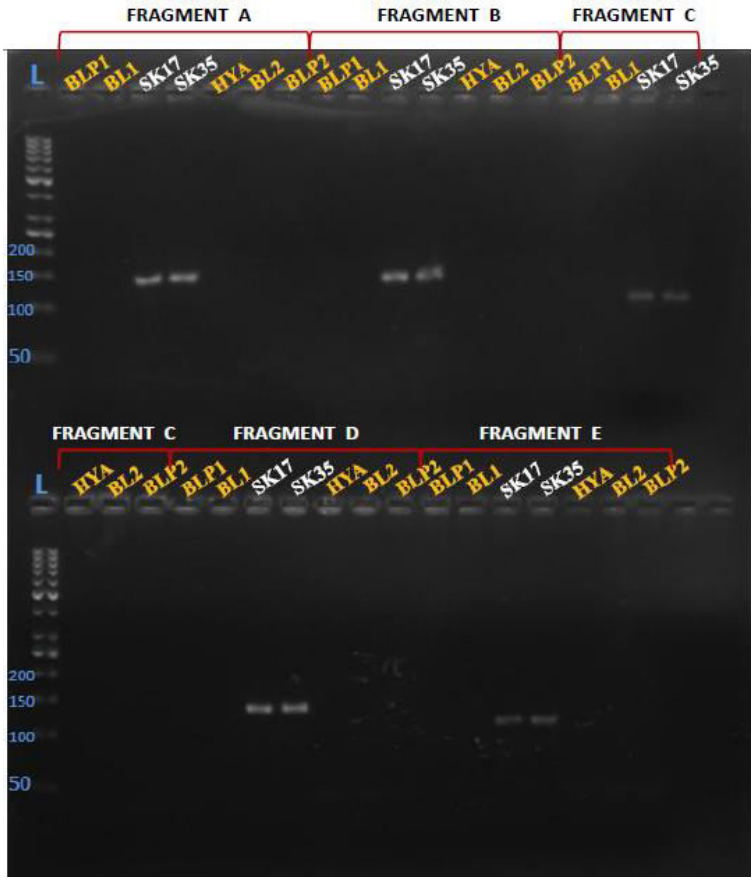


Figure 3.3 For the two ancient samples (SK17, SK35) their successfully amplified 5 overlapping aDNA fragments from the HVRI region. On the agarose gel; grinding, extraction and PCR blanks are also shown. For the fragments A, B, C, D and E sizes are: 140 bp, 139 bp, 109 bp, 133 bp and 118 bp, respectively. For each fragment 7 wells were used as shown in figure. L: DNA size ladder between 50-1000 bp, BLP1: PCR blank1, BLP2: PCR blank2, BL1: Extraction blank1, BL2: Extraction blank2, HYA: Grinding blank using synthetic Hydroxyapatite powder.

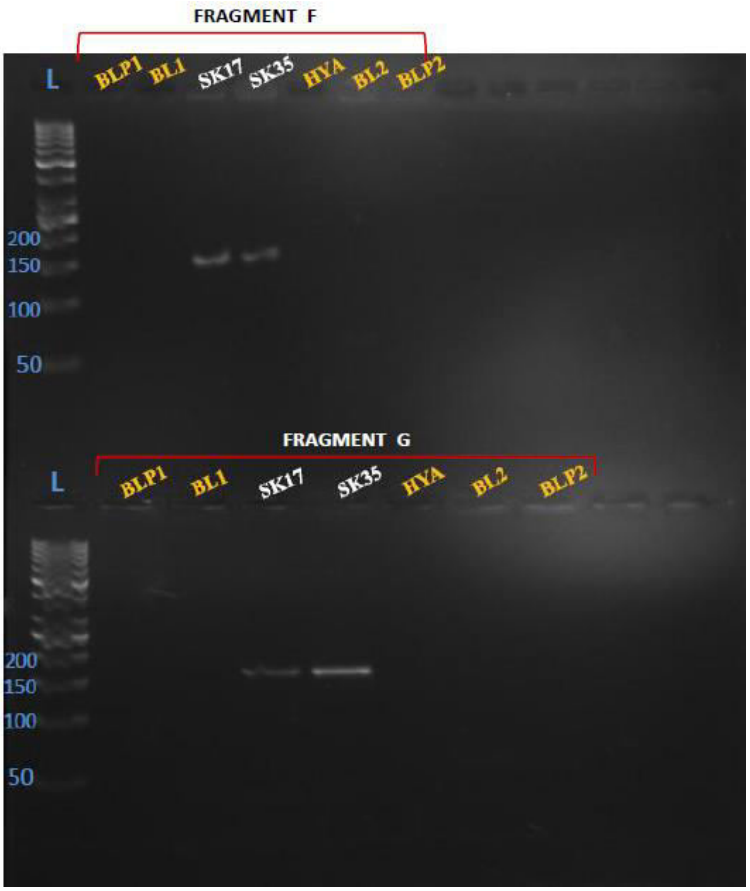


Figure 3.4 For the two ancient samples (SK17, SK35) their successfully amplified 2 overlapping aDNA fragments from the HVII region. On the agarose gel; grinding, extraction and PCR blanks are also shown. For the fragments F and G sizes are: 147 bp and 166 bp, respectively. For each fragment 7 wells were used as shown in figure. L: DNA size ladder between 50-1000 bp, BLP1: PCR blank1, BLP2: PCR blank2, BL1: Extraction blank1, BL2: Extraction blank2, HYA: Grinding blank using synthetic Hydroxyapatite powder.

Before going any further in the aDNA analysis it must be emphasized that modern mtDNA HVI and HVII regions of Reyhan Yaka, Nihan Dilşad Dağtaş, Eren Yüncü, Füsun Özer, Yasemin Yılmaz, Sidar Gündüzalp, Aliye Gündüzalp and Esra Çiftçi, who could have contaminated the ancient samples accidently, were sequenced. These sequences were given in Appendix A. Furthermore, their mutation motif profiles (one of the profiles as presented in Section 2.2.7.2 (Haplogroup Determination) in Materials and Methods Chapter with respect to CRS), thus, their mtDNA haplogroups are also presented in Appendix A.

As it can be seen from Figure 3.1, the bands of fragments C and E were pale which are the two shortest fragments of the HVRI. Moreover all of the blanks were clean which indicated that there was no contamination during the grinding, extraction and PCR amplification steps. After the amplification of the fragments, they were purified by an enzymatic reaction using the ExoSAP-IT and they were sent out for the sequencing.

3.3 Sequence Alignment of the Fragments

Sequencing was performed using ABI PRISM® 3100 Genetic Analyzer and sequences of the fragments were visualized with the BioEdit 5.0.9 (http://mbio.ncsu.edu/bioedit/bioedit.html) alignment program. Each fragment was sequenced in both forward and reverse directions. When the chromatogram peaks of the forward sequence had the low quality, the reverse sequence was used to find out the missing parts of the forward sequence. When both sequences were available, they were both evaluated for the accuracy of the target sequence. Then, the sequences of each fragment of the mtDNA HVRI and HVRII for the same individual were combined and aligned together with the corresponding sequences of CRS (Cambridge Reference Sequence, Anderson et al., 1981). Screenshots for the fragment sequences of sample SK28 after being aligned by the BioEdit 5.0.9 alignment program are shown in Figures 3.5- 3.11. In the alignment program each nucleotide has different colour (A:Green, T:Red, C:Blue, G:Black).

For the two individuals, SK29 and SK30 successful amplifications for all of the fragments could not be obtained (Table 3.1), therefore they were excluded from the further analysis. SK29 was a small tooth sample. Thus the sufficient amount of aDNA could not be obtained from the sample of individual SK29. Whereas the tooth sample of SK30 belonged to a young child and similarly, sufficient amount of aDNA could not be obtained from this sample.

In Figure 3.5 for individual with the identification number SK28, aligned sequences of its fragment A are shown.

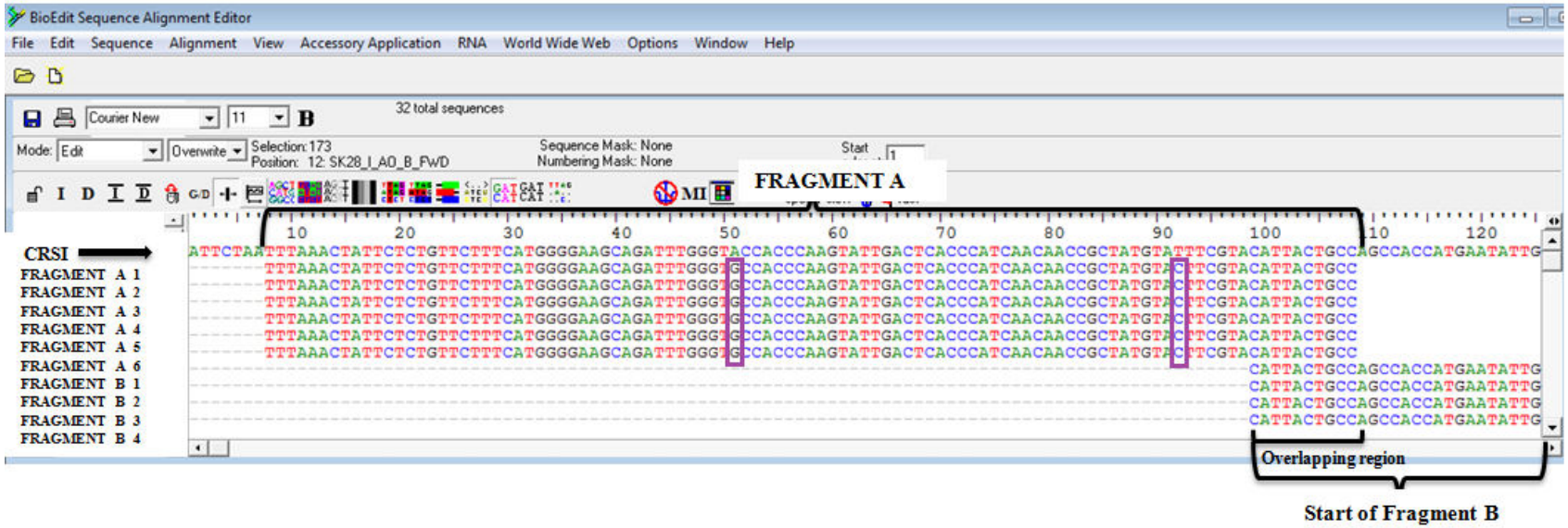


Figure 3.5 A screenshot from Bioedit 5.0.9 program shows the sequences (n=6) of fragment A which is the first fragment of HVI region and 140 bp in length. They are aligned with CRSI. A short part of fragment A which is 10 bp in length is overlapping with fragment B. Six sequences of fragment A belonged to two different extractions from the one tooth sample of individual SK28 and three independent PCRs from each extraction were made. The purple boxes indicate mutation motifs that are from A to G at 51 np (nucleotide position) and T to C at 92 np in reference to CRSI. These np are actually 16051st and 16092nd positions in the given CRS when the sequence of whole mtDNA is used.

The mutation motifs (indicated by the purple boxes) observed at the 51 np and 92 np in reference to CRSI in Figure 3.5 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments A and B indicate that, with a high probability, these two sequences belong to the same individual.

For the rest of fragments all observed (Figures 1-49, Appendix B) mutation motifs will be indicated on the figures. These are, then, going to be used for haplogroup determinations of the individuals. Pairwise overlapping regions of the fragments will be examined to support the continuity of aDNA of the same individual. In the sequences, on some nucleotide positions, differences between the repeats of the same individuals were obtained and they are indicated with the brown boxes. These transitions and transversions are evaluated in Section 3.3 (Putative Misincorporations and Authenticity of aDNA Sequences) of the Results Chapter and discussed in the Chapter 4 of the thesis. Presence of these is also considered as the confirmation of the authenticity of aDNA.

In Figure 3.6 for the same individual (SK28) sequences obtained for fragment B were shown.

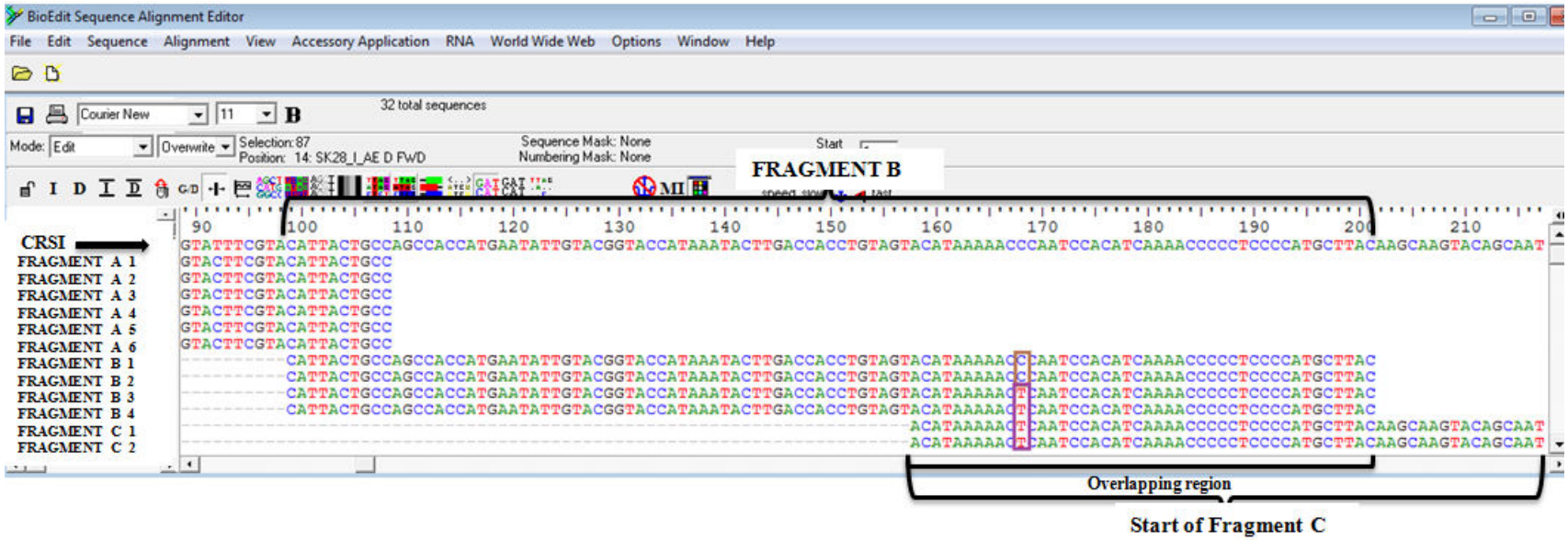


Figure 3.6 A screenshot from Bioedit 5.0.9 program shows the sequences (n=4) of fragment B which is the second fragment of HVI region and 139 bp in length. They are aligned with CRSI. A short part of fragment B which is 44 bp in length is overlapping with fragment C. Four sequences of fragment B belonged to two different extractions from the one tooth sample of individual SK28 and two independent PCRs from each extraction were made. The purple box indicates mutation motif that is from C to Tat168np in reference to CRSI. This np is actually 16168th position in the given CRS when the sequence of whole mtDNA is used. There were the "T" nucleotide on the 168th position of FRAGMENT B 3, FRAGMENT B 4 and also FRAGMENT C 1, FRAGMENT C 2, while it was "C" on the same position for FRAGMENT B 1 and FRAGMENT B 2 sequences (two B fragment sequences of the same sample) which was shown by brown box. Hence, the "C" on the FRAGMENT B 1 and FRAGMENT B 2 sequences, was considered as a nucleotide misincorporation (transition Type1). This nucleotide misincorporation is located at the overlapping region of fragment B and C.

The mutation motifs (indicated by the purple boxes) observed on the 168 np in reference to CRSI in Figure 3.6 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments B and C indicate that, with a high probability, these four sequences belong to the same individual.

For SK28, sequences for the rest of the fragments (C, D, E, F and G) are shown in Figures 3.7 and 3.11. For fragments F and G the reference sequences were CRSII.

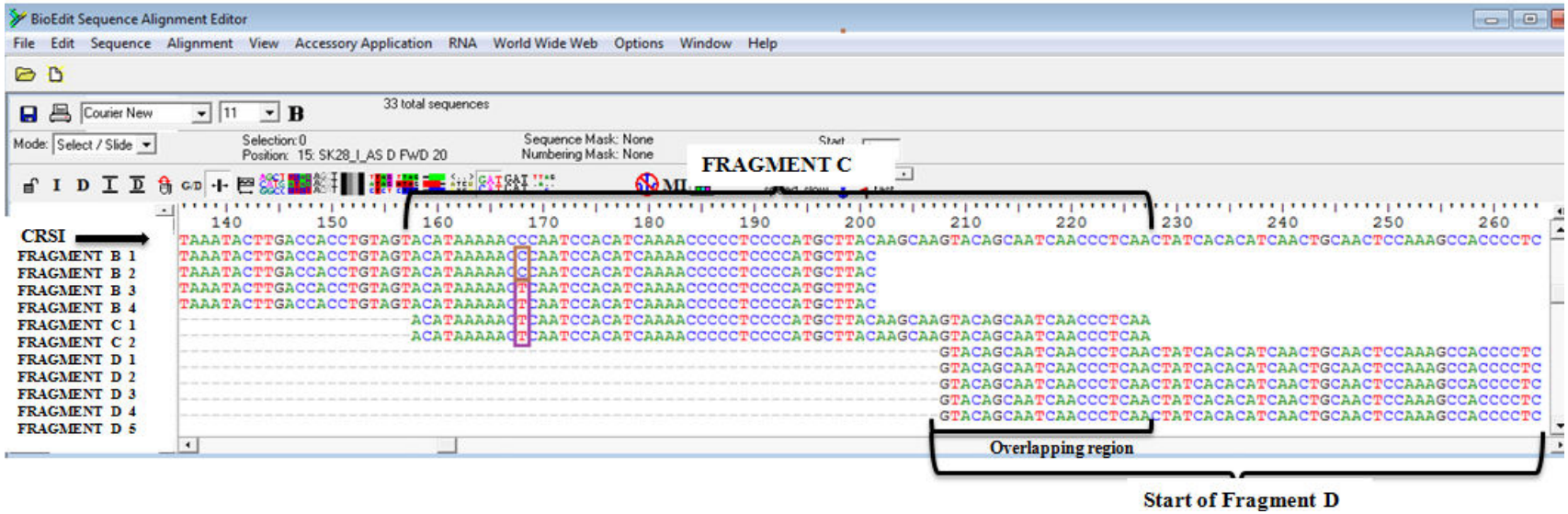


Figure 3.7 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment C which is the second fragment of HVI region and 109 bp in length. They are aligned with CRSI. A short part of fragment C which is 20 bp in length is overlapping with fragment D. Two sequences of fragment C belonged to two different extractions from the one tooth sample of individual SK28 and one independent PCR from each extraction was made. The purple box indicates mutation motif that is from C to T at168 np in reference to CRSI. This np is actually 16168th position in the given CRS when the sequence of whole mtDNA is used. This mutation motif is located at the overlapping part and both of the C fragment sequences had the same mutations at that position.

The mutation motifs (indicated by the purple boxes) observed on the 168 np in reference to CRSI in Figure 3.7 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments C and D indicate that, with a high probability, these two sequences belong to the same individual.

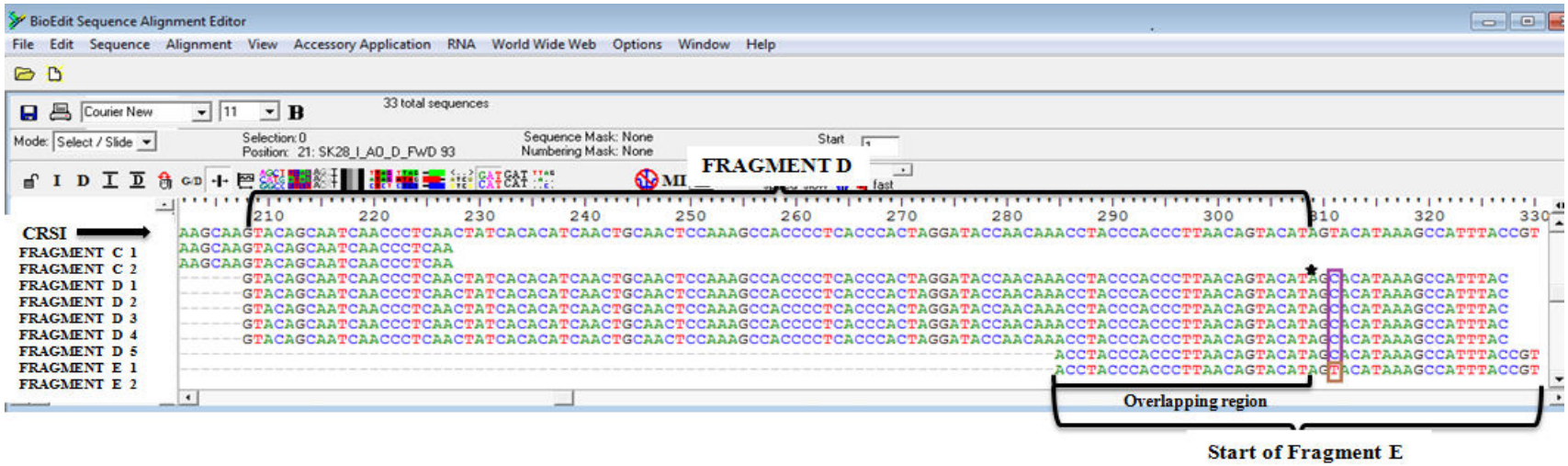


Figure 3.8 A screenshot from Bioedit5.0.9 program shows the sequences (n=5) of fragment D which is the fourth fragment of HVI region and 133 bp in length. They are aligned with CRSI. A short part of fragment D which is 24 bp in length is overlapping with fragment E. Five sequences of fragment D belonged to three different extractions from the one tooth sample of individual SK28 and two independent PCRs from each of the two extractions, one PCR from another extraction were made. The purple box indicates mutation motif that is from T to C at 311 np in reference to CRSI. This np is actually 16311th position in the given CRS when the sequence of whole mtDNA is used. Normally fragment D ends at the 309 np which was shown with the black star. However, in all of the repetitions of the fragment D sequences 19 additionally nucleotide could also be read. These, were on the reverse primer binding region of the sequences. Due to the limited number of repetitions of E fragment sequences, in the present study, the five D fragment sequences were examined up to the 327 np (3' end of D fragment was extended by 19 bp).The mutation motif at the overlapping part of the fragments D and E was observed.

The mutation motifs (indicated by the purple boxes) observed on the 311 np in reference to CRSI in Figure 3.8 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these five sequences belong to the same individual.

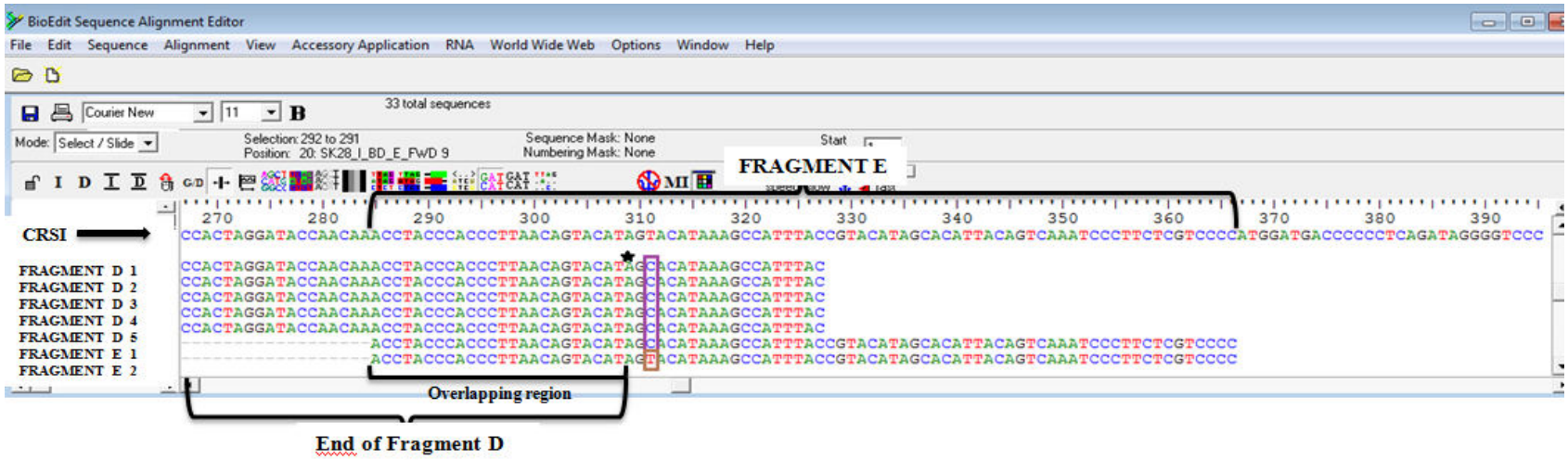


Figure 3.9 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment E which is the last fragment of HVI region and 118 bp in length. They are aligned with CRSI. A short part of fragment E which is 24 bp in length is overlapping with fragment D. Two sequences of fragment D belonged to two different extractions from the one tooth sample of individual SK28 and one independent PCR from each extraction. The purple box indicates mutation motif that is from T to C at 311 np in reference to CRSI. This np is actually 16311th position in the given CRS when the sequence of whole mtDNA is used. As it was explained in Figure 3.8 five sequences of fragment D were extended 19 bp more to check the accuracy of the sequences using the mutation motifs. One of the E fragment sequences had the same mutation at that position. There were the "C" nucleotide on the 311th position of all D fragment and FRAGMENT E 1 sequences while it was "T" on the same position for FRAGMENT E 2 sequence (second E fragment sequence of the same sample) which was shown by brown box. Hence, "T" on the FRAGMENT E 2 sequence was considered as a nucleotide misincorporation (transition Type2). This nucleotide misincorporation is located at the overlapping region of fragments D and E.

The mutation motifs (indicated by the purple boxes) observed on the 311 np in reference to CRSI in Figure 3.9 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments E and D indicate that, with a high probability, these two sequences belong to the same individual.

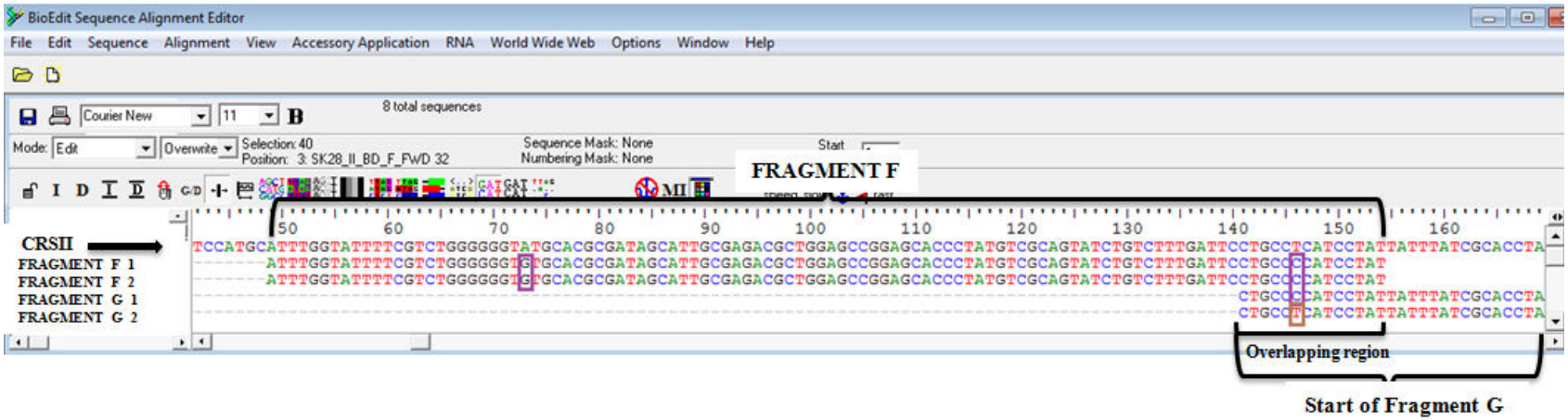


Figure 3.10 A screenshot from Bioedit5.0.9 program shows the sequences (n=2) of fragment F which is the first fragment of HVII region and 147 bp in length. They are aligned with CRSII. A short part of fragment F which is 14 bp in length is overlapping with fragment G. Two sequences of fragment F belonged to two different extractions from the one tooth sample of individual SK28 and one independent PCR from each extraction. The purple boxes indicate mutation motifs that are from A to G at 73 np and T to C at 146 np in reference to CRSII. These np are actually 73rd and146th positions in the given CRS when the sequence of whole mtDNA is used. T to C mutation motif is located at the overlapping part and both of the F fragment sequences had the same mutations at those two different positions.

The mutation motifs (indicated by the purple boxes) observed on the 73 np and 146 np in reference to CRSII in Figure 3.10 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

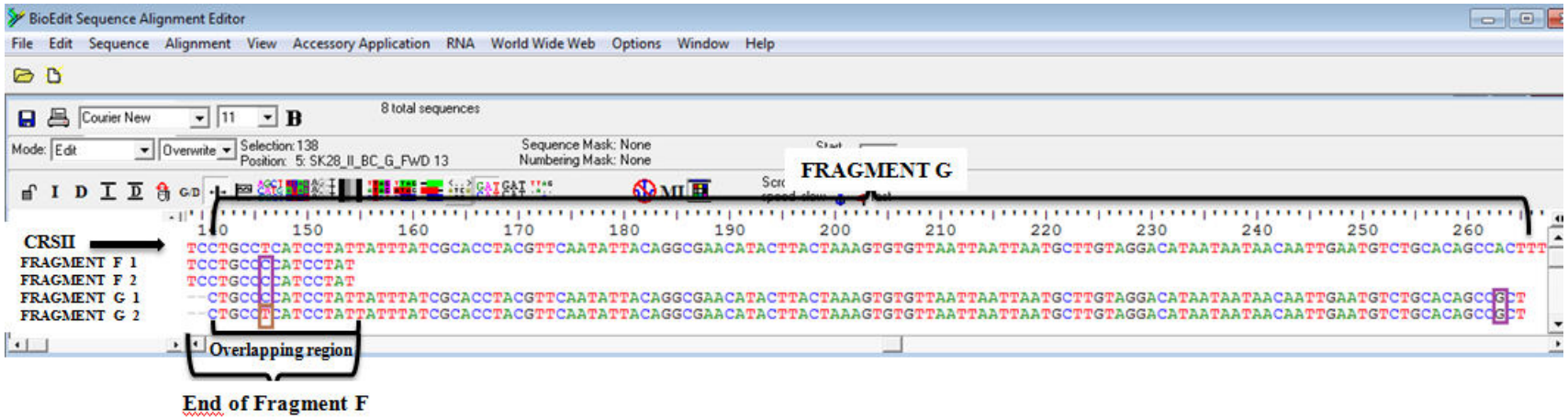


Figure 3.11 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment G which is the last fragment of HVII region and 166 bp in length. They are aligned with CRSII. A short part of fragment G which is 14 bp in length is overlapping with fragment F. Two sequences of fragment G belonged to two different extractions from the one tooth sample of individual SK28 and one independent PCR from each extraction. The purple boxes indicate mutation motifs that are from T to C at 146npand A to G at 263 np in reference to CRSII. These np are actually 146th and 263rd positions in the given CRS when the sequence of whole mtDNA is used. T to C mutation motif is located at the overlapping part and one of the G fragment sequences had the same mutation at that position while both of the G fragment sequences had the same A to G mutations at another position. There were the "C" nucleotide on the 146th position of FRAGMENT F 1, FRAGMENT F 2 and FRAGMENT G 1, while it was "T" on the same position for FRAGMENT G 2 sequence (second G fragment sequence of the same sample) which was shown by brown box. Hence, "T" on the FRAGMENT G 2 sequence was considered as a nucleotide misincorporation (transition Type2). This nucleotide misincorporation is located at the overlapping region of fragments F and G.

The mutation motifs (indicated by the purple boxes) observed on the 146 np and 263 np in reference to CRSII in Figure 3.11 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

Similar screenshots for samples identified as SK24, SK26, SK7, SK17, SK21, SK35, together with their mutation motifs are presented in Appendix B. For the sake of completeness SK28 was also given in Appendix B.

Furthermore, comparisons between ancient sequences and the modern DNA samples which were obtained from the archaeological team members (YY, SG, AG, AÇ) and aDNA laboratory researchers (RY, NDD, EY, FO) and presented in Appendix A, revealed that in individual SK26 in fragment D there was contamination from RY (in Appendix C, Figure 50).

All the sequences obtained from the same PCR were excluded from the analysis. For the rest of the aDNA sequences there were no contamination from the modern samples.

3.4 Putative Misincorporations and Authenticity of aDNA Fragment Sequences

In the present study putative misincorporations were determined by repeated (at least twice) and compared sequences of the fragments. The putative misincorporations with their transition types (Type1: A→G, T→C and Type2: C→T, G→A) based on the fragment types are given in Table 3.4.Transversion and indels were not observed for the samples presented in the present study.

The observed transitions on the basis of fragments were already indicated and explained on sample SK28in section 3.2 (Sequence Alignments of the Fragments) of the Results Chapter, in the present study. Total number of observed transitions and their distributions over the fragments for each sample are shown in Table 3.4 when UDG treatment was not applied and in Table 3.5 when UDG treatment was applied.

Table 3.4 The putative misincorporations with their transition types as found in each fragment sequences when UDG treatment was not applied. "→" **:** Changes from first nucleotide to second nucleotide. "/"**:** The ratio of number of transitions to number of repeated sequences.

Putative Misincorporations	**Transitions (Type1)**														**Transitions (Type2)**													
	A→G							**T→C**							**C→T**							**G→A**						
Fragments / **Samples**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**A**	**B**	**C**	**D**	**E**	**F**	**G**
SK24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1/2	-	1/2	-	-	-	-	-	-	-	-	-
SK26	-	1/4	-	-	-	-	-	-	-	1/2	-	-	-	-	1/3	-	-	-	-	-	-	-	-	-	-	-	-	-
SK28	-	-	-	-	-	-	-	-	1/2	-	-	-	-	-	-	-	-	-	1/2	-	1/2	-	-	-	-	-	-	-
SK7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1/3	-	-	-	-	-	-	-	-	-
SK17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1/3	-	-	-	-	-	-	-	-	-
SK21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SK35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 3.5 The putative misincorporations with their transition types as found in each fragment sequences when UDG treatment was applied. "→" **:** Changes from first nucleotide to second nucleotide. "/"**:** The ratio of number of transitions to number of repeated sequences.

Putative Misincorporations	**Transitions (Type1)**														**Transitions (Type2)**													
	A→G							**T→C**							**C→T**							**G→A**						
Fragments / **Samples**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**A**	**B**	**C**	**D**	**E**	**F**	**G**
SK24	-	-	-	-	-	1/3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SK26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SK28	-	-	-	-	-	-	-	-	1/2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SK7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SK17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SK21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SK35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1/3	-	-	-	-	-	-	-

From the Tables 3.4-3.5 it could be seen that,in the present study, there were misincorporations in almost all of the samples except in SK21;G to A transition (Type2) was not observed in any of the 7 Çemialo Sırtı samples, while another Type2 transition (from C to T) was highly frequent and it was especially observed in fragment E. Furthermore the rate of transition was 0.04% for transition Type 2 (C to T) without UDG treatment amplifications from the table 3.4, while it was 0.006% for the UDG treated sequences when all of the samples were considered (Table 3.5). However, the observed C to T transition with UDG treatment might be a damage from 5-methyl-cytosine (5mC) to thymine on the DNA sequences, whereas the non UDG treated sequences indicated cytosine deamination damage from the cytosine (C) to uracil (U). Other possibilities for this result will be discussed in next chapter (Discussion Chapter).The transition rates were calculated as the mean transitions per nucleotide sequenced per fragment and per sample for the studied seven Çemialo Sırtı samples. These calculations are given in Tables 3.6-3.7 and Table 3.8-3.9, respectively.

Table 3.6 Observed misincorporation rates (per nucleotide) with respect to each fragment when UDG treatment was not applied. "→" **:** Changes from first nucleotide to second nucleotide.

Putative Misincorporations	**Transitions (Type1)**		**Transitions (Type2)**	
	A→G	**T→C**	**C→T**	**G→A**
Fragments				
Fragment A	0	0	0.032	0
Fragment B	0.036	0.036	0	0
Fragment C	0	0.051	0.051	0
Fragment D	0	0	0	0
Fragment E	0	0	0.18	0
Fragment F	0	0	0	0
Fragment G	0	0	0.040	0

Table 3.7 Observed misincorporation rates (per nucleotide) with respect to each fragment when UDG treatment was applied. "→" **:** Changes from first nucleotide to second nucleotide.

Putative Misincorporations	**Transitions (Type1)**		**Transitions (Type2)**	
	A→G	**T→C**	**C→T**	**G→A**
Fragments				
Fragment A	0	0	0	0
Fragment B	0	0.036	0	0
Fragment C	0	0	0	0
Fragment D	0	0	0	0
Fragment E	0	0	0	0
Fragment F	0.045	0	0	0
Fragment G	0	0	0.040	0

When UDG treatment was not applied, the transition Type2 (C to T) was observed more than the UDG-treated amplifications.

Table 3.8 Observed misincorporation rates (per nucleotide) with respect to each sample when UDG treatment was not applied. "→" **:** Changes from first nucleotide to second nucleotide.

Putative Misincorporations	**Transitions (Type1)**		**Transitions (Type2)**	
	A→G	**T→C**	**C→T**	**G→A**
Samples				
SK24	0	0	0.092	0
SK26	0.043	0.043	0.043	0
SK28	0	0.033	0.066	0
SK7	0	0	0.049	0
SK17	0	0	0.039	0
SK21	0	0	0	0
SK35	0	0	0	0

Table 3.9 Observed misincorporation rates (per nucleotide) with respect to each sample when UDG treatment was applied. "→" **:** Changes from first nucleotide to second nucleotide.

Putative Misincorporations	**Transitions (Type1)**		**Transitions (Type2)**	
	A→G	**T→C**	**C→T**	**G→A**
Samples				
SK24	0.046	0	0	0
SK26	0	0	0	0
SK28	0	0.033	0	0
SK7	0	0	0	0
SK17	0	0	0	0
SK21	0	0	0	0
SK35	0	0	0.034	0

The highest rate of transition was 0.04% (the meanwhen UDG treatment was not applied) for transition Type2 (C to T), while it was 0.017% (the mean when UDG treatment was not applied)for transition Type1, when all of the samples were considered.

All of the Çemialo Sırtı sequences obtained in the present study are given in Appendix B.

3.5 Haplogroup Determination Based on mtDNA HVRI and HVRII

The mutation motifs on mtDNA HVRI and HVRII sequences which were exhibited by five and two overlapping fragments, respectively, were used to determine mtDNA haplogroups of seven Çemialo Sırtı individuals. If there were variations at the motif sites between the repeated sequences of an individual the most frequent variant is accepted as the "true" nucleotide. The successfully sequenced samples including their mtDNA haplogroups and mutation motifs were given in Table 3.10.

Table 3.10 Seven Çemialo Sırtı samples, their assigned mtDNA haplogroups and the mutation motifs which were used for their haplogroup determination in accordance with PhyloTreemt (http://www.phylotree.org/).

Sample	Haplogroup	HVRI Mutation Motifs	HVRII Mutation Motifs
SK24	H1z1	T16189C - T16311C	A263G
SK26	M1a1	T16093C - T16223C - T16249C - T16311C - T16359C	A73G - A263G
SK28	U2b1	A16051G - T16092C - C16168T - T16311C	A73G - T146C - A263G
SK7	H1a	A16162G - T16311C	A73G - A263G
SK17	HV	T16311C	A263G
SK21	R2	C16071T - C16278T	A73G - T152C - A263G
SK35	R6	C16266T - T16311C	T195C - A263G

3.6 mtDNA HVRI Haplogroups of Çemialo Sırtı Individuals' and Comparative Studies

In most of the previous studies, mtDNA haplogroups were determined on the basis of HVI region. Therefore, to be able to comparatively evaluate the results from Çemialo Sırtı with those from other populations, the information from HVI region is used. Thus, subhaplogroups H1z1 and H1a, M1a1, R2 and R6 which are the haplogroups of samples SK24, SK7, SK26, SK21, SK35, respectively, were considered as the representatives of main haplogroup H, M and R. Çemialo Sırtı is within the borders of modern Turkey, since Iran and Syria are the two close neighbours of Çemialo Sırtı, modern samples from Turkey and these countries were retrieved from databases. Since Anatolia was dominated by the Hellens in the years 330-270 BC, which is a period close to the dates of samples from Çemialo Sırtı, samples of modern Greece were also used for the comparative analysis together with modern data of Turkey and Iran. Furthermore, published data was available from Sagalassos excavation site which is dated 11th–13th century AD and from Neolithic sites the Tell Ramad, Tell Halula and Dja'de El Mughara of Northern Syria; thus these were

also used for the comparative purposes. To facilitate the visual comparisons pie charts showing the haplogroups under consideration (haplogroups observed in Çemialo Sırtı) were drawn for different populations and displayed in Figure 3.12.

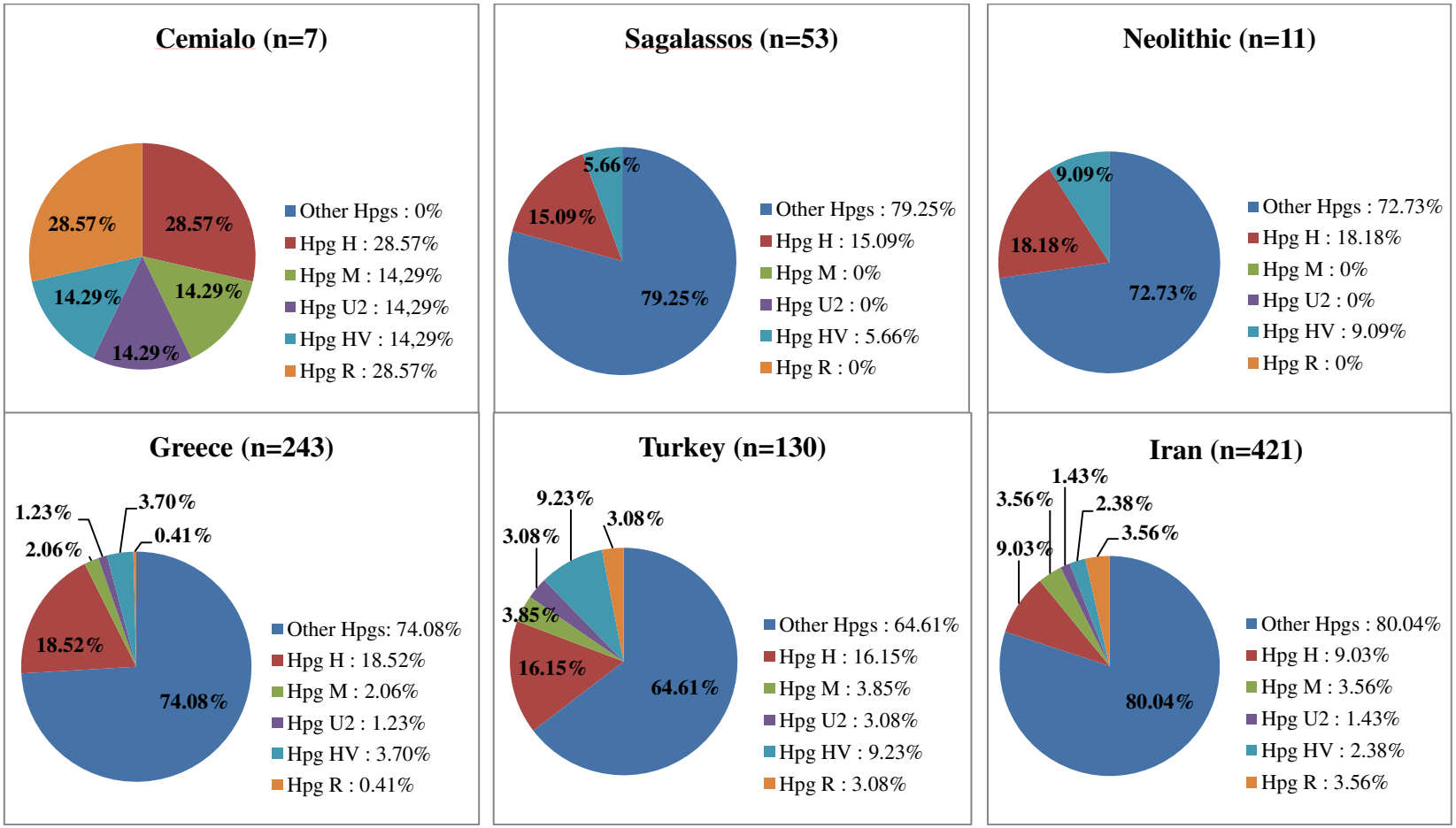


Figure 3.12 Pie charts of different populations (Greece, Iran, Turkey as modern and Sagalassos, Neolithic as ancient populations) depicting their mtDNA haplogroup compositions (on the basis of HVI region) with a special emphasis given to the haplogroups of Çemialo Sırtı. Hpg: Haplogroup, n: Total sample size.

The haplogroup H of SK24 and SK7 seems to be relatively high in all of the considered populations since the Neolithic times. Similarly haplogroup HV of SK17 might have been another frequently observed haplogroup of the region. However, especially R (SK21, SK35) also U2 (SK28) and M (SK26) seemed to be relatively rare haplogroups of the region since the ancient times. Pie charts indicate that Çemialo Sırtı represented by the seven individuals from the period 600-500 BC does not have a special affinity to any of the tested populations.

3.7 Principal Component Analysis (PCA) Based on mtDNA HVRI Haplogroups

The first Principle Component Analysis (PCA) was performed based on the frequencies data of 30 mtDNA haplogroups to visualize the haplogroup distribution of Çemialo Sırtı individuals in comparison to both ancient and modern populations by means of Principal Component Analysis. These 30 haplogroups were given in section 2.2.7.2 (Haplogroup Determination) of Materials and Methods Chapter in the present study. mtDNA HVI region data of 30 populations were used. The PCA was run in R v. 3.0.2. The modern populations were as follows: Albania (n=34), Bulgaria (n=668), Greece (n=243), Macedonia (n=160), Romania (n=91), Armenia (n=30), Azerbaijan (n=63), Cyprus (n=74) , Georgia (n=24), Iran (n=421), Iraq (n=156), Kazakh (n=43), Turkey (n=130), Turkmen (n=53), Syria (n=46), Uzbek (n=56). The ancient populations were Sagalassos (n=53), Neolithic (n=12) from Syria and other Neolithic populations from Southeastern/Central Europe; STA (n=43), LBKT (n=39), LBK (n=108), RSC (n=11), SCG (n=33), BAC (n=19), SMC (n=28), BEC (n=17), CWC (n=42), BBC (n=32), UC (n=88) and GURGY (n=31). Symbols of the populations are explained in the legend of Figure 3.13. For the ancient populations brief information was given in the first chapter of the present study.

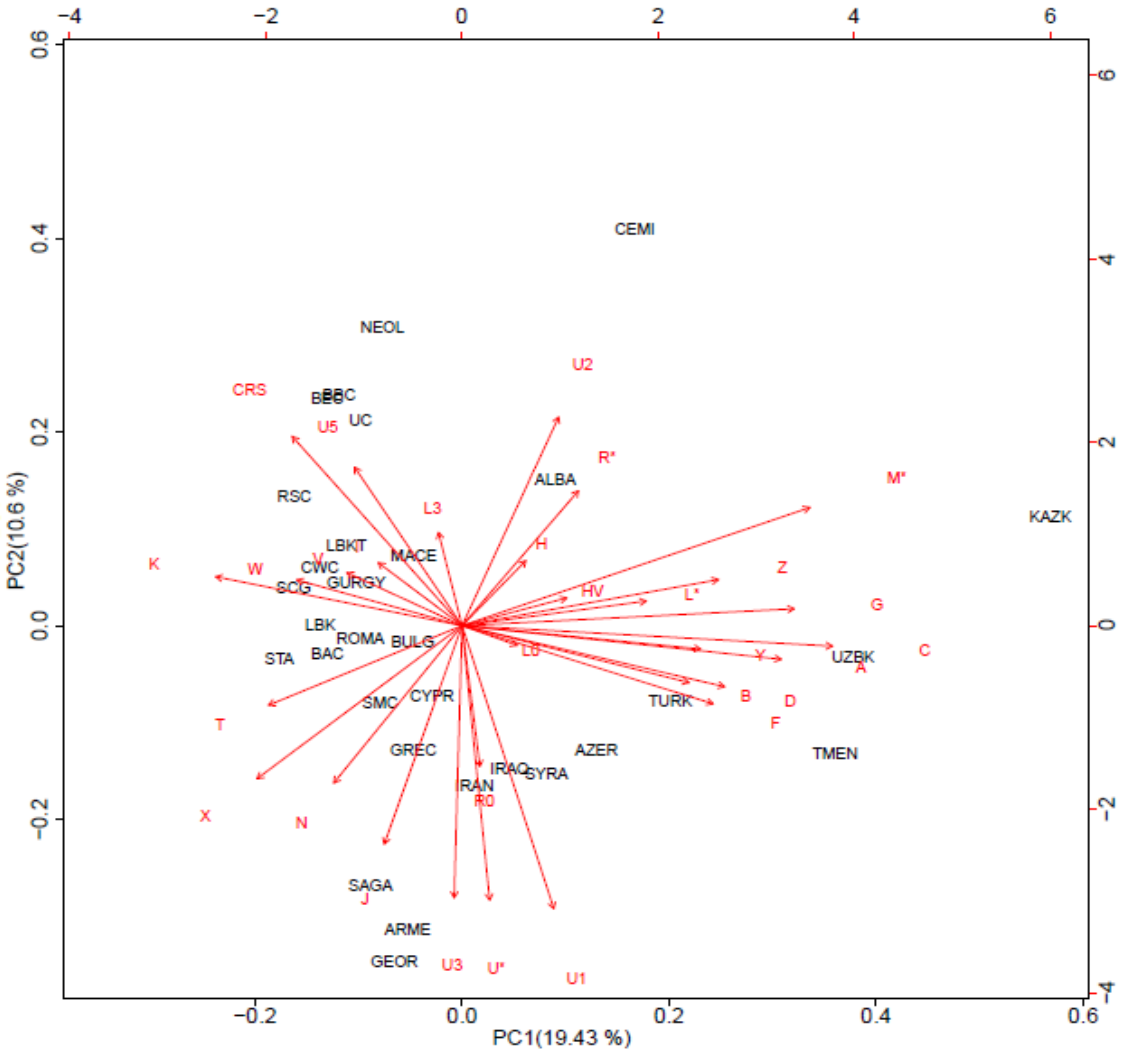


Figure 3.13 Two-dimensional plot of principle component analysis based on mtDNA HVRI haplogroup frequencies data. Alba: Albania, Bulg: Bulgaria, Grec: Greece, Mace: Macedonia, Roma: Romania, Kazk: Kazakh, Tmen: Turkmen, Turk: Turkey, Uzbk: Uzbek, Cypr: Cyprus, Iran: Iran, Iraq: Iraq, Geor: Georgia, Azer: Azerbaijan, Arme: Armenia, Saga: Sagalassos, Cemi: Cemialo, Neol: Neolithic, Syra: Syria, STA: STA (Early Neolithic Starčevo culture, 6000–5400 BC), LBKT: LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC), LBK: LBK (Linear Pottery culture, 5500-4775 cal BC), RSC: RSC (Rössen culture, 4625-4250 cal BC), SCG: SCG (Schöningen group, 4100-3950 cal BC), BAC: BAC (Baalberge culture, 3950-3400 cal BC), SMC: SMC (Salzmünde culture, 3400-3100/3025 cal BC), BEC: BEC (Bernburg culture, 3100-2650 cal BC), CWC: CWC (Corded Ware culture, 2800-2200/2050 cal BC), BBC: BBC (Bell Beaker culture, 2500-2200/2050 cal BC), UC: UC (Unetice culture, 2200-1550 cal BC), GURGY: Gurgy (5000-4000 cal. BC).

In Figure 3.13 the first principle component (PC1) explained 19.43% of the variation, while the PC2 covered 10.6%. Other PCs (principle components) was not shown in the Figure 3.13. The 30.03% of the total variation was accounted by the first two components together.

The red arrows indicate the 30 haplogroups and their distributions based on the haplogroup frequencies which were observed in each population. It can be observed from the Figure 3.13 that Çemialo Sırtı individuals differ from the Neolithic Syria population according to the R and U2 haplogroups which were not found in the Neolithic Syria population. On the other hand the Neolithic Syria population is genetically similar to Çemialo Sırtı due to the haplogroups H and HV, while it also shows genetic affinity to the other Neolithic populations from south eastern and central Europe due to the haplogroups K and CRS. Thus, the Neolithic population from Syria is located between the Çemialo Sırtı and other Neolithic populations from south eastern/central Europe according to the PCA which was carried out on the basis of haplogroups frequencies.

3.8 Principal Component Analysis (PCA) Based on mtDNA HVRI Sequences

In this section genetic relatedness between Çemialo Sırtı and 19 populations on the basis of 240bp long sequences between 16126-16366 np of the mtDNA HVRI sequence according to the CRS (Cambridge Reference Sequence, Anderson et al., 1981) were visualized by means of Principal Component Analysis (PCA). For the analysis GenAlEx (http://biology-assets.anu.edu.au/GenAlEx/Welcome.html) package program was used. The modern populations were as follows: Armenia (n=30), Azerbaijan (n=69), Cyprus (n=91), Georgia (n=83), Greece (n=320), Iran (n=50), Kazakh (n=44), Turkey (n=93), Turkmen (n=20), Syria (n=48) and Uzbek (n=20). The ancient populations were Sagalassos (n=53), Neolithic (n=10) from Syria and other Neolithic populations from south eastern/central European regions; LBKT (n=39), LBK (n=54), RSC (n=7), CWC (n=32), BBC (n=20) and UC (n=88). Symbols of the populations are explained in the legend of Figure 3.14.

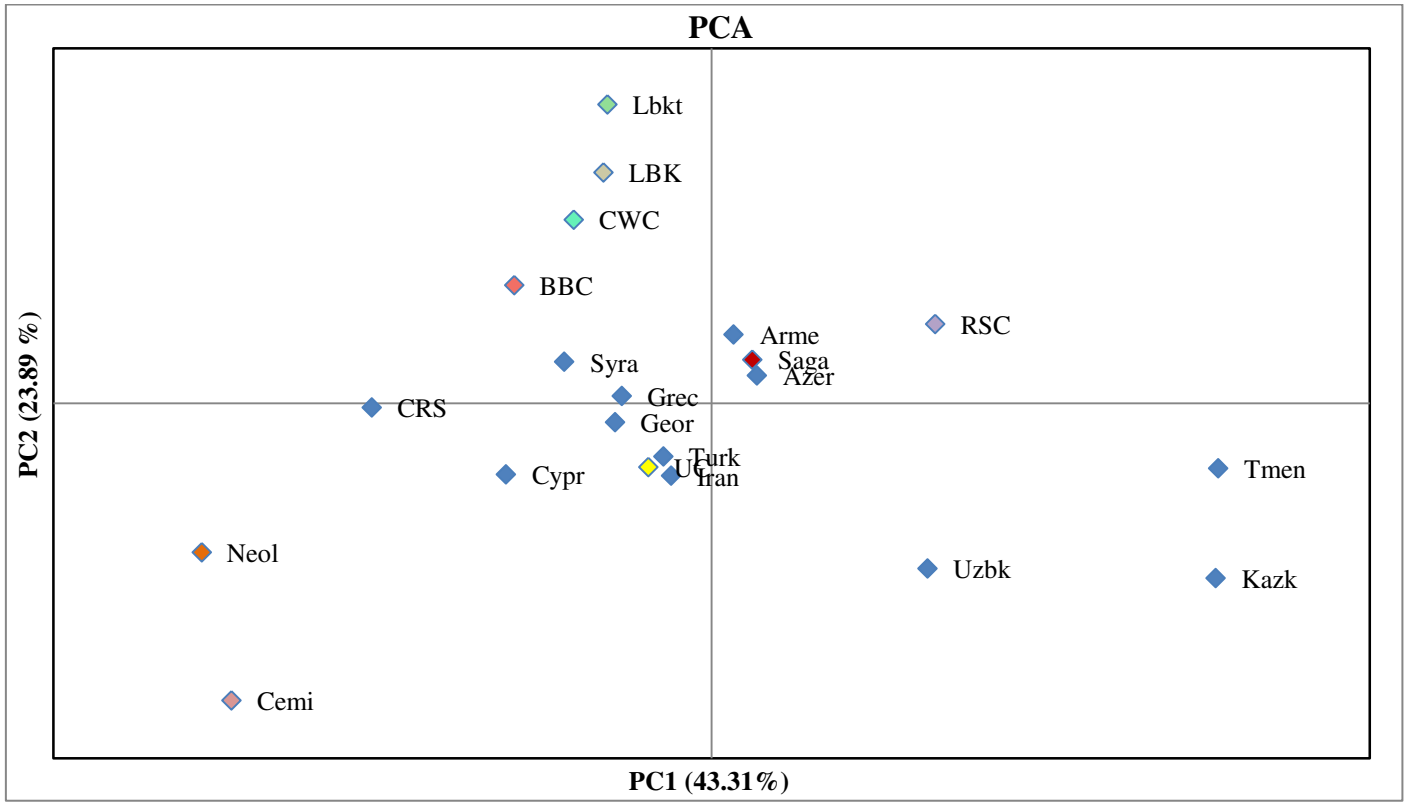


Figure 3.14 Two-dimensional plot of principle component analysis based on mtDNA HVRI sequences. The lighter and different colored dots indicate the different ancient populations. Grec: Greece, Kazk: Kazakh, Tmen: Turkmen, Turk: Turkey, Uzbk: Uzbek, Cypr: Cyprus, Iran: Iran, Iraq: Iraq, Geor: Georgia, Azer: Azerbaijan, Arme: Armenia, Saga: Sagalassos, Cemi: Çemialo Sırtı, Neol: Neareastern Neolithic population from Syria, Syra: Syria, Lbkt: LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC), LBK: LBK (Linear Pottery culture, 5500-4775 cal BC), RSC: RSC (Rössen culture, 4625-4250 cal BC), CWC: CWC (Corded Ware culture, 2800-2200/2050 cal BC), BBC: BBC (Bell Beaker culture, 2500-2200/2050 cal BC), UC: UC (Unetice culture, 2200-1550 cal BC), CRS: Cambridge Reference Sequence (Anderson et al., 1981).

The first principle component (PC1) explained 43.31% of the variation, while the PC2 covered 23.89%. The PC3 was not shown in the Figure 3.14 and it covered 17.88% of the variation. The 85.08% of the total variation was explained by the first three components together.

Furthermore to visualize genetic proximity of the Çemialo Sırtı population to the other populations under consideration, the pairwise mean genetic distances between Çemialo Sırtı and other populations were given in Table 3.11.This genetic distance matrix was part of the larger matrix calculated automatically by GenAlEx package program to plot the PCA which was shown in Figure 3.14.

Table 3.11 The population genetic (240 bp in length mtDNA HVI) distances between the Çemialo Sırtı and 19 populations based on the genetic distance matrix of the PCA. Grec: Greece, Kazk: Kazakh, Tmen: Turkmen, Turk: Turkey, Uzbk: Uzbek, Cypr: Cyprus, Iran: Iran, Iraq: Iraq, Geor: Georgia, Azer: Azerbaijan, Arme: Armenia, Saga: Sagalassos, Cemi: Çemialo Sırtı, Neol: Neareastern Neolithic population from Syria, Syra: Syria, Lbkt: LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC), LBK: LBK (Linear Pottery culture, 5500-4775 cal BC), RSC: RSC (Rössen culture, 4625-4250 cal BC), CWC: CWC (Corded Ware culture, 2800-2200/2050 cal BC), BBC: BBC (Bell Beaker culture, 2500-2200/2050 cal BC), UC: UC (Unetice culture, 2200-1550 cal BC), CRS: Cambridge Reference Sequence (Anderson et al., 1981).

	Cemi
Neol	2,343
Arme	4,490
Azer	4,801
Iran	4,000
Kazk	5,023
Geor	4,124
Tmen	5,000
Uzbk	4,693
Cypr	4,009
Grec	4,141
Turk	4,163
Saga	4,431
Syra	4,232
Lbkt	4,308
UC	4,045
BBC	3,943
CWC	4,080
RSC	5,020
LBK	4,302
CRS	2,000

It can be observed from Figure 3.14 and Table 3.11 that Çemialo Sırtı individuals are close to the samples from the near eastern Neolithic population. Moreover, among the tested modern samples Çemialo Sırtı individuals seemed to be closer to the Cyprus population according to the PCA and distance table. The whole pairwise population matrix of population genetic distance was given in Appendix D.

Both of the PCA results indicated that Çemialo Sırtı population showed an affinity to the near eastern Neolithic population from Syria but not to Sagalassos population. The Neolithic populations from Europe were genetically similar to each other and also they seemed to be genetically similar to the populations from south and southeast Europe such as Greece, Bulgaria, Romania and Macedonia based on the mtDNA haplogroup frequencies data.

3.9 Population Continuity

In the previous sections, based on the comparative analysis, Çemialo Sırtı population indicated a genetic affinity to the near eastern Neolithic population from Syria. The approximate distance between the two sites is 900 km. In this section, it was asked if the relatively small difference between the two populations can be explained by the effect of random drift only in which case continuity between two ancient populations represented by 7 Çemialo Sırtı individuals and 10 neareastern Neolithic individuals from Syria cannot be rejected. "Continuity" between the populations was tested by the set of simulations. The simulations were performed using Fastsimcoal2 coalescence simulator as reported in section 2.2.8.2 (Fastsimcoal2 Simulations) of the Materials and Methods Chapter. In all of the simulations the generation time for human was assumed as 25 years (Sverrisdóttir et al., 2014). For all of the simulations effective population sizes (Ne) were assumed as a range from 250-5000 for the near eastern Neolithic population, while Ne was within the range of 5000-50000 for the Çemialo Sırtı population. The overlapping region of the mtDNA HVRI sequences was detected as 240 bp in length (between 16126-16366 np on the CRS (Anderson et al., 1981)). The archaeological dates of the 10 Neolithic and 7 Çemialo Sırtı individuals were assumed as 7300 BC and 500 BC, respectively. F_{ST} value between two ancient populations was calculated as 0.06310 using Arlequin 3.5.2 (Excoffier and Lischer 2010). Pairwise F_{ST} value was not statistically significant ($P>0.05$) by permutation test carried out by the same software. Then 1000 simulations for each combination of population sizes were performed. The proportion of simulated F_{ST} values that are greater than the observed F_{ST} value was calculated under the three

growth models. Plots of the three simulations were given in Figures 3.15-3.17. In the present study all the fastsimcoal2 simulations were run by Assist. Prof. Dr. Ayşegül Birand.

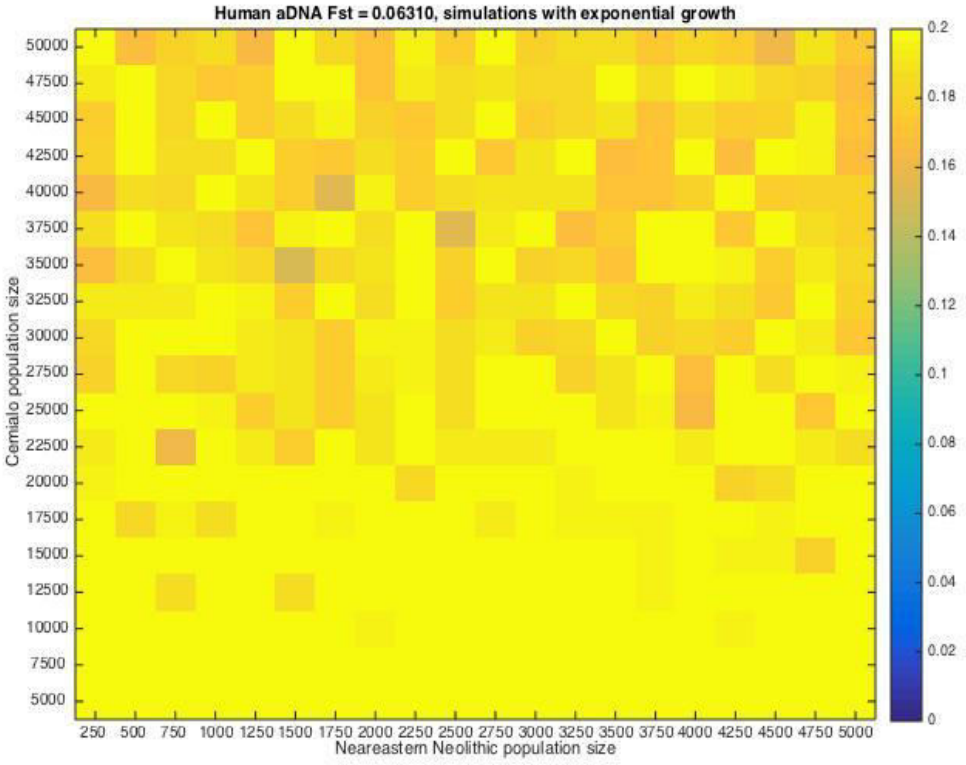


Figure 3.15 Probabilities (P) of obtaining simulated molecular F_{ST} values greater than the observed value of F_{ST} under the assumptions of "continuity" and exponential growth between Çemialo Sırtı and near eastern Neolithic individuals. P-values are shaded in accordance with the given scale by the side of the figure.

Figure 3.15 suggested that for all of the assumed sizes of two populations, observed F_{ST} value is in the expected region ($P>0.05$) of the distribution under random drift. Hence, one cannot reject the continuity between the two populations for the given parameters.

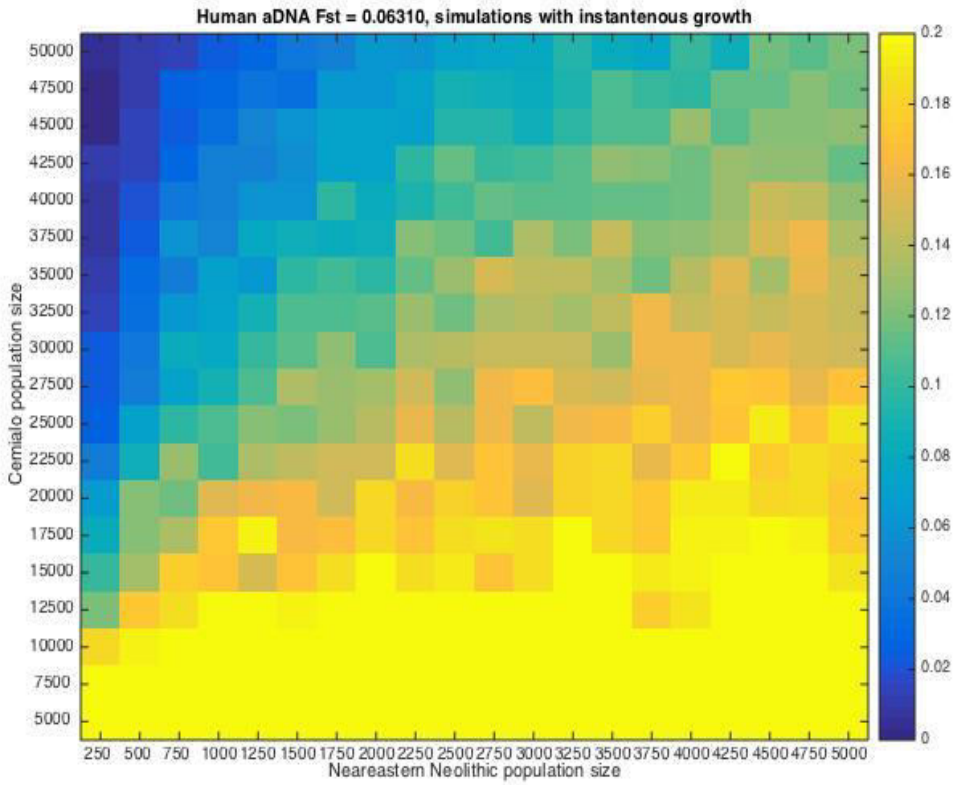


Figure 3.16 Probabilities (P) of obtaining simulated molecular F_{ST} values greater than the observed value of F_{ST} under the assumption of "continuity" and instantaneous growth between Çemialo Sırtı and near eastern Neolithic individuals. P-values are shaded in accordance with the given scale by the side of the figure.

Figure 3.16 indicated that if population was growing instantaneously from Neolithic population size of 250-500 to 22.500 or above of the Çemialo Sırtı in that case continuity can be rejected because observed F_{ST} value is an unlikely observation ($P<0.05$) under the assumption of presence of random drift only.

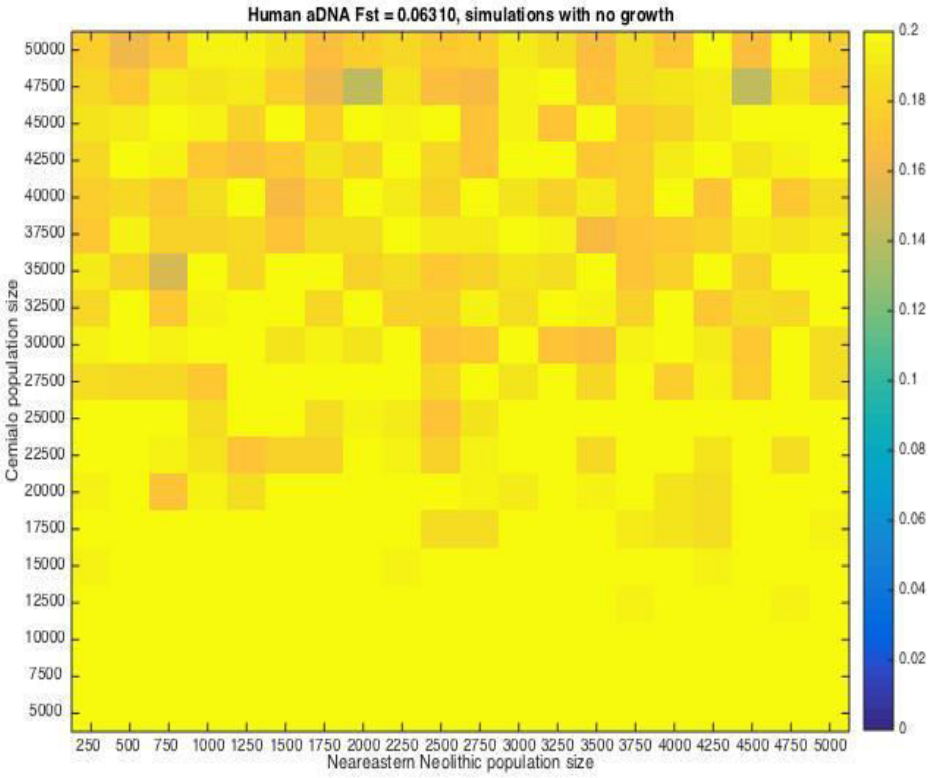


Figure 3.17 Probabilities (P) of obtaining simulated molecular F_{ST} values greater than the observed value of F_{ST} under the assumption of "continuity" and no growth between Çemialo Sırtı and near eastern Neolithic individuals. P-values are shaded in accordance with the given scale by the side of the figure.

Figure 3.17 suggested that for all of the assumed sizes of two populations, observed F_{ST} value is in the expected region ($P>0.05$) of the distribution under random drift. Hence, one cannot reject the continuity between the two populations for the given parameters.

CHAPTER 4

DISCUSSION

In the present study, mtDNA HVI and HVII region sequences (359 bp and 217 bp, respectively) of the 7 ancient samples collected from Çemialo Sırtı (Batman) excavation site were obtained. The mtDNA HPG compositions of ancient human samples from Çemialo Sırtı were determined and also the analysis of mtDNA HVI-HVII region sequences was performed.

4.1 Avoiding Contamination in Human aDNA Studies

For the reliability of the aDNA results, a dedicated aDNA laboratory is needed and the necessary measures to control the contaminations should be taken. The dedicated aDNA laboratory which is in the MODSIMMER Building at METU campus, was built in 2012 and all of the precautions (such as UV irradiation before the experiments, strictly limited access to the laboratory, cleaning all the benches with bleach, using laminar flow hood, DNA-free laboratory equipments and reagents) were taken. Additionally grinding of samples, aDNA extractions and PCR preparations were conducted in three different cabinets in two distinct rooms of the dedicated aDNA laboratory. All of the negative controls (of grinding, extraction and PCR steps) were clean.

Another important step in human aDNA studies is handling of ancient human samples, especially, when the basic Sanger Sequencing method is performed. Thus, the sample collecting method of the Ottoni et al.'s (2011) study was followed, in the field. Furthermore, to avoid the modern human contamination from ancient human

samples, additional handling methods were followed by archaeological team members, given in Section 2.1.1 (Sample Collection) of the Materials and Methods Chapter.

As another measure of contamination, modern mtDNA HVI and HVII regions of archaeological team members and researchers who had access to aDNA laboratory were studied to rule out their possible contamination to ancient samples.

In order to learn the experimental works for human aDNA extraction and analysis such as haplogroup determination, the author of the presented thesis visited Dr. Ottoni's laboratory in Leuven, Belgium and gained hands-on experiences.

4.2 Optimization of the Extraction and PCR Amplification Steps

When a method was adapted from another laboratory/study, it is quite likely that some modifications would be needed in the laboratory where the method is going to be established. These modifications are needed partly because the equipments of two laboratories are different. Moreover, features of the examined samples might be another factor that may cause to change in the protocols. The modifications were performed for some steps of both aDNA extraction and PCR amplification protocols. After the optimizations, indeed the results were better than previous results prior to optimizations, i.e. bands and sequences of mtDNA fragments were seen more clearly and the more efficient amount of aDNA was yielded.

4.3 Retrieval of aDNA

The success rate for the DNA recovery and amplification was 77.8% (7 out of 9 individuals) yielded amplifiable DNA for the Çemialo Sırtı samples (600-500 BC) and the success rate in extractions 100% (9 out of 9 individuals). Recovery rate for Sagalassos human population from Southwest Anatolia (Ottoni et al., 2011) was 62.4% and it was 23.8 % for the Near Eastern Neolithic population (approximately

8000 BC) from Syria (Fernandez et al., 2014). Additionally, in the study of Szécsényi-Nagy et al. (2015), for the STA (Early Neolithic Starčevo culture, 6000–5400 BC) and LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC) populations in total 76.1% and 22.9% recovery rates of the HVI and HVII regions were obtained, respectively. The recovery rates for the Central European Neolithic populations; LBK (Linear Pottery culture, 5500-4775 cal BC), RSC (Rössen culture, 4625-4250 cal BC), SCG (Schöningen group, 4100-3950 cal BC), BAC (Baalberge culture, 3950-3400 cal BC), SMC (Salzmünde culture, 3400-3100/3025 cal BC), BEC (Bernburg culture, 3100-2650 cal BC), CWC (Corded Ware culture, 2800-2200/2050 cal BC), BBC (Bell Beaker culture, 2500-2200/2050 cal BC) and UC (Unetice culture, 2200-1550 cal BC) are 88%, 52.4%, 100%, 67.9%, 96.7%, 85%, 77.2%, 93.5% and 83.2%, respectively, while it was % 84.1 in total (364 out of 433 samples) (Brandt et al., 2013). In the Rivollat et al.'s (2015) study, for the Gurgy group (5000-4000 cal. BC) which was an Early/Middle Neolithic population from the southern region of the Paris Basin, for HVI region 38.2% was the observed recovery rate. It can be anticipated that, age of the material and the latitude of the site of sample together with the micro environmental conditions at the site affect the recovery rate. Since samples of the present study were nearly 7500 years younger than those of northern Syria (Fernandez et al., 2014) and 3500 years younger of those obtained from Paris Basin (Rivollat et al., 2015), samples of the present study had considerably better recovery rate. Although the samples were older among the other European samples' (from Balkans and Central Europe) recovery rate of the present study is close to the high end of the range of observed values (100%-22.9%). The high recovery rates of the central European Neolithic samples might be due to the climate of the sites which are cooler than the climates of Anatolia and Near East. Therefore, the aDNA of central European samples might have been well-preserved. Recovery rate observed in the present study is slightly better than the Sagalassos' which may be by chance, because the number of studied samples were small (n=9) in the present study.

Incidentally, in the study of Fernandez et al. (2014) the amplified sequences of mtDNA HVI region of each sample were shorter in comparison to similar studies (Ottoni et al., 2011; Brandt et al., 2013; Rivollat et al., 2015; Szécsényi-Nagy et al., 2015) and this might be due to the ages of samples and the conditions of excavation sites. In comparative aDNA analysis in the present study, the minimum size of the common sequence is determined by the sequence length of Neolithic northern Syrian samples (Fernandez et al., 2014).

Additionally, the age of each ancient individual and the size of each sample might be some of the other factors affecting the recovery rates and the effective recovery of aDNA (Ottoni et al., 2011). For example, in the present study enough amount of aDNA could not be obtained from a small tooth sample of individual SK29. Similarly, for SK30 which belonged to a small child results cannot be observed. Thus, sufficient amount of aDNA which was necessary for the repeated amplifications of mtDNA HVI and HVII regions, could not be obtained from individuals SK29 and SK30.

In addition the quality of the sequences and their peaks in different fragments were not the same even for the same individual. For instance, success rate was relatively low for fragment E than fragment D. However, in the present study longer sequences than the targeted ones of the fragment D were obtained and this result was helpful to compensate for lower success rate of fragment E. Longer sequences provided an opportunity for the observations of the putative transitions at the overlapping regions of fragments E and D.

4.4 Authenticity of aDNA

In the present study, to confirm that every fragment belongs to the same individual seven short overlapping fragments were used to amplify human mtDNA HVI and HVII regions. Overlapping sequences supported the existence of just one unique sequence of each individual. Since there were at least two repeated sequences for

each overlapping fragment, each repeated sequence of the fragments was from an independent PCR amplification. Also PCR amplifications were from at least two different aDNA extractions of each sample. The sequences from these repeated steps were obtained for each individual and supported the existence of just one unique sequence of each individual. Moreover it was ensured that any of the observed sequences does not match with any of the mtDNA sequences of research team members'.

These were the first set of requirements in relation to the authenticity of extracted aDNA. There was an additional criterion to be fulfilled after comparatively examining the repeated sequences of the individual. It was the presence of misincorporations on the repeated sequences which is discussed in the next section of the present chapter.

4.4.1 Nucleotide misincorporations

Nucleotide differences are observed between the repeated sequences of the ancient samples much more frequently than it is observed in the modern sequences (Hansen et al., 2001; Hofreiter et al., 2001; Pääbo et al., 2004). Only one of these alternative nucleotides is the authentic one whereas the other one is incorporated as a result of some biochemical processes. The incorporation of non-authentic nucleotide is called as misincorporation, As some of the misincorporations, nucleotide transitions could take place in two different ways on the aDNA sequences and are described as: misincorporations of Type1 and Type2. The first type (Type1) occurs during the PCR amplification as the PCR artefacts by the Taq DNA polymerase irrespective of the quality of DNA template (Pääbo et al., 2004; Olivieri et al., 2010). The second type (Type2) of nucleotide transitions is defined as postmortem nucleotide changes on aDNA sequences and these are considered as an indication of authenticity (Pääbo et al., 2004; Olivieri et al., 2010). Since the C to T (C→ T) transitions among the Type2 transitions were the most commonly observed ones in aDNA, their presence indicates existence of postmortem nucleotide changes. Whereas, the transversions

were observed less frequently than the transitions not only in aDNA studies but also in other DNA studies (Hofreiter et al., 2001).

To be able to detect the misincorporations, repeated sequences of the same fragments are needed. Repeated sequences can be obtained by different methods: amplifying and sequencing the same region of the aDNA several times, by means of cloning, shotgun sequencing and next generation sequencing (references for the latter case can be seen in Sawyer et al., 2012; Rohland et al., 2015; Knapp et al., 2015).
When the numbers of obtained sequences are high, for instance 10, the nucleotide type (e.g. C) which has the highest frequency is accepted as the nucleotide of the authentic sequence. In the present study, amplification numbers were not high (2-6), and mostly 2, however, the misincorporations could be detected when there were more than 2 repeats of each fragment sequences. The presence of misincorporations was used to confirm that the DNA sequenced is the aDNA. The misincorporations were further used to estimate the original sequence of the aDNA. Except in the fragment sequences of one individual (SK21), the misincorporations could be observed in most of the fragment sequences of other individuals. There were only C→ T transitions among the Type2 transitions, but no G→A. There might be few mistakes in assigning the authentic nucleotides in misincorporation positions. For instance, for a C→ T transition event, in all of the amplified sequences T might be observed by chance. In this case, instead of C nucleotide, T nucleotide might be assumed as the authentic nucleotide. However, it is believed that most of the assignments are correct since the frequencies of the two alternative nucleotides were not equal at any misincorporation sites. In the present study, frequency of misincorporations is reduced, because an enzyme uracil–DNA–glycosylase (UDG) is used. Among the misincorporation types especially C→ T transitions were observed. The expected effect of UDG treatment (Pääbo et al., 2004; Rohland et al., 2015) is explained in the next section.

4.4.2 UDG (Uracil–DNA–Glycosylase) treatment

Uracil–DNA–glycosylase (UDG) is an enzyme which removes deaminated cytosines (uracils) from the DNA and causes an abasic site on the DNA sequence (Pruvost et al., 2005; Rohland et al., 2015). In aDNA studies the UDG is used for two reasons: 1) One is to avoid carryover contamination (contamination from the previously amplified or cloned products such as "jumping PCR") from the previously amplified or cloned sequences (Pruvost et al., 2005). When the UDG treatment is carried out during the PCR amplification procedure, the dUTP should be used instead of dTTP, so that PCR products contain uracil rather than thymine. Then, in the next PCR, if the uracil-containing carryover PCR products from previous amplifications contaminate the new PCR, UDG degrades the contaminant PCR products (Longo et al., 1990; Pruvost et al., 2005). 2) Another reason of using UDG in the aDNA studies is removal of the deaminated cytosine (C) nucleotide which changes into the uracil (U) residue. In the case of UDG treatment, firstly UDG removes the uracils from DNA molecule. Then Taq DNA polymerase enzyme uses the other strand of DNA molecule which does not contain uracil. With this mechanism UDG and Taq DNA polymerase enzyme, together, prevents the transitions from C to T or G to A by removing the uracil nucleotide at the beginning of PCR amplification (Hansen et al 2001; Pääbo et al., 2004; Rohland et al., 2015).

Recently, Rohland et al. (2015) performed three types of library preparation for next generation sequencing; with no UDG treatment, with partial UDG treatment and with regular (full) UDG treatment. This study revealed that in partial UDG treatment case some deaminated cytosines of aDNA molecule which were located on the terminal sides of aDNA, were not efficiently removed from the aDNA molecule (Rohland et al., 2015). However, with the full UDG treatment uracils from aDNA molecules were removed successfully (Rohland et al., 2015). In the case of without UDG treatment all C to T transitions were observed and its ratio was high at the 5' terminal, while it was decreasing towards the centre of aDNA molecule (Varshney et al., 1991; Meyer et al., 2012; Rohland et al., 2015).

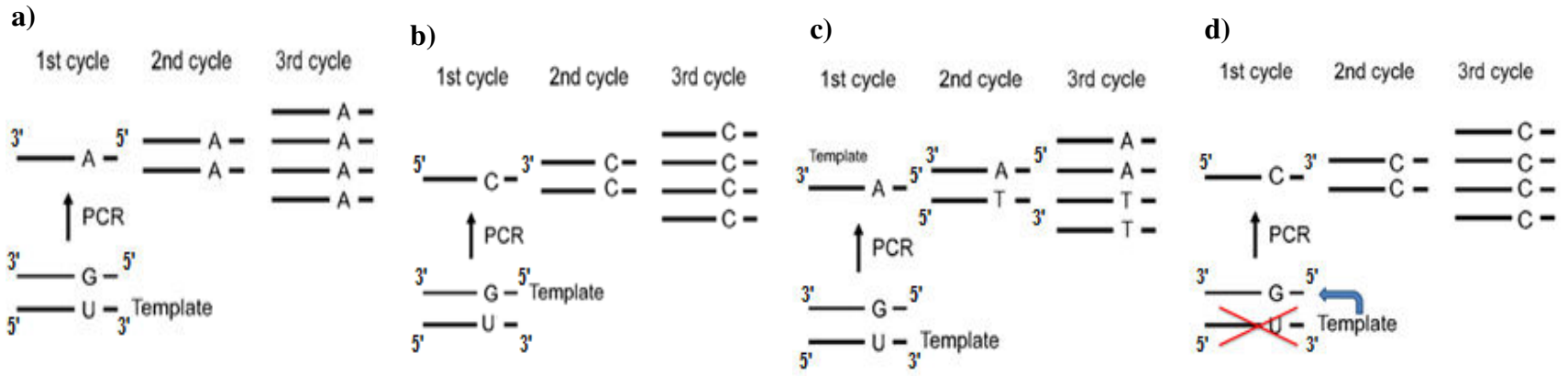
However, if the C to T transition happens due to the deamination of 5-methyl-cytocine (5mC), the UDG treatment cannot detect and remove the damages, as transitions, on aDNA molecule (Pedersen et al., 2014). In the case of methylation, first cytosine is methylated and causes 5-methyl-cytocine (5mC); then the C to T transition happens by the deamination of 5mC. Thus, the changes directly from 5mC to thymine could also cause the C to T transition on aDNA molecule and they may be observed even when full UDG treatment is applied (Pedersen et al., 2014).

In the present study full UDG treatment was performed in most of the amplified sequences (n=118, 92.2%). Thus, C to T transitions which were due to deamination of cytosine nucleotide should not have been observed in these UDG treated amplified sequences. However, in a few of the amplified sequences (n=7, 7.8%) (one sequence of each fragment C and E of SK24, one sequence of fragment A of SK26, one sequence of each fragment E and G of SK28, one sequence of fragment E of SK7 and one sequence of fragment E of SK17) C to T transitions were obtained without UDG treatment. Excess C to T transitions which were observed mostly on fragment E must be due to amplified sequences of E fragment which were without UDG treatment. As in Rohland et al.'s (2015) study, in UDG untreated amplifications (for instance 4 times in fragment E) the C to T (Type2) misincorporations are very frequent (n=7, 0.04%) when all of the samples are considered in presented study, as it can be seen in Tables 3.4-3.9.

During the experiments it was observed that the PCR amplifications without UDG treatment were more successful than those with fully UDG-treated ones in the case of present study. It might be due to the degradation of aDNA molecules which contains uracil residues, right before the start of PCR amplification (just at the beginning of PCR reaction). Furthermore on some of the amplified sequences (n=1, 0.78%) C to T transitions were also detected, even when they were treated with full UDG method. Thus, in the present study I suggest that the observed C to T transitions from the UDG-treated sequences might be due to the nucleotide damages from 5-methyl-cytocine (5mC) to thymine on the aDNA molecule. For this case, another possible

explanation may be the lack of proper reaction of UDG enzyme. Also there may be a mistake during the incorporation of nucleotide at the thermocycling step by chance. The two mechanisms (cytocine deamination and another plausible 5mC deamination) of C to T transition which were observed in the present study, are presented in Figures 4.1-4.2, schematically.

In the present study, within the limits of repeat numbers of sequences, high number of Type2 transitions and especially C to T transitions can be considered as supporting authenticity of the sequences. Sawyer et al. (2012) found that the samples older than 500 years, might have at least 10% nucleotide damage rate without UDG treatment and this could suggest the plausible authenticity of aDNA. Furthermore, Rohland et al. (2015) suggested that when the partially UDG treatment was performed, on average, the damage (from 5' terminal side C to T or on the reverse complement from 3' terminal side G to A) rate can be 3% or higher, also after the partially UDG treatment. The partially UDG treatment mostly removes the uracils within the aDNA molecule similar to the full UDG treatment, but generally leaves the uracils at the terminals of DNA molecules (Rohland et al., 2015). In another study by Haak et al. (2015) the library preparations were also performed in three ways, without UDG treatment, with full UDG treatment and partial UDG treatment. Experiments without UDG treatment were to screen the authenticity of the aDNA and in this case the C toT transition rate, in first base, was observed in a range between 0.002-0.482 using the mtDNA capture method, with a mean of 0.124 (Haak et al., 2015). The full and partial UDG treatments were also carried out using nuclear DNA obtained using the 390k SNP capture method (Haak et al., 2015). Then, in Haak et al.'s (2015) study C to T transition rates, in first base, were observed between 0.001-0.15 according to samples, with a mean of 0.05. Thus, it must be noted that in the present study, when full UDG treatment method was used, the observed Type2 (C to T) transition rate was (n=1, 0.005%) within the ranges of the previous two studies.



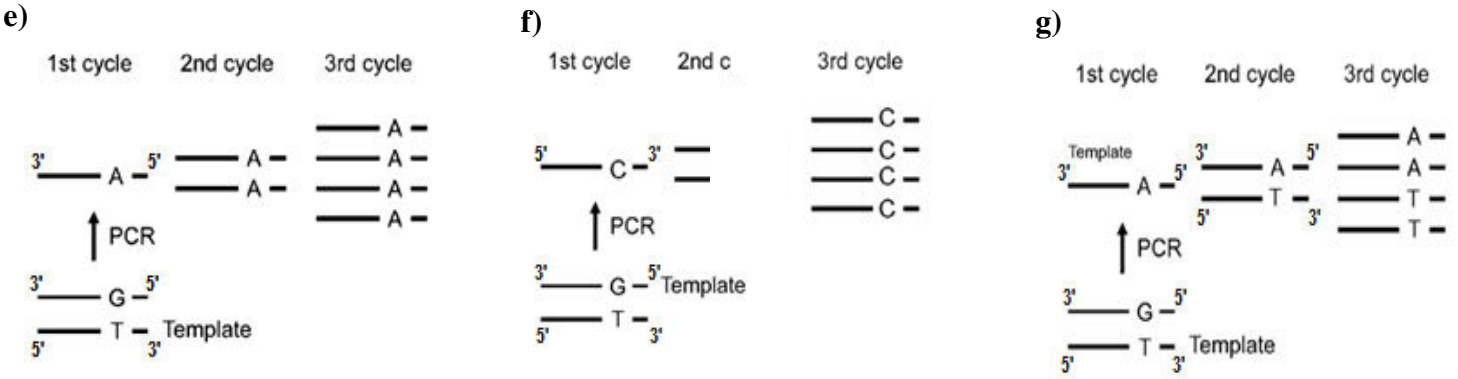
a)The amplification of aDNA molecules during the PCR when there is cytosine deamination and thus it is converted to uracil. This strand of aDNA molecule is taken as the template for amplification, then G to A transition is observed. PCR amplification is without UDG treatment.

b)The amplification of aDNA molecules during the PCR when there is cytosine deamination and thus it is converted to uracil. The other strand of aDNA molecule is taken as the template for amplification. It shows the replications of aDNA strand which contains cytosine. PCR amplification is without UDG treatment.

c)The amplification of aDNA molecules during the PCR when there is cytosine deamination and thus it is converted to uracil. This aDNA molecule is taken as the template for amplification. Then the adenine-contain aDNA strand was taken as template at the end of 1st cycle and replicated. Hence C to T transition is observed. PCR amplification is without UDG treatment.

d)The amplification of aDNA molecules during the PCR when there is cytosine deamination and thus it is converted to uracil. It shows that the UDG enzyme removed the aDNA strand which contained uracil. Hence other strand of aDNA molecule is taken as the template for amplification and C to T transition is not observed. It shows the removal mechanism of UDG enzyme. PCR amplification is with full UDG treatment.

Figure 4.1 Schematic representations (a, b, c and d) of the two kind of Type2 transitions during the PCR amplifications with and without UDG treatment when there was the cytosine deamination (U) on the aDNA molecule. The figures were taken from Pääbo et al.'s (2004) study and modified to show possible cases that might have been observed in the present study.



e)The amplification of aDNA molecules during the PCR when there is cytosine metylation and thus it is converted to thymine. This aDNA molecule is taken as the template for amplification, then G to A transition is observed. In this case of methylation, PCR amplifications are the same with and without UDG treatment. Because there is no uracil and thus UDG does not affect to aDNA molecule.

f)The amplification of aDNA molecules during the PCR when there is cytosine metylation and thus it is converted to thymine. Another strand of aDNA molecule is taken as the template for amplification. It shows the replications of aDNA strand which contains cytosine. In this case of methylation PCR amplifications are the same with and without UDG treatment. Because there is no uracil and thus UDG does not affect to aDNA molecule.

g)The amplification of aDNA molecules during the PCR when there is cytosine metylation thus it is converted to thymine. This strand of aDNA molecule is taken as the template for amplification. The adenine-contain aDNA strand was taken as template at the end of 1st cycle and replicated. Then the both C to T and G to A transitions are observed. In this case of methylation PCR amplifications are the same with and without UDG treatment. Because there is no uracil and thus UDG does not affect to aDNA molecule.

Figure 4.2 Schematic representations (e, f and g) of the two kind of Type2 transitions during the PCR amplifications with and without UDG treatment when there was the cytosine methylation (T) on the aDNA molecule. The figures were taken from Pääbo et al.'s (2004) study and modified to show possible cases that might have been observed in the present study.

4.5 mtDNA Haplogroups of Ancient Çemialo Sırtı Individuals and Their Worldwide Distributions

When both of the sequences of mtDNA HVI and HVII regions were used not only the macrohaplogroups but some minor haplogroups could also be determined. In the studies of Ottoni et al. (2011) and Szécsényi-Nagy et al. (2015) both of the mtDNA HVI-HVII regions were amplified and sequenced for haplogroup determination. On the other hand the mtDNA haplogroups can also be determined based on only HVRI (HVI region) which was the case in Fernandez et al.'s (2014) study.

In the present study, both regions of HVR (Hypervariable region) were employed and certain informative mutation motifs of HVRII (HVII region) were used for the determination and confirmation of mtDNA haplogroups of Çemialo Sırtı individuals. For instance, the A73G! ("!" indicates the back mutation) mutation motif on HVRII helped to determine the haplogroup of the individual coded as SK7 as haplogroup H1a, while the absence of the same mutation motif on HVRII would imply that SK24 is H1z1. If there were no information from HVRII both of them would be labelled as haplogroup H. Likewise, for SK17, absence of A73G! mutation motif on the HVRII region indicated that SK17 is haplogroup HV. The mutation motifs T146C!, T152C! and T195C! on the HVRII region were used for the determination and confirmation of haplogroups U2b1 (SK28), R2 (SK17) and R6 (SK35), respectively, which are the other haplogroups found among Çemialo Sırtı individuals. However, for haplogroup M1a1 (SK26) the mutation motifs on HVRI (especially the T16359C) was informative enough for the haplogroup determination and the motifs of HVRII did not contribute to the haplogroup determination in this individual. In the present study the list of observed haplogroups are as follows H1z1, M1a1, U2b1, H1a, HV, R2 and R6.

In terms of macrohaplogroups H, HV, M, R and U2 are observed with the frequencies: 28.57% (2/7), 14.29% (1/7), 14.29% (1/7), 28.57% (2/7) and 14.29% (1/7), respectively. The frequencies of macrohaplogroups indicate that haplogroup H

and R are the most common haplogroups among the studied Çemialo Sırtı individuals. By studying the geographic distributions of mtDNA haplogroups in modern populations, certain haplogroups are related with some groups of modern human populations (Metspalu et al., 2004; Quintana-Murci et al., 2004). For instance, haplogroup H is the most common haplogroup in Europe (Ghezzi et al., 2005) and it is also frequently found in North Africa and Middle East (Richards et al. 2000; Ennafaa et al., 2009). Haplogroups R0, H, V, HV, U, K, J and T are the branches of macrohaplogroup R and it is known as the haplogroup of west Eurasia (Quintana-Murci et al., 2004; Ottoni et al., 2011).

The haplogroup M1a1 could be identified by a specific mutation motif as T16359C (from T to C at 16359 np) of the mtDNA HVRI (Gonzalez et al., 2007) and it is frequently found in Ethiopia and East Africa (Gonzalez et al., 2007).

On the other hand, in a recent study, by Brandt et al. (2013), it was found that the LBK (Linear Pottery culture, 5500-4775 cal BC) is characterized with a special group of haplogroups N1a, T2, K, J, HV, V, W, X therefore, collectively they are called as a mitochondrial "Neolithic package" (Brandt et al., 2013). The haplogroup H was common in Neolithic and modern populations but, since it is also observed in pre-agricultural groups in Iberia, it was not in the Neolithic package (Hervella et al., 2012; Szécsényi-Nagy et al., 2015). It is suggested that in relation to haplogroup HV which is from the Neolithic package, originated in the Near East and Caucasus region (Malyarchuk et al., 2008; Szécsényi-Nagy et al., 2015). Haplogroup HV is observed today in south western Asia, especially in Iran, Anatolia regions (Richards et al. 2000; Tambets et al. 2000; Malyarchuk et al., 2008). The haplotype of Çemialo Sırtı individual (SK17) which contains only the mutation at 16311np (T16311C!) of the mtDNA HVRI, was also included in the haplogroup HV. This haplotype is observed with low frequencies in north, central and north western Europe populations (Richards et al. 2000; Al-Zahery et al. 2003; Metspalu et al. 2004; Quintana-Murci et al. 2004; Malyarchuk et al., 2008). Additionally, this haplotype

was also observed in one individual from LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC) populations (Szécsényi-Nagy et al., 2015).

The central European hunter-gatherers were defined by mostly U lineages such as U, U4, U5, and U8 with respect to mtDNA data (Bramanti et al., 2009; Fu et al., 2013). In the LBK (Linear Pottery culture, 5500-4775 cal BC) population mtDNA lineages of the hunter-gatherers are found rarely (Brandt et al., 2013). Likewise, the hunter-gatherer haplogroups were also rare in both STA (Early Neolithic Starčevo culture, 6000–5400 BC) and LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC) populations (Szécsényi-Nagy et al., 2015). It was suggested that this shift between hunter-gatherers and Neolithic populations in central Europe could be due to the rapid transition with a genetic invasion from the Near East, Anatolia and Caucasus (Brandt et al., 2013).

Haplogroup U2 was determined for one individual (SK28) among the studied Çemialo Sırtı samples as subhaplogroup U2b1. In Szécsényi-Nagy et al.'s (2015) study, haplogroup of an individual from LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC) population was determined as U2. Haplogroup U2 was also found in CWC (Corded Ware culture, 2800-2200/2050 cal BC), BBC (Bell Beaker culture, 2500-2200/2050 cal BC) and UC (Unetice culture, 2200-1550 cal BC) which are three groups of the central European Neolithic populations, while it was absent in Gurgy (Neolithic Population from southern of Paris Basin) and LBK (Linear Pottery culture, 5500-4775 cal BC) populations (Brandt et al., 2013; Rivollat et al., 2015). It was suggested that CWC is characterized by haplogroups I and U2 and occurred with the hunter-gatherer haplogroups U4 and U5 and haplogroup T1which is the Late Neolithic/EBA (Early Bronze Age) lineage (Brandt et al., 2013). Thus, in CWC group the frequency of Early/Middle Neolithic haplogroups are lower than other the Central European Neolithic groups (Brandt et al., 2013). In the study of Brandt et al. (2013) their analysis indicated that the CWC has a genetic similarity to two ancient Kurgan groups of South Siberia (Keyser et al., 2009) and Kazakhstan (Lalueza-Fox et al., 2004) due to the I, U2, and T1 haplogroup composition (Brandt

et al., 2013). It is known that haplogroup U2 was especially found in Palaeolithic, Mesolithic, and Bronze Age samples from Russia (Keyser et al., 2009; Krause et al., 2010; Der Sarkissian et al., 2013; Brandt et al., 2013).

The haplogroup U2 was also observed in other ancient samples. For example, in the study of Krause et al. (2010) mtDNA haplogroup of very old human sample, 30000-year-old hunter-gatherer, from the site on Don River in Russia, was determined as U2. Furthermore, the haplogroup U2 was also found in the samples among Mesolithic European hunter-gatherers such as 11,000 year-old sample (U2e) from Blätterhöhle in Germany (Bollongino et al. 2013), two 9500 year-old samples (U2e) from Karelia in Russia (Der Sarkissian, 2011) and two 8000 year-old samples (U2e1) from Motala in Sweden (Lazaridis et al. 2014). The haplogroup U2 is mostly common in modern south Asian populations while it is found in low frequency in the central and western Asia (Metspalu et al., 2004). The U2e which is a subclade of U2, is also found in Europe (Maji et al., 2008).

Another macrohaplogroup observed among the studied Çemialo Sırtı individuals is R, which was detected in two Çemialo Sırtı individuals, SK21 and SK35, as two subhaplogroups R2 and R6, respectively. The haplogroup R was not found in the Near Eastern Neolithic individuals (Fernandez et al., 2014), while the subhaplogroup R0 were observed in the Sagalassos population (Ottoni et al., 2011). Additionally in the study of Szécsényi-Nagy et al. (2015), R haplogroup was detected for only one individual (1/39) from LBKT population. Haplogroup R was found for few individuals in RSC, SMC, BEC, CWC, BBC and UC groups from Central European Neolithic populations, while it was absent in LBK and Gurgy Neolithic populations (Brandt et al., 2013; Rivollat et al., 2015; Szécsényi-Nagy et al., 2015). The macrohaplogroup R is the most common haplogroup in west Eurasian modern populations, while the subhaplogroup R2 is observed at low frequencies in the Near/Middle East and India (Metspalu et al. 2004).

In the Ottoni et al.'s (2011) study: X, W and N1b haplogroups of Sagalassos (Byzantium samples from southwest Anatolia) samples are part of the macrohaplogoup N (none of them are observed in Çemialo Sırtı yet). The observed H, HV (present in the Neolithic package) haplogroups of Çemialo Sırtı were also found in the Sagalassos population. In parallel to R in Çemialo Sırtı, R0 was observed as a subhaplogroup of R, in Sagalassos. In addition, the East/South Asian macrohaplogroup such as M and subSaharan haplogroups such as L1, L2, L3 were not found in Sagalassos population, while the macrohaplogroup M was observed among Çemialo Sırtı individuals.

In the Fernandez et al.'s (2014) study, among the Neolithic samples from Northern Syria, the following haplogroups were observed: H (3/12), HV (1/12), K (4/12), L3 (1/12), CRS (3/12). Presence of high frequency of haplogroup H and haplogroup HV is similar to results from Çemialo Sırtı. The macrohaplogroup M was not detected in the study by Fernandez et al. (2014), African HPG (Haplogroup) L was observed, HV, K and N components of Neolithic package were also present in Neolithic northern Syrian population.

In general, haplogroups observed in Çemialo Sırtı so far are the ones expected in the region. The macrohaplogroup N is one of the haplogroups that is found in Levant region and was not detected among the Çemialo Sırtı individuals yet, while it was found in the Sagalassos population, in Neolithic population from Northern Syria, STA and LBKT populations including five Central European Neolithic (LBK, RSC, SCG, BAC, SMC) and Gurgy Neolithic groups (Ottoni et al., 2011; Brandt et al., 2013; Fernandez et al., 2014; Szécsényi-Nagy et al., 2015; Rivollat et al., 2015). Moreover, it was observed in a higher frequency in LBK than other Neolithic groups. It can be expected that when further samples will be examined from Çemialo Sırtı N haplogroup may also be seen.

Furthermore, it can be seen that Çemialo Sırtı individuals, based on their 7 haplogroups, do not seem to be very close to neither ancient nor modern populations

which were used for the comparative analysis in presented study. However, one may ask the following question: among the studied modern and ancient populations, which populations are relatively more similar to Çemialo Sırtı genetically? With the help of a multivariate analysis, for instance PCA, one can obtain an answer to the question. PCA analyses were carried out on the basis of two different markers: mtDNA HVRI haplogroups and mtDNA HVRI-HVRII sequences, presented in the following sections.

4.6 PCAs Based on Ancient Samples and Modern Populations

The PCA with haplogroup frequencies data showed that mtDNA haplogroup composition of Çemialo Sırtı population is relatively similar to the near eastern Neolithic population from Syria. This could be based on the similar frequency of haplogroups H and HV, while the Neolithic population from Syria is close to other Neolithic populations from central Europe due to the its high frequencies of haplogroups K and CRS.

According to Szécsényi-Nagy et al.'s (2015) PCA results, STA population showed an affinity to the Near East and the Caucasus populations, whereas the LBKT was between the populations from South and southeast Europe such as Greeks, Bulgarians and Italians. They also suggested this genetic affinity of LBKT population might be due to the higher frequency of haplogroup H in the LBKT Neolithic population.

PCA is a useful multivariate method by which one can visualize the pairwise genetic proximity between many populations on the basis of many variables. Here 30 populations and 30 haplogroup frequencies (Figure 3.13) or 240 bp long sequences (Figure 3.14) were used. For instance when pie charts for haplogroups of selected populations were used it was not possible to observe the existing genetic similarities between the populations. However both of the PCA results indicated that Çemialo Sırtı population showed genetic affinity to the Neolithic population from Syria.

Nevertheless, by PCA graphs, only certain portion of the total genetic diversity that exists between the populations can be presented. In the present study, for haplogroup frequencies the total genetic diversity accounted by the graph was 30.03% (Figure 3.13), but, it was 67.20% when the sequences were used (Figure 3.14). Therefore, Figure 3.14, accounting higher proportion of the total variation, can be considered as a better summary of the genetic relationship between the populations under consideration. In Figure 3.14 it was observed that as well as Çemialo Sırtı individuals had a genetic affinity to the Near Eastern Neolithic population from Syria, it was also closer to Cyprus than to any other modern populations that were tested. The observed proximity between Cyprus and Neolithic population from Syria was already considered as the evidence of spread of Neolithic life style by Seafaring (Fernandez et al., 2014). Results of the present study might be suggesting that Neolithic population of the Northern Syria, Çemialo Sırtı population and modern Cyprus population had relations perhaps by descent and Çemialo Sırtı at least by 500BC was highly related with Neolithic population of Northern Syria despite the fact that it is approximately 900 km away and had been under the rules of different civilizations (e.g. Median Empire, Persian Achaemenid Empire).

Possible continuity between the Çemialo Sırtı population and Neolithic population from Syria is tested by simulations which were carried out using the Fastsimcoal2 software (Excoffier et al., 2013).

As the last two notes: 1) Neolithic populations from Central Europe are close to each other in both of the PCAs, whereas based on the PCA with haplogroup frequencies data they are between the populations from south and southeast Europe such as Greece, Bulgaria, Romania and Macedonia; similar to PCA results of Szécsényi-Nagy et al.'s (2015); 2) Modern Turkish population as a whole is quite different than Çemialo Sırtı population. Together with Iran and Azerbaijan populations, it is slightly closer to western Central Asian populations, for instance to Uzbekistan, from where it received nomadic Turks (Roux, 1997; Zerjal et al., 2002; Salman, 2004) compared to other populations of the region such as Armenia, Georgia and Syria.

One must remember that results of these comparative studies are very much depends on the samples employed and the number of sample sizes.

4.7 Population Continuity

Analysis revealed that under two growth models (no growth/exponential growth) from Neolithic population of Syria to Çemialo Sırtı population the differences in their sequences are similar to the ones that can be seen in a randomly drifting continuous population. In other words, the results of simulations under two growth models (no growth / exponential growth) suggested that, for the mtDNA (maternal lineage hence the females) does not seemed to be affected by the migrations since the Neolithic ages in the Batman- Northern Syrian region. This does not mean that there was no migration to the region but if there were migrations perhaps the immigrants were not genetically different from the hosts. However, for some of the parameters of the populations, population continuity can be rejected under the assumption of instantaneous growth model. Yet, in relation to instantaneous growth model one may not expect to have a sudden effective population size increase from 250-500 of Neolithic population to 22500 or more of Çemialo Sırtı population. If there was such a change, it can be explained by a sudden very effective migration to the region. More accurate effective population size estimations, for instance, by whole genome sequencing, may give a deeper insight about the evolutionary forces operating on the populations. Very similar observation, under the assumption of exponential growth, was made in Iberia. In north eastern Iberia the continuity could not be rejected between the Neolithic and modern populations (Sverrisdóttir et al., 2014) using 18 combined Early/Middle Neolithic and 118 modern mtDNA HVI region sequences. Sverrisdóttir et al.'s (2014) study suggested that population continuity in northeast Iberia was not rejected. Because, the ancient samples were from differentiated populations, Sverrisdóttir et al. (2014) also performed coalescent simulations for each of the ancient populations separately, using 7 mtDNA HVI region sequences from early Neolithic and 11 sequences from middle Neolithic with the modern northeast Iberian samples. Again the population continuity in northeast Iberia was not rejected (Sverrisdóttir et al., 2014). In the same study, Sverrisdóttir et al. (2014)

also observed the population continuity in south eastern France between the late Neolithic and present day populations using 29 ancient and 106 modern mtDNA HVI region sequences (Sverrisdóttir et al., 2014).

The number of the observations in the present study was low (just 7). When the number is low uncertainty in the estimates could be high and therefore results might be easily accommodating the continuity between the populations. In other words, continuity between the populations could have been observed because of the low sample sizes of the populations. In many of the studies based on mtDNA regions continuity between the populations were observed (for instance in Sverrisdóttir et al., 2014) and in many of these studies at least one of the population sizes were low (7, 11,29). In Sverrisdóttir et al.'s (2014) study, the overlapping region was 310 bp in length (between 16053-16362 np on the CRS) for northeast Iberian samples and 308 bp in length (between 16055-16362 np on the CRS) for south eastern France samples. In the presented study, the overlapping region of the sequences was 240 bp in length (between 16126-16366 np on the CRS (Anderson et al., 1981)) due to the shorter sequences of Neolithic population from Syria. This overlapping region (240 bp in length) is shorter than other overlapping regions employed in two previous examples (Sverrisdóttir et al., 2014), which could have further increased the chance of observing continuity between the two ancient populations.

If the exponential growth model is the appropriate growth model for the humans and if the observed continuities between the Neolithic and more recent populations (500 BC for the present study and modern populations in Sverrisdóttir et al.'s(2014) study) were real, the results suggest that females on both ends of Mediterranean had not changed much at least after the Neolithic.

4.8 The importance of the present study

This study is the first study where the questions about the human aDNA are addressed in the dedicated ancient DNA laboratory of Middle East Technical

University (METU), Ankara, Turkey. Despite the small number of samples, Çemialo Sırtı was also examined for the first time in terms of human mtDNA HPGs. Moreover by using comparative analysis between Çemialo Sırtı population, various modern populations (such as modern Turkish, Cypriot and Iranian) and ancient populations from the region (ancient Sagalassos population, Neolithic population from Syria, other Neolithic populations from Central European regions) contributed genetically to the understanding of human population history in Anatolia.

Furthermore the optimization of some steps of both extractions and PCR amplifications protocols were carried out for the Çemialo Sırtı sample set. Then it was observed that after the optimizations better results for aDNA were obtained for all of the samples in the present study.

By utilizing aDNA from archaeological human material as in the present study, it is possible to infer past human population changes, such as the continuity and similarity between two past populations in the geographically closer regions.

It is believed that by the present study and by the similar studies a new avenue will be established where numerous joint and significant studies between anthropologists, archaeologists and biologists will be carried out in Turkey, soon.

4.9 Future Studies

To increase the mtDNA data of ancient human populations in Anatolia, the presented study here needs to be further expanded with more samples (at least 5 more) which were collected previously from the Çemialo Sırtı excavation site. By increasing the number of samples, the conclusion of the Çemialo Sırtı population will be more robust and informative. Also preliminary information about the mtDNA haplogroup composition of the Çemialo Sırtı population would be better understood. Further genetic studies on ancient samples from southeast Anatolia, in turn, will contribute to the understanding of the ancient history and the peopling of Anatolia. Moreover three

samples from the Çemialo Sırtı excavation site, very soon, are going to be sent to C14 aging for the accurate timing of the period.

Very soon, the SNP (Single Nucleotide Polymorphism) analysis of the human mtDNA this time based on not HVI-HVII regions of mtDNA, but on coding region will be carried out for all of the Çemialo Sırtı samples. Results will enable us to confirm the previous observations independently. The SNP analysis method can provide additional information for the unambiguous assignment of the not only mtDNA haplogroups but also subhaplogroups.

Further tests of population continuity will be carried out with different values of sample sizes of the two populations. These tests will be done to see the effect of sample sizes on the results of continuity tests. For instance, different sample sizes will be considered between the samples of Sagalossos and two different modern populations. Lastly, modern samples from the southeast Anatolia will be employed to test the maternal continuity between Çemialo Sırtı and modern population of the region and Neolithic Syrian population to modern population of the region.

REFERENCES

Akurgal, E. (2003). Anadolu Uygarlıkları. Net Turistik Yayınlar AŞ.

Al-Zahery, N., Semino, O., Benuzzi, G., Magri, C., Passarino, G., Torroni, A., Santachiara-Benerecetti, A. S. (2003). Y-chromosome and mtDNA polymorphisms in Iraq, a crossroad of the early human dispersal and of post-Neolithic migrations.Molecular Phylogenetics and Evolution, 28(3):458-72.

Anderson, S., Bankier, T., Barrel, B.G., De Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, S., Schreier, P.H., Smith, A.J.H., Staden, R., Young, I.G. (1981). Sequence and organization of the human mitochondrial genome. Nature, 290: 457–465.

Bollongino, R., Nehlich, O., Richards, M.P., Orschiedt, J., Thomas, M.G., Sell, C., Fajkosova, Z., Powell, A., Burger, J. (2013). 2000 Years of Parallel Societies in Stone Age Central Europe. Science, 342 (6157): 479-481.

Bramanti, B., Thomas, M.G., Haak, W., Unterlaender, M., Jores, P., Tambets, K., Antanaitis-Jacobs, I., Haidle, M.N., Jankauskas, R., Kind, C.-J., Lueth, F., Terberger, T., Hiller, J., Matsumura, S., Forster, P., Burger, J. (2009). Genetic Discontinuity Between Local Hunter-Gatherers and Central Europe's First Farmers. Science, 326 (5949): 137-140.

Brandt, G., Haak, W., Adler, C.J., Roth, C., Szecsenyi-Nagy, A., Karimnia, S., Möller-Rieker, S., Meller, H., Ganslmeier, R., Friederich, S., Dresely, V., Nicklisch, N., Pickrell, J.K., Sirocko, F., Reich, D., Cooper, A., Alt, K.W., The Genographic Consortium. (2013). Ancient DNA Reveals Key Stages in the Formation of Central European Mitochondrial Genetic Diversity. Science, 342 (6155): 257-261.

Cinnioglu, C., King, R., Kivisild, T., Kalfolu, E., Atasoy, S., Cavalleri, G.L., Lillie, A.S., Roseman, C.C., Lin, A.A., Prince, K., Oefner, P.J., Shen, P., Semino, O., Cavalli-Sforza, L.L., Underhill, P.A. (2004). Excavating Y-chromosome haplotype strata in Anatolia. Human Genetics, 114 : 127–148.

Sacks, D., Murray, O., Lisa R. Brody R. L. (2005). Encyclopedia of the ancient Greek world. Infobase Publishing, pp. 256.

Del Pozzo, G., Guardiola, J. (1989). Mummy DNA Fragment Identified. Nature, 339, 431-432.

Der Sarkissian, C., Balanovsky, O., Brandt, G., Khartanovich, V., Buzhilova, A., Koshel, S., Zaporozhchenko, V., Gronenborn, D., Moiseyev, V., Kolpakov, E., Shumkin, V., Alt, K.W., Balanovska, E., Cooper, A., Haak, W., the Genographic

Consortium. (2013). Ancient DNA Reveals Prehistoric Gene-Flow from Siberia in the Complex Human Population History of North East Europe. PLOS Genetics, 9(2): e1003296.

Diakonoff, I. M. (1985), "Media", The Cambridge History of Iran 2 (Edited by Ilya Gershevitch ed.), Cambridge, England: Cambridge University Press, pp. 36–148.

Eglinton G, Logan GA. 1991. Molecular preservation. Philos. Trans. R. Soc. London Ser. B, 333:315–27; discussion 27–28.

Excoffier, L. and Lischer, H.E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10: 564-567.

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.

Fernandez, E., Perez-Perez, A., Gamba, C., Prats, E., Cuesta, P., Josep Anfruns, J., Molist, M., Arroyo-Pardo, E., Turbo´n, D. (2014). Ancient DNA Analysis of 8000 B.C. Near Eastern Farmers Supports an Early Neolithic Pioneer Maritime Colonization of Mainland Europe through Cyprus and the Aegean Islands. PLoS Genet, 10(6): e1004401.

Friedberg, EC., Walker, GC., Siede, W. (1995). DNA Repair and Mutagenesis. Washington, DC: ASM Press, 698 pp.

Fu, Q., Mittnik, A., Johnson P.L.F., Bos, K., Lari, M., Bollongino, R., Sun, C., Giemsch, L., Schmitz, R., Burger, J., Ronchitelli A.M., Martini, F., Cremonesi, R.G., rı´ Svoboda, J., Bauer, P., Caramelli, D., Castellano, S., Reich, D., Pa¨a¨bo, S., Krause, J. (2013). A Revised Timescale for Human Evolution Based on Ancient Mitochondrial Genomes. Current Biology 23, 553–559.

Ghezzi, D., Marelli, C., Achilli, A., Goldwurm, S., Pezzoli, G., Barone, P., Pellecchia, MT., Stanzione, P., Brusa, L., Bentivoglio, AR., Bonuccelli, U., Petrozzi, L., Abbruzzese, G., Marchese, R., Cortelli, P., Grimaldi, D., Martinelli, P., Ferrarese, C., Garavaglia, B., Sangiorgi, S., Carelli, V., Torroni, A., Albanese, A., Zeviani, M. (2005). Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. European Journal of Human Genetics, 13(6): 748–52.

Gilbert, M. Thomas P., Bandelt H-J., Hofreiter M., Barnes I. 2005. Assessing ancient DNA studies. TRENDS in Ecology and Evolution, 20(10).

Gonzalez, A., Larruga, J.M., Abu-Amero, K.K., Shi, Y., Pestano, J., Cabrera, V.M. (2007). Mitochondrial lineage M1 traces an early human backflow to Africa. BMC Genomics, 8: 223.

Green, P. (1990). Alexander to Actium, the historical evolution of the Hellenistic age. University of California Press, Pages 7-8.

Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., Fu, Q., Mittnik, A., Banffy, E., Economou, C., Francken, M., Friederich, S., Pena, R.G., Hallgren, F., Khartanovich, V., Khokhlov, A., Kunst, M., Kuznetsov, P., Meller H., Mochalov, O., Moiseyev, V., Nicklisch, N., Pichler, S.L., Risch, R., Guerra M.A.R., Roth, C., Sze´cse´nyi-Nagy, A., Wahl, J., Meyer, M., Krause, J., Brown, D., Anthony, D., Cooper, A., Alt, K.W., Reich, D. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. Nature, 522, 207-211.

Hansen, A., Willerslev, E., Wiuf, C., Mourier, T., Arctander, P. (2001). Statistical evidence for miscoding lesions in ancient DNA templates. Mol. Biol. Evol., 18:262–265.

Hervella, M., Izagirre, N., Alonso, S., Fregel, R., Alonso, A., Cabrera, V.M. (2012). Ancient DNA from Hunter-Gatherer and Farmer Groups from Northern Spain Supports a Random Dispersion Model for the Neolithic Expansion into Europe. PLOS One, 7(4):e34417.

Hewitt, G. (2004). Genetic Consequences of Climatic Oscillations in the Quaternary.Philosophical Transactions of the Royal Society of London Series Biological Sciences, 359 (1442): 183-195.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A., Wilson, A.C. (1984). DNA-sequences from the Quagga, an extinct member of the horse family. Nature, 312:282–284.

Hofreiter, M., Serre, D., Poinar, HN., Kuch, M., Pääbo, S. (2001). Ancient DNA. Nat Rev Genet., 2: 353–359.

İnalcık, K. (1997). Turkey and Europe: A historical Perspective. Perceptions, Journal of International Affairs. Volume II

Jobling, M.A., Tyler-Smith, C., Hurles, M. (2004). Human Evolutionary Genetics: Origins, Peoples and Disease. Garland Science.

Keyser, C., Bouakaze, C., Crubezy, E., Nikolaev, V.G., Montagnon, D., Reis, T., Ludes, B. (2009). Ancient DNA provides new insights into the history of south Siberian Kurgan people. Human Genetics, 126(3): 395-410.

Knapp, M., Lalueza-Fox, C., Hofreiter, M. (2015). Re-inventing ancient human DNA. Investigative Genetics, 6:4.

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, 894–897.

Lacan, M., Keyser, C., Ricaut, F.X., Brucato, N., Duranthon, F., Guilaine, J., Crubézy, E., Ludesa, B. (2011). Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. PNAS, USA. 108: 9788–9791.

Lalueza-Fox, C., Sampietro, M.L., Gilbert, M.T.P., Castri, L., Facchini, F., Pettener, D., Bertranpetit, J. (2004). Unravelling migrations in the steppe: mitochondrial DNA sequences from ancient Central Asians. Proceedings of the Royal Society B: Biological Sciences, 271(1542):941-7.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., Berger, B., Economou, C., Bollongino, R., Fu, Q., Bos, K.I., Nordenfelt, S., Li, H., de Filippo, C., Prüfer, K., Sawyer, S., Posth, C., Haak, W., Hallgren, F., Fornander, E., Rohland, N., Delsate, D., Francken, M., Guinet, JM.,Wahl, J., Ayodo, G., Babiker, HA., Bailliet, G., Balanovska, E., Balanovsky, O., Barrantes, R., Bedoya, G., Ben-Ami, H., Bene, J., Berrada, F., Bravi, CM., Brisighelli, F., Busby, GB., Cali, F., Churnosov, M., Cole, DE., Corach, D., Damba, L., van Driem, G., Dryomov, S., Dugoujon, JM., Fedorova, SA., Gallego Romero, I., Gubina, M., Hammer, M., Henn, BM., Hervig, T., Hodoglugil, U., Jha, AR., Karachanak-Yankova, S., Khusainova, R., Khusnutdinova, E., Kittles, R., Kivisild, T., Klitz, W., Kučinskas, V., Kushniarevich, A., Laredj, L., Litvinov, S., Loukidis, T., Mahley, RW., Melegh, B., Metspalu, E., Molina, J., Mountain, J., Näkkäläjärvi, K., Nesheva, D., Nyambo, T., Osipova, L., Parik, J., Platonov, F., Posukh, O., Romano, V., Rothhammer, F., Rudan, I., Ruizbakiev, R., Sahakyan, H., Sajantila, A., Salas, A., Starikovskaya, EB., Tarekegn, A., Toncheva, D., Turdikulova, S., Uktveryte, I., Utevska, O., Vasquez, R., Villena, M., Voevoda, M., Winkler CA., Yepiskoposyan, L., Zalloua, P., Zemunik, T., Cooper, A., Capelli, C., Thomas, MG., Ruiz-Linares, A., Tishkoff, SA., Singh, L., Thangaraj, K., Villems, R., Comas, D., Sukernik, R., Metspalu, M., Meyer, M., Eichler, EE., Burger, J., Slatkin, M., Pääbo, S., Kelso, J., Reich, D., Krause, J. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature, 513,409–413.

Lewis, B. (1995). The Middle East: A brief history of the last 2000 years. A Touchstone book, New York.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature, 362:709–15.

Long, J.C. (1991). The genetic structure of admixed populations.Genetics, 127: 417 – 428.

Longo, M.C., Berninger, M.S., Hartley, J.L. (1990). Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. Gene., 93:125-128.

Maji, S., Krithika, S., Vasulu, T.S. (2008). Distribution of mitochondrial DNA macrohaplogroup N in India with special reference to haplogroup R and its sub-haplogroup U. Int J Hum Genet., 8:85–96.

Malhi and Smith. (2002). Brief Communication: Haplogroup X Confirmed in Prehistoric North America, American Journal of Physical Anthropology, 119(1), Pages 84-86.

Malmström, H., Vretemark, M., Tillmar, A., Durling, M.B., Skoglund, P., Gilbert, M.T.P., Willerslev, E., Holmlund, G., Götherström, A. (2012). Finding the founder of Stockholm – A kinship study based on Y-chromosomal, autosomal and mitochondrial DNA. Annals of Anatomy - Anatomischer Anzeiger, 194(1), 138–145.

Malyarchuk, B., Grzybowski, T., Derenko, M., Perkova, M., Vanecek, T., Lazur, J., Gomolcak, P., Tsybovsky, I. (2008). Mitochondrial DNA Phylogeny in Eastern and Western. Mol. Biol. Evol., 25(8):1651–1658.

Matney, T., Algaze G., Dulik, M. C., Erdal, Ö. D., Erdal, Y. S., Gokcumen, O., Lorenz, J., Mergen, H. (2012). Understanding Early Bronze Age Social Structure Through Mortuary Remains: A Pilot aDNA Study From Titris¸ Höyük, Southeastern Turkey. International Journal of Osteoarchaeology, 22:338–351.

McLean, B.H. (2002). An introduction to Greek epigraphy of the Hellenistic and Roman periods from Alexander the Great down to the reign of Constantine (323 BC-AD 337). University of Michigan Press.

Metspalu, M., Kivisild, T., Metspalu, E., Parik, J., Hudjashov, G., Kaldma, K., Serk, P., Karmin, M., Behar, D.M., Gilbert, M.T.P., Endicott, P., Mastana, S., Papiha, S.S., Skorecki, K., Torroni, A., Villems, R. (2004). Most of the extant mtDNA boundaries in South and Southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. BMC Genet., 5, 26.

Meyer, M., Kircher, M., Gansauge, MT., Li, H., Racimo, F., Mallick, S., Schraiber, JG., Jay, F., Prüfer, K., de Filippo, C., Sudman,t PH., Alkan, C., Fu, Q., Do, R., Rohland, N., Tandon, A., Siebauer, M., Green, RE., Bryc, K., Briggs, AW., Stenzel, U., Dabney, J., Shendure, J., Kitzman, J., Hammer, MF., Shunkov, MV., Derevianko, AP., Patterson, N., Andrés, AM., Eichler, EE., Slatkin, M., Reich, D., Kelso, J., Pääbo, S. (2012). A high-coverage genome sequence from an archaic Denisovan individual. Science, 338(6104):222-6.

Olivieri, C., Ermini, L., Rizzi, E., Corti, G., Bonnal, R., Luciani, S., Marota, I., De Bellis, G., Rollo, F. (2010). Characterization of Nucleotide Misincorporation Patterns in the Iceman's Mitochondrial DNA. PLoS ONE, 5(1): e8629.

Ottoni, C., Ricaut, FX., Vanderheyden, N., Brucato, N., Waelkens, M., Decorte, R. (2011). Mitochondrial analysis of a Byzantine population reveals the differential impact of multiple historical events in South Anatolia. Eur J Hum Genet., 19: 571–576.

Pääbo, S., Poinar, H.,Serre, D., Jaenicke-Despr´es, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., Hofreiter, M. (2004). Genetic analyses from ancient DNA, Annu. Rev. Genet., 38:645–79.

Pääbo, S. (2014). The Human Condition-A Molecular Approach. Cell, 157.

Pääbo, S. (1985). Molecular cloning of Ancient Egyptian mummy DNA. Nature, 314, 644 – 645.

Pääbo, S. (1989). Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci USA, 86: 1939-1943.

Pedersen, J. S., Valen, E., Vargas Velazquez, A.M., Parker, B.J., Rasmussen, M., Lindgreen, S., Lilje, B., Tobin, D.J., Kelly, T.K., Vang, S., Andersson, R., Jones, P.A., Hoover, C.A.,
Tikhonov, A., Prokhortchouk, E., Rubin, E.M., Sandelin, A., Gilbert, M.T.P., Krogh, A.,
Willerslev, E., Orlando, L. (2013). Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. Genome Res., 24:454–466.

Cherni, L., Fernandes, V., Pereira, JB., Costa, MD., Goios, A., Frigi, S., Yacoubi-Loueslati, B., Amor, MB., Slama, A., Amorim, A., El Gaaied, AB., Pereira, L. (2009). Post-last glacial maximum expansion from Iberia to North Africa revealed by fine characterization of mtDNA H haplogroup in Tunisia. American Journal of Physical Anthropology, 139 (2): 253–60.

Pruvost, M., Grange, T., Geigl, E-M. (2005). Minimizing DNA contamination by using UNG-coupled quantitative real-time PCR on degraded DNA samples: application to ancient DNA studies. Biotechniques, 38:569-575.

Quintana-Murci, L., Chaix, R., Wells, RS., Behar, DM., Sayar, H., Scozzari, R., Rengo, C., Al-Zahery, N., Semino, O., Santachiara-Benerecetti, AS., Coppa, A., Ayub, Q., Mohyuddin, A., Tyler-Smith, C., Qasim Mehdi, S., Torroni, A., McElreavey, K. (2004). Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. Am J Hum Genet., 74, pp. 827–845.

Richards, M., Macaulay,V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., Villems, R., Thomas, M., Rychkov, S., Rychkov, O., Rychkov, Y., Gölge, M., Dimitrov, D., Hill, E., Bradley, D., Romano, V., Calı`, F., Vona, G., Demaine, A., Papiha, S., Triantaphyllidis, C., Stefanescu, G., Hatina, J., Belledi, M., Di Rienzo, A., Novelletto, A., Oppenheim, A., Nørby, S., Al-Zaheri, N., Santachiara-Benerecetti, S., Scozzari, R., Torroni, A., Bandelt, H. (2000). Tracing European Founder Lineages in the Near Eastern mtDNA Pool. American Journal of Human Genetics, 67: 1251-1276.

Rivollat, M., Mendisco, F., Pemonge, M-H., Safi, A., Saint-Marc, D., Bremond, A., Couture-Veschambre, C., Rottier, S., Deguilloux, M-F. (2015). When the Waves of European Neolithization Met: First Paleogenetic Evidence from Early Farmers in the Southern Paris Basin. PLOS One, 10(4): e0125521.

Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., Reich, D. (2015). Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. Philosophical Transactions B 370(1660).

Rossabi, M. (1994). The Legacy of the Mongols in Manz, B.F (ed) Central Asia in Historical perspective. Boulder, Colorado. Westview Press.

Roux, J-P. (1997). Türklerin Tarihi: Büyük Okyanus'tan Akdeniz'e İki Bin Yıl. Milliyet Yayınları.

Salman, H. (2004). "Türk adı, Türklerin ana yurdu ve göçleri" in Öztürk, C (ed) Türk Tarihi ve kültürü. Pegem A Yayıncılık.

Sampson, G.C. (2008). The Defeat of Rome: Crassus, Carrhae and the Invasion of the East. Cyrus the Great, founder of the First Persian Empire (c. 550–330 BCE). Pen & Sword Books Limited, p. 33. ISBN 9781844156764.

Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., Pääbo, S. (2012). Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PLoS ONE, 7(3), e34131.

Schönberg, A., Theunert, C., Li, M.,BStoneking, M., Nasidze, I. (2011). High-throughput sequencing of complete human mtDNA genomes from the Caucasus and West Asia: high diversity and demographic inferences. European Journal of Human Genetics, 19, 988–994.

Sverrisdóttir, OÓ., Timpson, A., Toombs, J., Lecoeur, C., Froguel, P., Carretero, JM., Arsuaga Ferreras, JL., Götherström, A., Thomas, MG. (2014). Direct estimates of natural selection in Iberia indicate calcium absorption was not the only driver of lactase persistence in Europe. Mol. Biol. Evol., 31,975–983.

Sykes, B. (2001). The Seven Daughters of Eve. London; New York: Bantam Press, ISBN 0393020185.

Szécsényi-Nagy, A., Brandt, G., Haak, W., Keerl, V., Jakucs, J., Möller-Rieker, S., Köhler, K., Mende, BG., Oross, K., Marton, T., Osztás, A., Kiss, V., Fecher, M., Pálfi, G., Molnár, E., Sebők, K., Czene, A., Paluch, T., Šlaus, M., Novak, M., Pećina-Šlaus, N., Ősz, B., Voicsek, V., Somogyi, K., Tóth, G., Kromer, B., Bánffy, E., Alt, KW. (2015). Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. Proc. R. Soc. B., 282: 20150339.

Tambets, K., Kivisild, T., Metspalu, E., Parik, J., Kaldma, K., Laos, S., Tolk, H.V., Golge, M., Demirtas, H., Geberhiwot, T., Papiha, S.S., de Stefano, G.F., Villems, R., (2000). The topology of the maternal lineages of the Anatolian and Trans-Caucasus populations and the peopling of Europe: some preliminary considerations. In: Renfrew, C., Boyle, K. (Eds.), Archaeogenetics: DNA and the Population Prehistory of Europe. Cambridge, pp. 219–235.

Tavernier, J. (2007). Iranica in the Achaemenid Period (ca. 550-330 B.C.). Lexicon of Old Iranian Proper Names and Loanwords, Attested in Non-Iranian Texts. Leuven â Paris â Dudley.

Underhill, P.A., Passarino, G., Lin, A.A., Shen, P., Lahr, M.M., Foley, R.A., Oefner, P.J., Cavalli-Sforza, L.L. (2001). The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations.Annals of Human Genetics, 65: 43-62.

Van Oven, M., Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat., 30, E386–E394.

Varshney, U., van de Sande, JH. (1991). Specificities and kinetics of uracil excision from uracil-containing DNA oligomers by Escherichia coli uracil DNA glycosylase. Biochemistry, 30, 4055–4061.

Vryonis, S. (1971). The Decline of medieval Hellenism in Asia Minor and the process of Islamization from the eleventh through the fifteenth century. University of California press.

APPENDICES

APPENDIX A: mtDNA HVRI-HVRII Sequences and Haplogroups of Research Members

>İnci_HVRI
TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCC
TCAACTATCACGCATCAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGATA
CCAACAAACCTACTCACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTA
CATAGCACATTACAGTCAAATCCCTTCTCGTCCCC

>Füsun_HVRI
TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCC
TCAACTATCACACATCAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGATA
CCAACAAACCTACCCACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTA
CATAGCACATTACAGTCAAATCTCTTCTCGTCCCC

>Eren_HVRI
TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCCCCCCATGCTTACAAGCAAGTACAGCAATCAACCT
TCAACTATCACACATCAACTGCAATTCCAAAGCCACCCCTCACCCACTAGGATAT
CAACAAACCTACCCACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTAC
ATAGCACATTACAGTCAAATCCCTTCTCGTCCCC

>Dilşad_HVRI

TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGCACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCC
TCAACTATCACACATCAACTGCAACTCCAAAGCCACCTCTCACCCACTAGGATA
CCAACAAACCTACCCATCTTTAACAGTACATAGTACATAAAGCCATTTACCGTAC
ATAGCACATTACAGTCAAATCCCTTCTCGTCCCC

>Reyhan_HVRI

TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGTACAGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCTCCTCATGCTTACAAGCAAGTACAGCAATCAACCC
TCAACTATCACACATCAACTGCAACTCCAAAGTCACCCCTCACCCATTAGGATAC
CAACAAACCTACTCACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTAC
ATAGCACATTACAGTCAAATCCCTTCTCGTCCCC

>Yasemin_HVRI

TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCCCCCCATGCTTACAAGCAAGTACAGCAATCAACCC
TCAACTATCACACATCAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGATA
CCAACAAACCTACCCACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTA
CATAGCACATTACAGTCAAATCCCTTCTCGCCCCC

>Sidar_HVRI

TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCC
TCAACTATCACACATCAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGATA
CCAACAAACCTACCCACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTA
CATAGCACATTACAGTCAAATCCCTTCTCGCCCCC

>Aliye_HVRI

TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCC
TCAACTATCACACATCAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGATA
CCAACAAACCTACCCACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTA
CATAGCACATTACAGTCAAATCCCTTCTCGCCCCC

>Esra_HVRI

TTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAG
TATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC
ATGAATATTGTACGGTACCATAAATACTTGACTACCTGTAGTACATAAAAACCC
AATCCACATCAAAACCCCCCCCCCATGCTTACAAGCAAGTACAGCAATCAACCC
TCAACTATCACACATCAACTGCAACTCCAAAGCCACCCCTCACCCATTAGGATA
CCAACAAACCTACCCACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTA
CATAGCACATTACAGTCAAATCCCTTCTCGCCCCC

>İnci_HVRII

ATTTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTGTTAATT
AATTAATGCTTGTAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCT

>Füsun_HVRII

ATTTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTGTTAATT
AATTAATGCTTGTAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCT

>Eren_HVRII

ATTTGGTATTTTCGTCTGGGGGGTGTGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATCCTGTTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACCTACTAAAGTGTGTTAATT
AATTAATGCTTATAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCT

>Dilsad_HVRII

ATTTGGTATTTTCGTCTGGGGGGTGTGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTGTTAATT
AATTAATGCTTGTAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCT

>Reyhan_HVRII

ATTTGGTATTTTCGTCTGGGGGGTGTGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATTCTATTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTGTTAATT
AATTAATGCTTGTAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCT

>Yasemin_HVRII

ATTTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTGTTAATT
AATTAATGCTTGTAGGACATAGTAATAACAATTGAATGTCTGCACAGCCGCT

>Sidar_HVRII

ATTTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATCCCATTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTGTTAATT
AATTAATGCTTGTAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCT

>Aliye_HVRII

ATTTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATCCCATTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTGTTAATT
AATTAATGCTTGTAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCT

>Esra_HVRII

ATTTGGTATTTTCGTCTGGGGGGTGTGCACGCGATAGCATTGCGAGACGCTGGA
GCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATTCCATTATT
TATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTATTAATT
AATTAATGCTTGTAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCT

Individual	**Haplogroup**	**HVRI Mutation Motifs**	**HVRII Mutation Motifs**
İnci	H2a2b	A16235G - C16291T	A263G
Füsun	H2a1	C16354T	A263G
Eren	X2b1	T16189C - C16223T - C16248T - C16278T	A73G - A153G - T195C G225A - A263G
Dilşad	T2	T16126C - C16261T - C16294T - C16296T	A73G - A263G
Reyhan	U5a1b1c2	G16129A - C16192T - C16256T - C16270T - C16291T	A73G - C150T - A263G
Yasemin	H1c3b	T16189C - T16362C	A235G - A263G
Sidar	R0b	T16362C	T152C - A263G
Aliye	R0b	T16362C	T152C - A263G
Esra	D5b1c1	C16148T - T16189C - C16270T - T16362C	A73G - C150T - T152 - G207A - A263G

APPENDIX B: Screenshots of Samples SK24, SK26, SK28, SK7, SK17, SK21, SK35 and Their Mutation Motifs

SK24

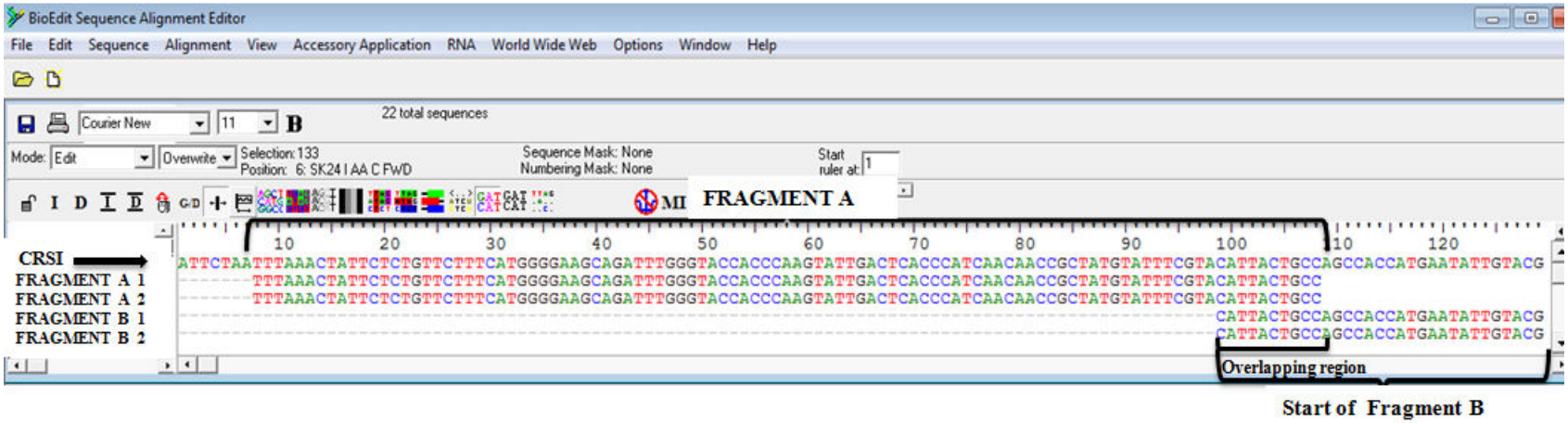


Figure 1 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment A which is the first fragment of HVI region and 140 bp in length. They are aligned with CRSI. A short part of fragment A which is 10 bp in length is overlapping with fragment B. Two sequences of fragment A belonged to two different extractions from the two different teeth samples of individual SK24 and one PCR from each extraction.

For this individual, in this fragment there was no mutation motif in relation to haplogroup determination in reference to CRSI. Once more, identical sequences observed for the overlapping parts of the fragments A and B indicate that, with a high probability, these two sequences belong to the same individual.

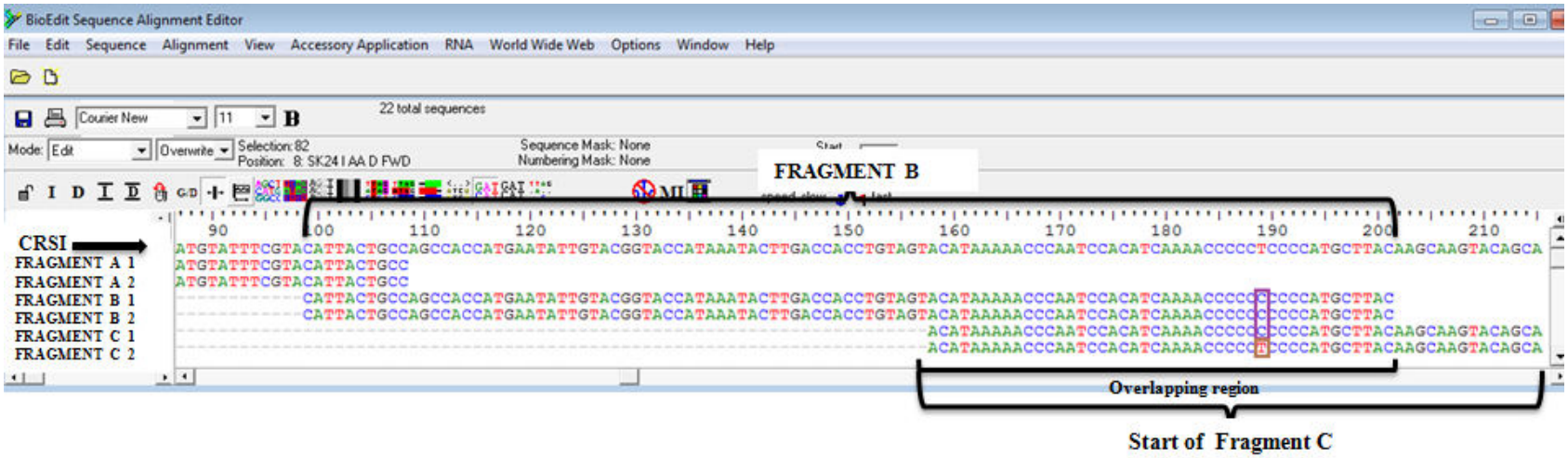


Figure 2 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment B which is the second fragment of HVI region and 139 bp in length. They are aligned with CRSI. A short part of fragment B which is 44 bp in length is overlapping with fragment C. Two sequences of fragment B belonged to two different extractions from the two different teeth samples of individual SK24 and one PCR from each extraction. The purple box indicates mutation motif that is from T to C at 189np in reference to CRSI. This np is actually 16189th position in the given CRS when the sequence of whole mtDNA is used.

The mutation motif (indicated by the purple boxes) observed on the 189 np in reference to CRSI in Figure 2 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments B and C indicate that, with a high probability, these two sequences belong to the same individual.

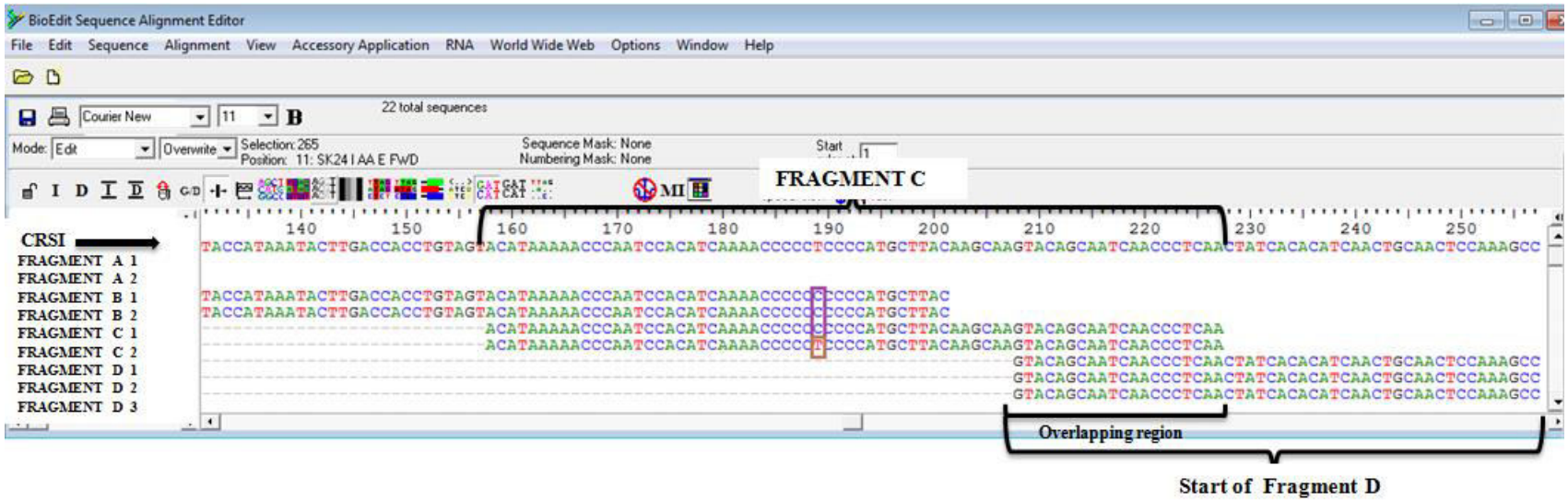


Figure 3 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment C which is the third fragment of HVI region and 109 bp in length. They are aligned with CRSI. A short part of fragment C which is 20 bp in length is overlapping with fragment D. Two sequences of fragment C belonged to two different extractions from the two different teeth samples of individual SK24 and one PCR from each extraction. The purple box indicates mutation motif that is from T to C at189 np in reference to CRSI. This np is actually 16189th position in the given CRS when the sequence of whole mtDNA is used. There were the "C" nucleotide on the 189th position of FRAGMENT B 1, FRAGMENT B 2 and also FRAGMENT C 1, while it was "T" on the same position for FRAGMENT C 2, sequence (second C fragment sequences of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type2). This nucleotide misincorporation is located at the overlapping region of fragment B and C.

The mutation motif (indicated by the purple box) observed on the 189 np in reference to CRSI in Figure 3 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments C and D indicate that, with a high probability, these two sequences belong to the same individual.

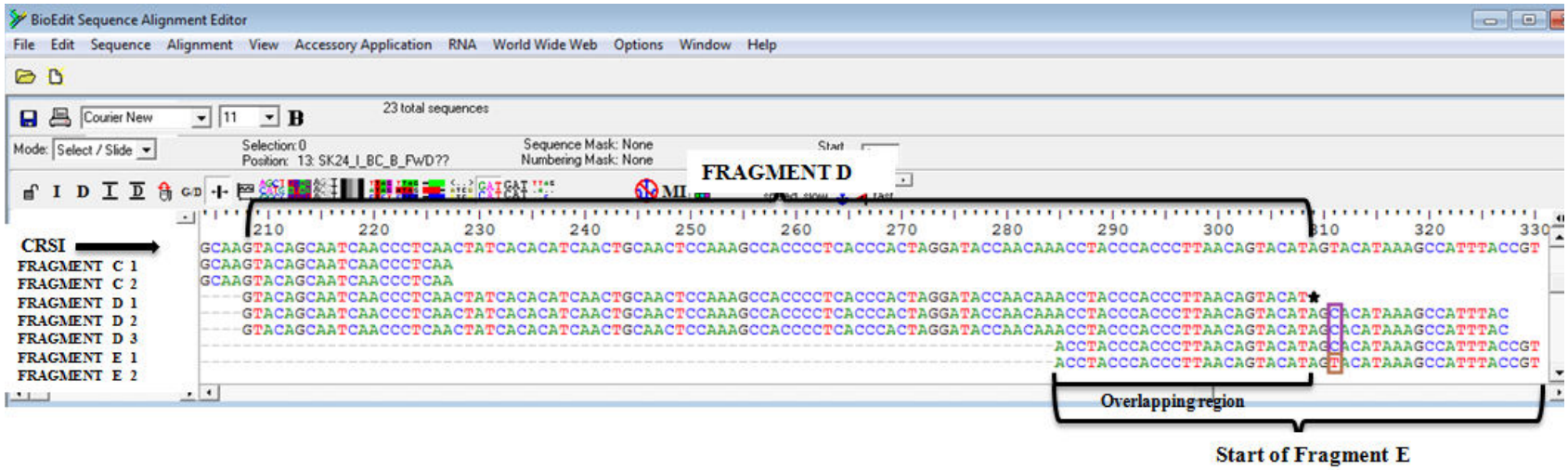


Figure 4 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment D which is the fourth fragment of HVI region and 133 bp in length. They are aligned with CRSI. A short part of fragment D which is 24 bp in length is overlapping with fragment E. Three sequences of fragment D belonged to two different extractions from the two different teeth samples of individual SK24 and two independent PCRs from one extraction, one PCR from another extraction. The purple box indicates mutation motif that is from T to C at 311 np in reference to CRSI. This np is actually 16311th position in the given CRS when the sequence of whole mtDNA is used. Normally fragment D ends at the 309 np which was shown with the black star. However two repetitions of the fragment D sequences were able to be read 19 additionally nucleotide, which were on the reverse primer binding region of the sequences. Due to the limited number of repetitions of E fragment sequences, in the present study, the five D fragment sequences were examined up to the 327 np (3' end of D fragment was extended 19 bp). This cases will be explained in the Discussion Chapter. The mutation motif at the overlapping part of fragments D and E was observed on the first E fragment sequences.

The mutation motif (indicated by the purple box) observed on the 311 np in reference to CRSI in Figure 4 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these three sequences belong to the same individual.

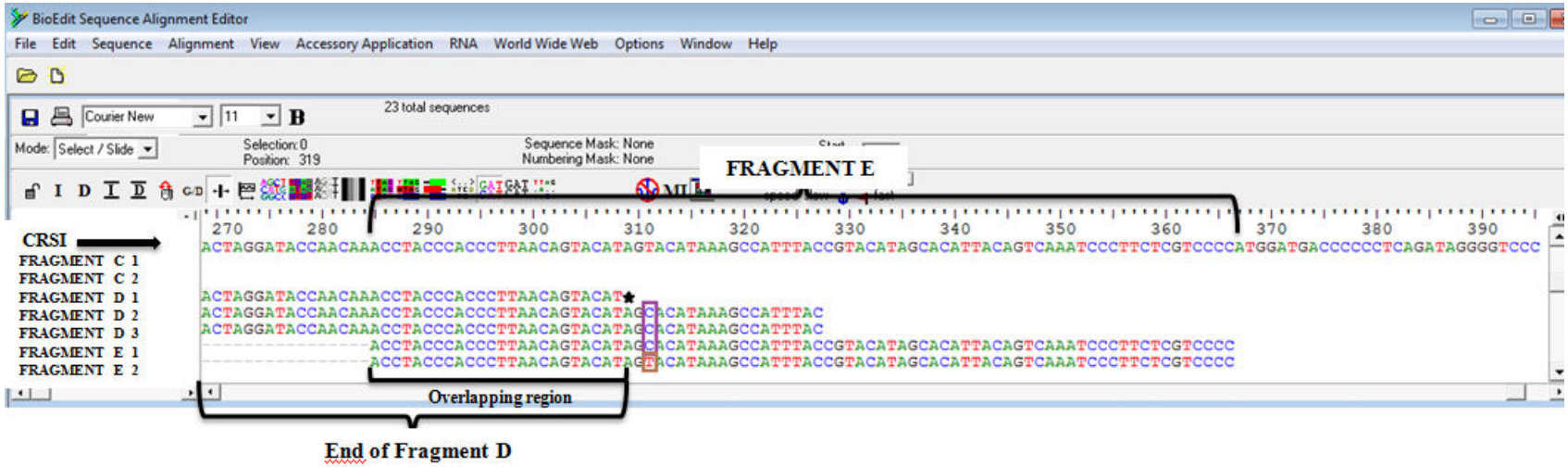


Figure 5 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment E which is the last fragment of HVI region and 118 bp in length. They are aligned with CRSI. A short part of fragment E which is 24 bp in length is overlapping with fragment D. Two sequences of fragment E belonged to two different extractions from the two different teeth samples of individual SK24 and one PCR from each extraction. The purple box indicates mutation motif that is from T to C at 311 np in reference to CRSI. This np is actually 16311th position in the given CRS when the sequence of whole mtDNA is used. As it was explained in Figure 4 two sequences of fragment D were extended 19 bp more to check the accuracy of the sequences using the mutation motifs. One of the E fragment sequences had the same mutation at that position. There were the "T" nucleotide on the 311th position of FRAGMENT D 2, FRAGMENT D 3 and FRAGMENT E 1 sequences while it was "T" on the same position for FRAGMENT E 2 sequence (second E fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type2). This nucleotide misincorporation is located at the overlapping region of fragment D and E.

The mutation motif (indicated by the purple box) observed on the 311 np in reference to CRSI in Figure 5 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments E and D indicate that, with a high probability, these two sequences belong to the same individual.

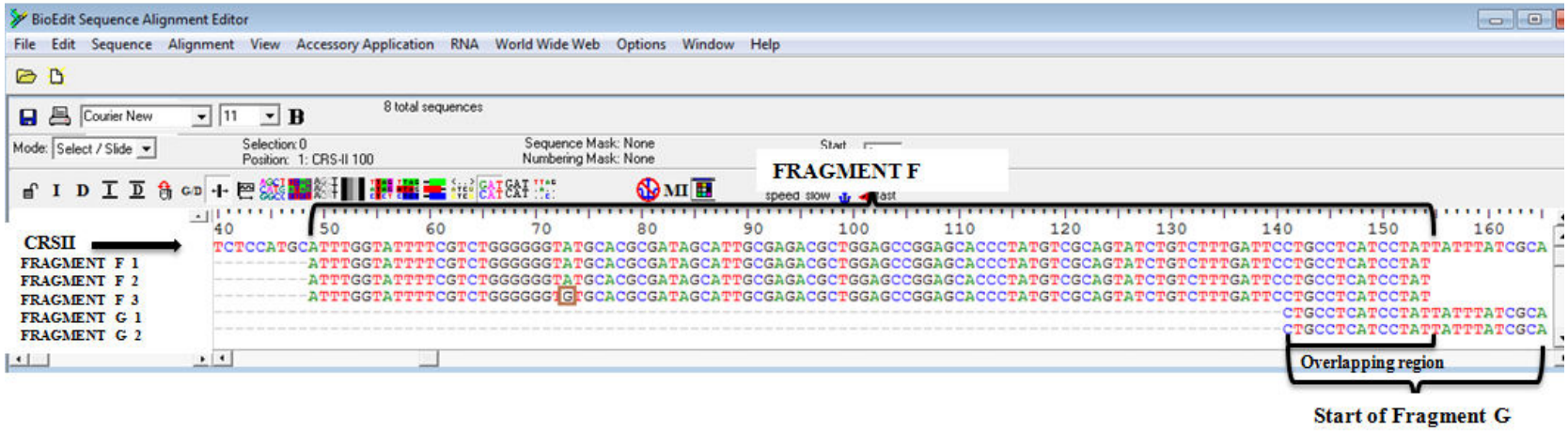


Figure 6 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment F which is the first fragment of HVII region and 147 bp in length. They are aligned with CRSII. A short part of fragment F which is 14 bp in length is overlapping with fragment G. Three sequences of fragment D belonged to two different extractions from the two different teeth samples of individual SK24 and two independent PCRs from one extraction, one PCR from another extraction. For this individual, in this fragment there was no mutation motif in relation to haplogroup determination in reference to CRSII. However, there were the "A" nucleotide on the 73rd position of FRAGMENT F 1 and FRAGMENT F 2, while it was "G" on the same position for FRAGMENT F 3 sequence (third F fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type1). This nucleotide misincorporation is located at the overlapping region of fragment F and G.

Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these three sequences belong to the same individual.

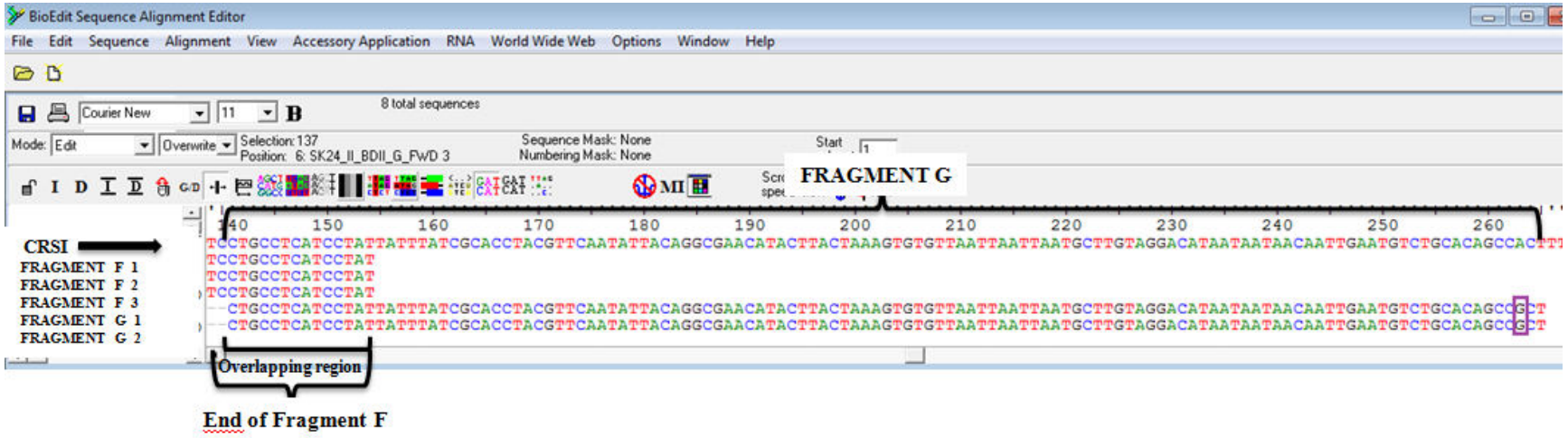


Figure 7 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment G which is the last fragment of HVII region and 166 bp in length. They are aligned with CRSII. A short part of fragment G which is 14 bp in length is overlapping with fragment F. Two sequences of fragment E belonged to two different extractions from the two different teeth samples of individual SK24 and one PCR from each extraction. The purple box indicates mutation motif that is from A to G at 263 np in reference to CRSII. This np are actually 263rd positions in the given CRS when the sequence of whole mtDNA is used. Both of the G fragment sequences had the same A to G mutations at this position.

The mutation motif (indicated by the purple box) observed on the 263 np in reference to CRSII in Figure 7 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments G and F indicate that, with a high probability, these two sequences belong to the same individual.

SK26

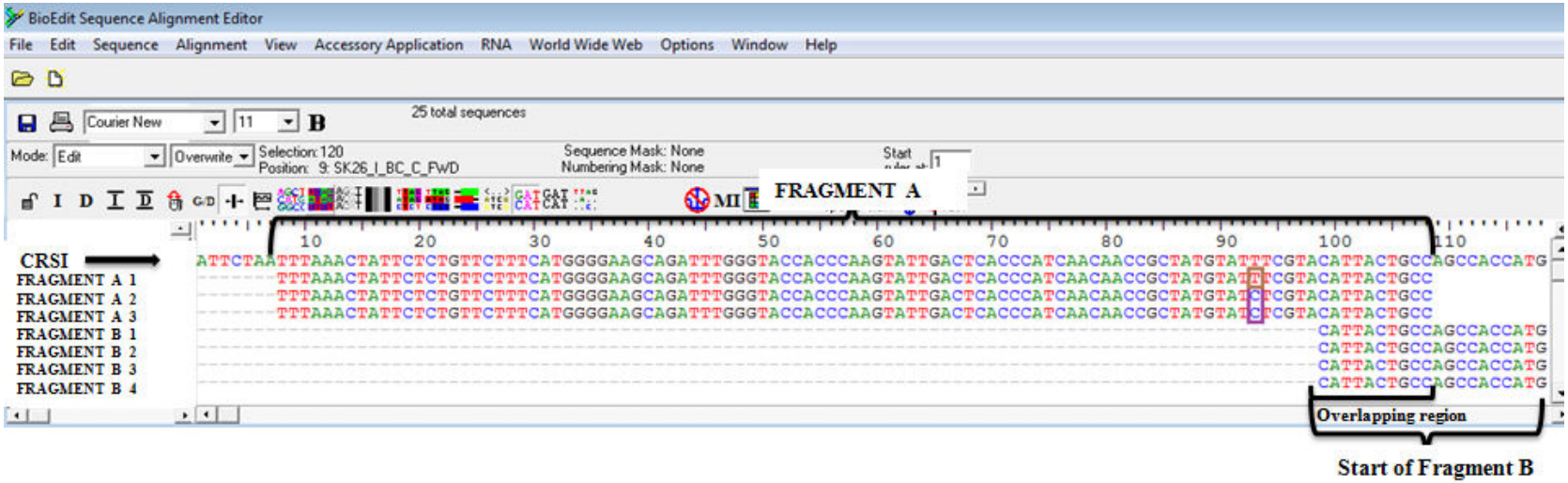


Figure 8 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment A which is the first fragment of HVI region and 140 bp in length. They are aligned with CRSI. A short part of fragment A which is 10 bp in length is overlapping with fragment B. Three sequences of fragment A belonged to two different extractions from the two different teeth samples of individual SK26 and two independent PCRs from one extraction, one PCR from another extraction. The purple box indicates mutation motif that is from T to C at 93 np in reference to CRSI. This np are actually 16093rd positions in the given CRS when the sequence of whole mtDNA is used. There were the "C" nucleotide on the 93rd position of FRAGMENT A 2 and FRAGMENT A 3, while it was "T" on the same position for FRAGMENT A 1 sequence (first A fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type2).

The mutation motif (indicated by the purple box) observed on the 93 np in reference to CRSII in Figure 8 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments A and B indicate that, with a high probability, these three sequences belong to the same individual.

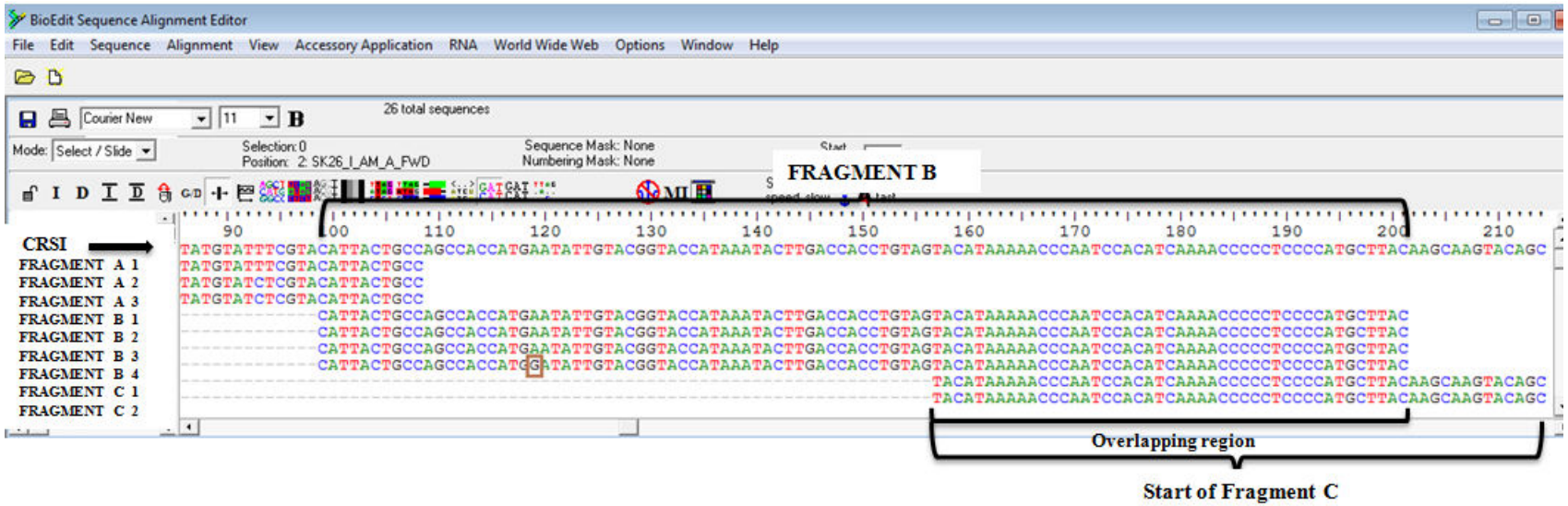


Figure 9 A screenshot from Bioedit 5.0.9 program shows the sequences (n=4) of fragment B which is the second fragment of HVI region and 139 bp in length. They are aligned with CRSI. A short part of fragment B which is 44 bp in length is overlapping with fragment C. Four sequences of fragment B belonged to three different extractions from the two different teeth samples of individual SK26 and two independent PCRs from one extraction, one PCR from each of other two extractions. For this individual, in this fragment there was no mutation motif in relation to haplogroup determination in reference to CRSI. However, there were the "A" nucleotide on the 119th position of FRAGMENT B 1, FRAGMENT B 2 and FRAGMENT B 3, while it was "G" on the same position for FRAGMENT B 4 sequence (fourth B fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type1).

Once more, identical sequences observed for the overlapping parts of the fragments B and C indicate that, with a high probability, these four sequences belong to the same individual.

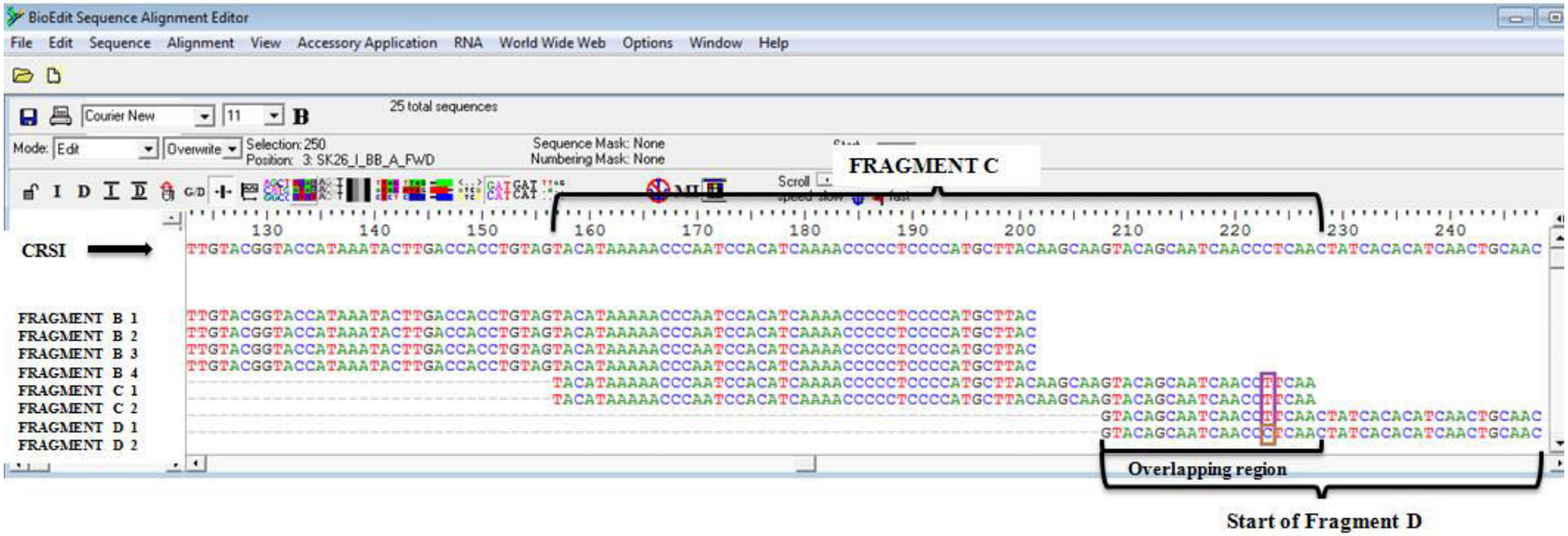


Figure 10 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment C which is the third fragment of HVI region and 109 bp in length. They are aligned with CRSI. A short part of fragment C which is 20 bp in length is overlapping with fragment D. Two sequences of fragment C belonged to two different extractions from the two different teeth samples of individual SK26 and one PCR from each extraction. The purple box indicates mutation motif that is from C to T at 223 np in reference to CRSI. This np is actually 16223rd position in the given CRS when the sequence of whole mtDNA is used. This mutation motif is located at the overlapping part and both of the C fragment sequences had the same mutations at that position.

The mutation motif (indicated by the purple box) observed on the 223 np in reference to CRSI in Figure 10 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments C and D indicate that, with a high probability, these two sequences belong to the same individual.

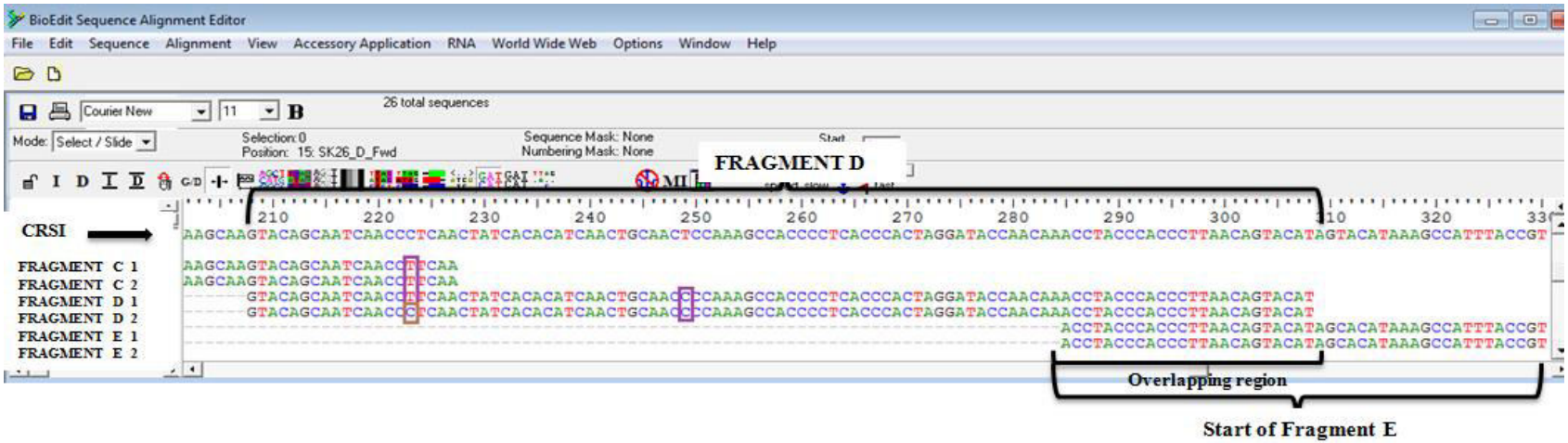


Figure 11 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment D which is the fourth fragment of HVI region and 133 bp in length. They are aligned with CRSI. A short part of fragment D which is 24 bp in length is overlapping with fragment E. Two sequences of fragment D belonged to two different extractions from the two different teeth samples of individual SK26 and one PCR from each extraction. The purple boxes indicate mutation motifs that are from C to T at 223 np and T to C at 249 np in reference to CRSI. These np are actually 16223rd and 16249th positions in the given CRS when the sequence of whole mtDNA is used. C to T mutation motif is located at the overlapping part of fragment C and D. Also one of the D fragment sequences had the same mutation at 223rd position while both of the D fragment sequences had the same T to C mutations at 249th position. There were the "T" nucleotide on the 223rd position of FRAGMENT C 1, FRAGMENT C 2 and FRAGMENT D 1, while it was "C" on the same position for FRAGMENT D 2 sequence (second D fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type1). This nucleotide misincorporation is located at the overlapping region of fragment C and D.

The mutation motifs (indicated by the purple boxes) observed on the 223 np and 249 np in reference to CRSI in Figure 11 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these two sequences belong to the same individual.

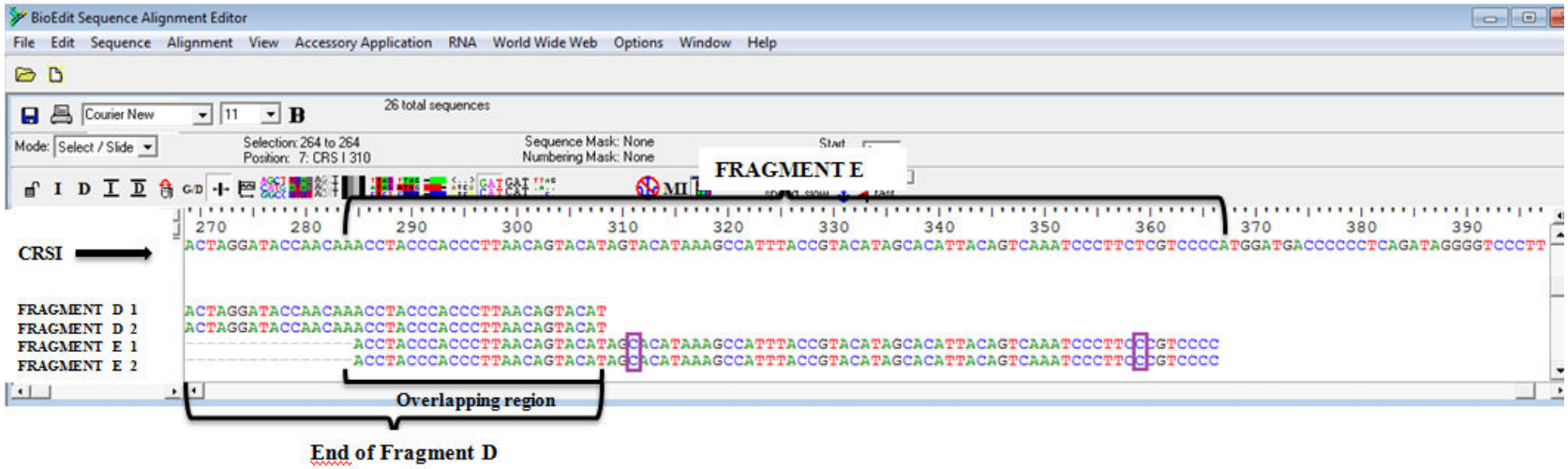


Figure 12 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment E which is the last fragment of HVI region and 118 bp in length. They are aligned with CRSI. A short part of fragment E which is 24 bp in length is overlapping with fragment D. Two sequences of fragment E belonged to two different extractions from the two different teeth samples of individual SK26 and one PCR from each extraction. The purple boxes indicate mutation motifs that are from T to C at 311 np and at 359 np in reference to CRSI. These np are actually 16311th and 16359th positions in the given CRS when the sequence of whole mtDNA is used. Both of the E fragment sequences had the same mutations at that positions.

The mutation motifs (indicated by the purple boxes) observed on the 311 np and 359 np in reference to CRSI in Figure 12 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these two sequences belong to the same individual.

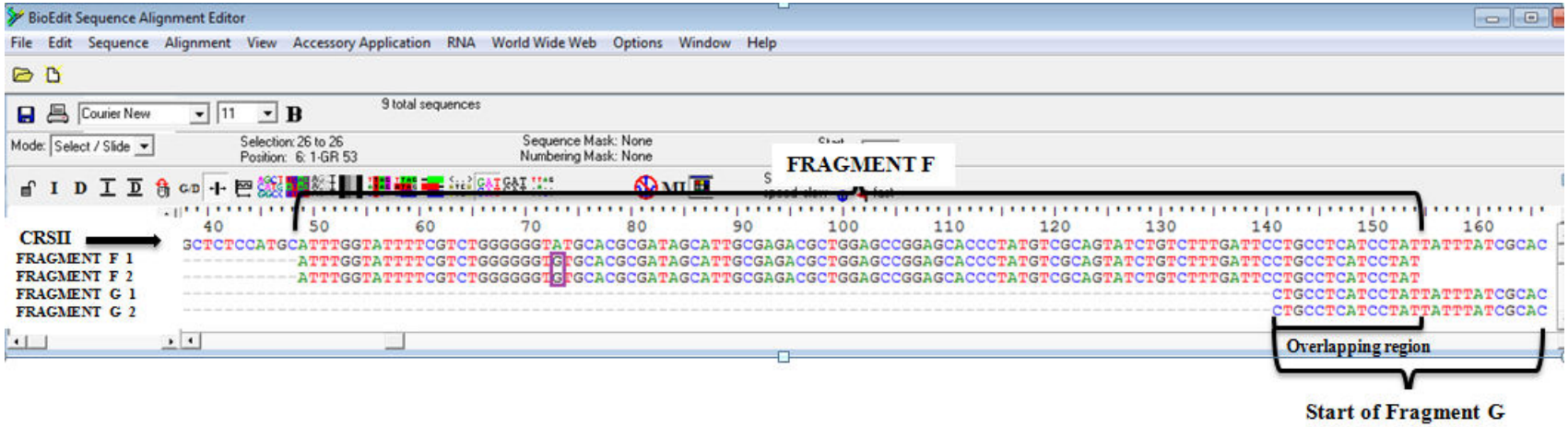


Figure 13 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment F which is the first fragment of HVII region and 147 bp in length. They are aligned with CRSII. A short part of fragment F which is 14 bp in length is overlapping with fragment G. Two sequences of fragment F belonged to two different extractions from the two different teeth samples of individual SK26 and one PCR from each extraction. The purple box indicates mutation motif that is from A to G at 73 np in reference to CRSII. This np are actually 73rd positions in the given CRS when the sequence of whole mtDNA is used. Both of the F fragment sequences had the same A to G mutations at this position.

The mutation motif (indicated by the purple box) observed on the 73 np in reference to CRSII in Figure 13 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

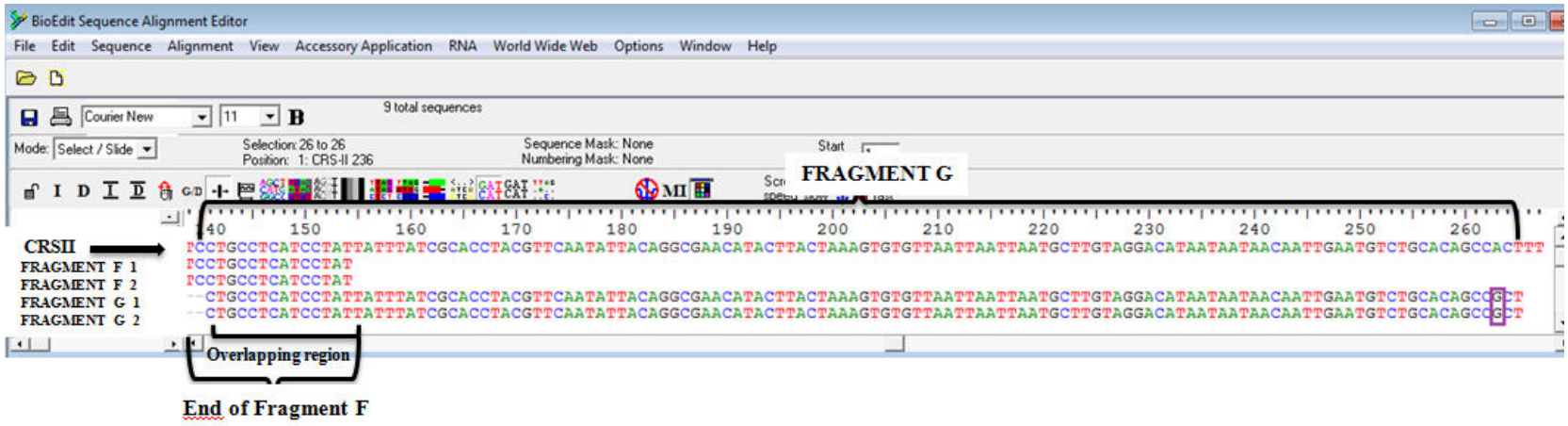


Figure 14 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment G which is the last fragment of HVII region and 166 bp in length. They are aligned with CRSII. A short part of fragment G which is 14 bp in length is overlapping with fragment F. Two sequences of fragment G belonged to two different extractions from the two different teeth samples of individual SK26 and one PCR from each extraction. The purple box indicates mutation motif that is from A to G at 263 np in reference to CRSII. This np are actually 263rd positions in the given CRS when the sequence of whole mtDNA is used. Both of the G fragment sequences had the same A to G mutations at this position.

The mutation motif (indicated by the purple box) observed on the 263 np in reference to CRSII in Figure 14 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments G and F indicate that, with a high probability, these two sequences belong to the same individual.

SK28

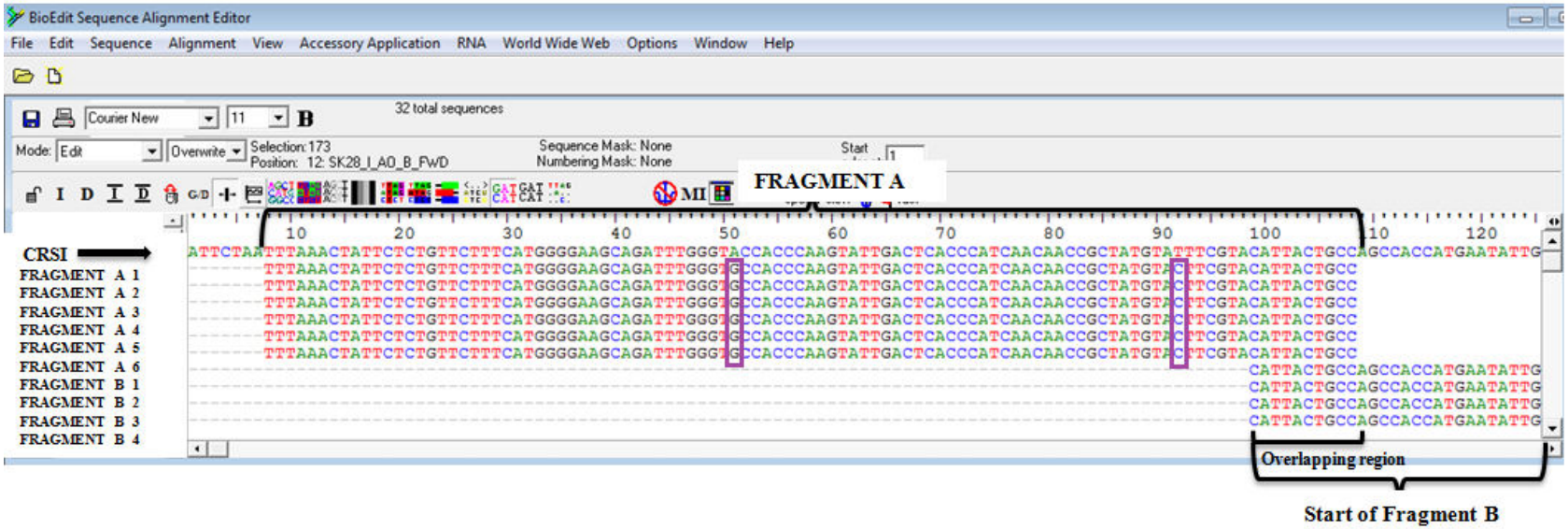


Figure 15 A screenshot from Bioedit 5.0.9 program shows the sequences (n=6) of fragment A which is the first fragment of HVI region and 140 bp in length. They are aligned with CRSI. A short part of fragment A which is 10 bp in length is overlapping with fragment B. Six sequences of fragment A belonged to two different extractions from the one tooth sample of individual SK28 and three independent PCRs from each extraction. The purple boxes indicate mutation motifs that are from A to G at 51 np (nucleotide position) and T to C at 92 np in reference to CRSI. These np are actually 16051th and 16092th positions in the given CRS when the sequence of whole mtDNA is used.

The mutation motifs (indicated by the purple boxes) observed on the 51 np and 92 np in reference to CRSI in Figure 15 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments A and B indicate that, with a high probability, these two sequences belong to the same individual.

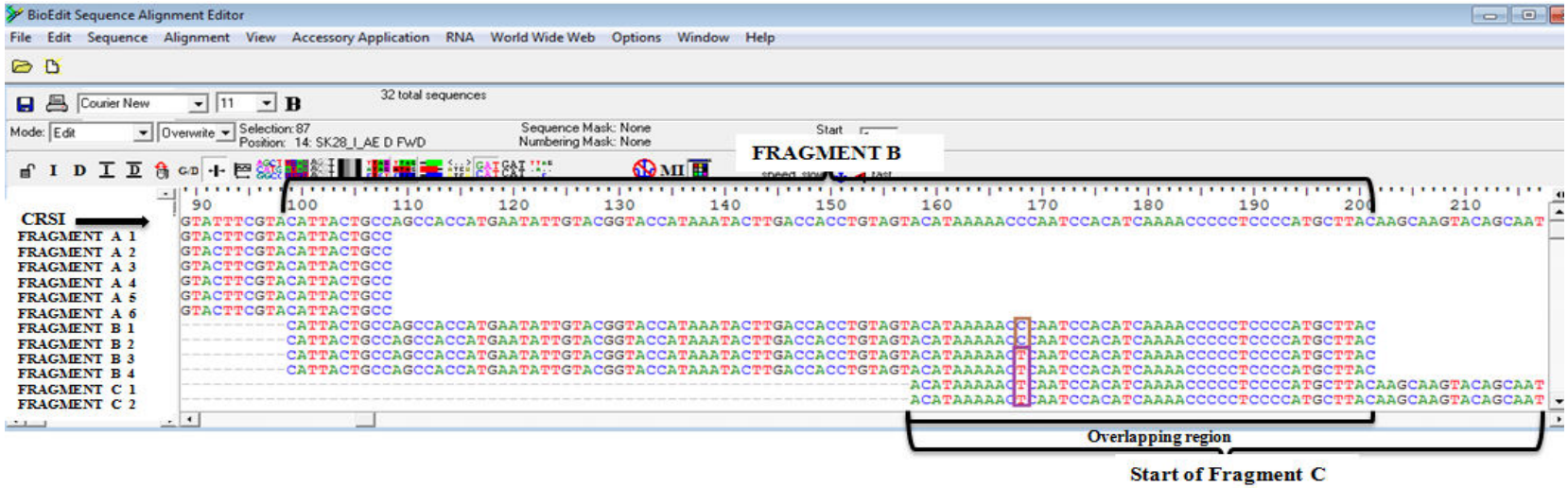


Figure 16 A screenshot from Bioedit 5.0.9 program shows the sequences (n=4) of fragment B which is the second fragment of HVI region and 139 bp in length. They are aligned with CRSI. A short part of fragment B which is 44 bp in length is overlapping with fragment C. Four sequences of fragment B belonged to two different extractions from the one tooth sample of individual SK28 and two independent PCRs from each extraction. The purple box indicates mutation motif that is from C to T at168 np in reference to CRSI. This np is actually 16168th position in the given CRS when the sequence of whole mtDNA is used. There were the "T" nucleotide on the 168th position of FRAGMENT B 3, FRAGMENT B 4 and also FRAGMENT C 1, FRAGMENT C 2, while it was "C" on the same position for FRAGMENT B 1 and FRAGMENT B 2 sequences (two B fragment sequences of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type1) . This nucleotide misincorporation is located at the overlapping region of fragment B and C.

The mutation motifs (indicated by the purple box) observed on the 168 np in reference to CRSI in Figure 16 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments B and C indicate that, with a high probability, these four sequences belong to the same individual.

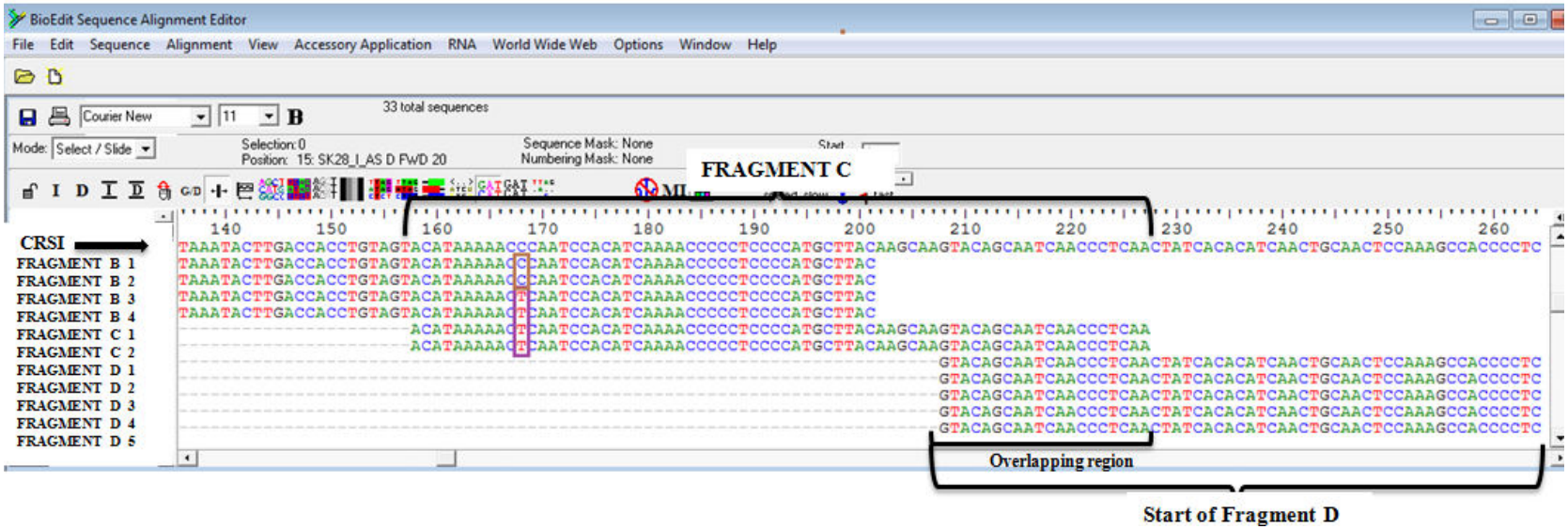


Figure 17 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment C which is the second fragment of HVI region and 109 bp in length. They are aligned with CRSI. A short part of fragment C which is 20 bp in length is overlapping with fragment D. Two sequences of fragment B belonged to two different extractions from the one tooth sample of individual SK28 and one independent PCR from each extraction. The purple box indicates mutation motif that is from C to T at 168 np in reference to CRSI. This np is actually 16168th position in the given CRS when the sequence of whole mtDNA is used. This mutation motif is located at the overlapping part and both of the C fragment sequences had the same mutations at that position.

The mutation motifs (indicated by the purple box) observed on the 168 np in reference to CRSI in Figure 17 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments C and D indicate that, with a high probability, these two sequences belong to the same individual.

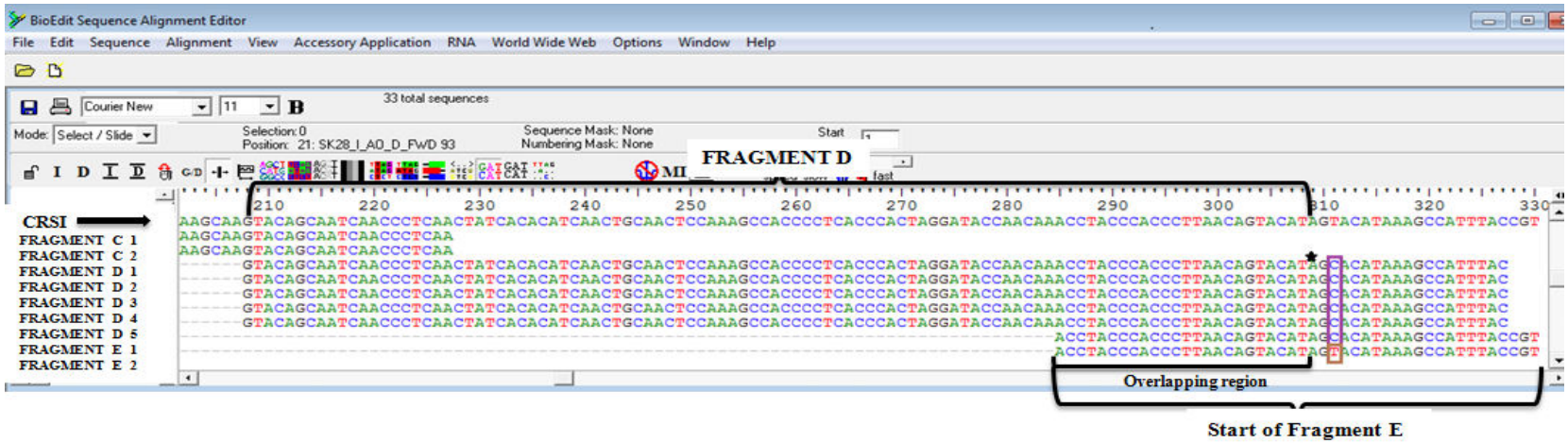


Figure 18. A screenshot from Bioedit 5.0.9 program shows the sequences (n=5) of fragment D which is the fourth fragment of HVI region and 133 bp in length. They are aligned with CRSI. A short part of fragment D which is 24 bp in length is overlapping with fragment E. Five sequences of fragment D belonged to three different extractions from the one tooth sample of individual SK28 and two independent PCRs from each of the two extractions, one PCR from another extraction. The purple box indicates mutation motif that is from T to C at 311 np in reference to CRSI. This np is actually 16311th position in the given CRS when the sequence of whole mtDNA is used. Normally fragment D ends at the 309 np which was shown with the black star. However all of the repetitions of the fragment D sequences were able to be read 19 additionally nucleotide, which were on the reverse primer binding region of the sequences. Due to the limited number of repetitions of E fragment sequences, in the present study, the five D fragment sequences were examined up to the 327 np (3' end of D fragment was extended 19 bp). This cases will be explained in the Discussion Chapter. The mutation motif at the overlapping part of fragments D and E was observed on one of the E fragment sequences.

The mutation motifs (indicated by the purple boxes) observed on the 311 np in reference to CRSI in Figure 18 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these five sequences belong to the same individual.

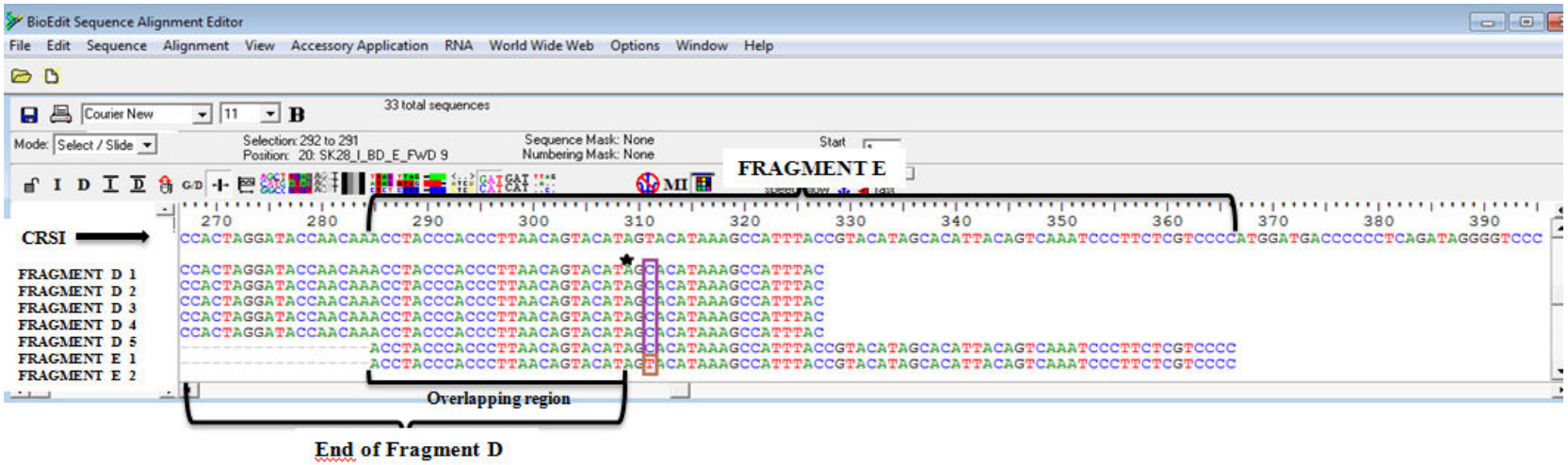


Figure 19 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment E which is the last fragment of HVI region and 118 bp in length. They are aligned with CRSI. A short part of fragment E which is 24 bp in length is overlapping with fragment D. Two sequences of fragment D belonged to two different extractions from the one tooth sample of individual SK28 and one independent PCR from each extraction. The purple box indicates mutation motif that is from T to C at 311 np in reference to CRSI. This np is actually 16311th position in the given CRS when the sequence of whole mtDNA is used. As it was explained in Figure 18 five sequences of fragment D were extended 19 bp more to check the accuracy of the sequences using the mutation motifs. One of the E fragment sequences had the same mutation at that position. There were the "C" nucleotide on the 311th position of all D fragment and FRAGMENT E 1 sequences while it was "T" on the same position for FRAGMENT E 2 sequence (second E fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type2). This nucleotide misincorporation is located at the overlapping region of fragment D and E.

The mutation motifs (indicated by the purple boxes) observed on the 311 np in reference to CRSI in Figure 19 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments E and D indicate that, with a high probability, these two sequences belong to the same individual.

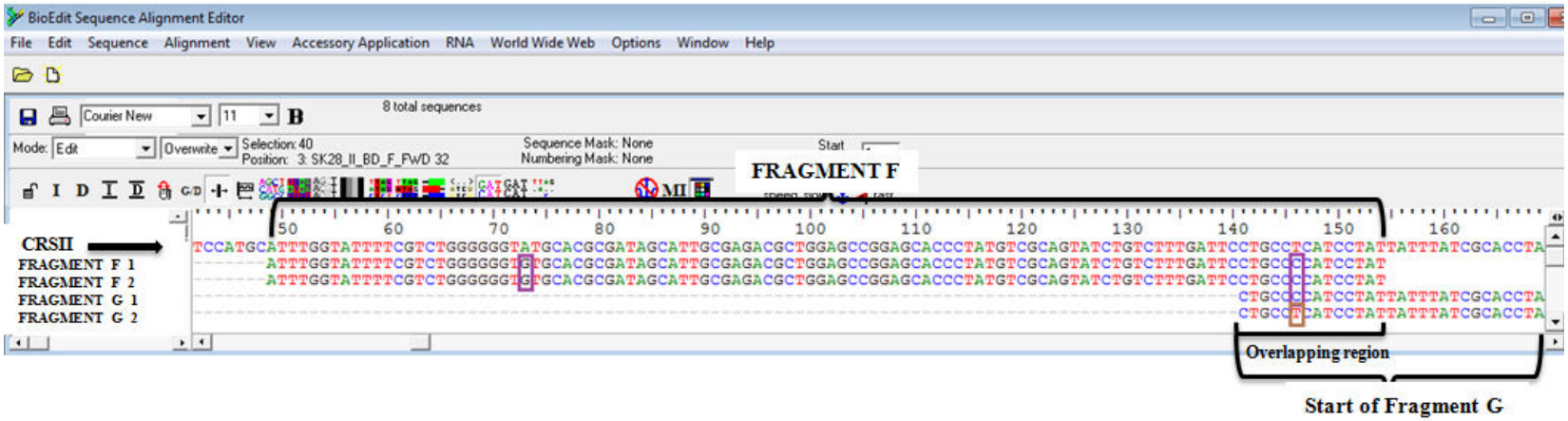


Figure 20 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment F which is the first fragment of HVII region and 147 bp in length. They are aligned with CRSII. A short part of fragment F which is 14 bp in length is overlapping with fragment G. Two sequences of fragment F belonged to two different extractions from the one tooth sample of individual SK28 and one independent PCR from each extraction. The purple boxes indicate mutation motifs that are from A to G at 73 np and T to C at 146 np in reference to CRSII. These np are actually 73rd and 146th positions in the given CRS when the sequence of whole mtDNA is used. T to C mutation motif is located at the overlapping part and both of the F fragment sequences had the same mutations at that two different positions.

The mutation motifs (indicated by the purple boxes) observed on the 73 np and 146 np in reference to CRSII in Figure 20 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

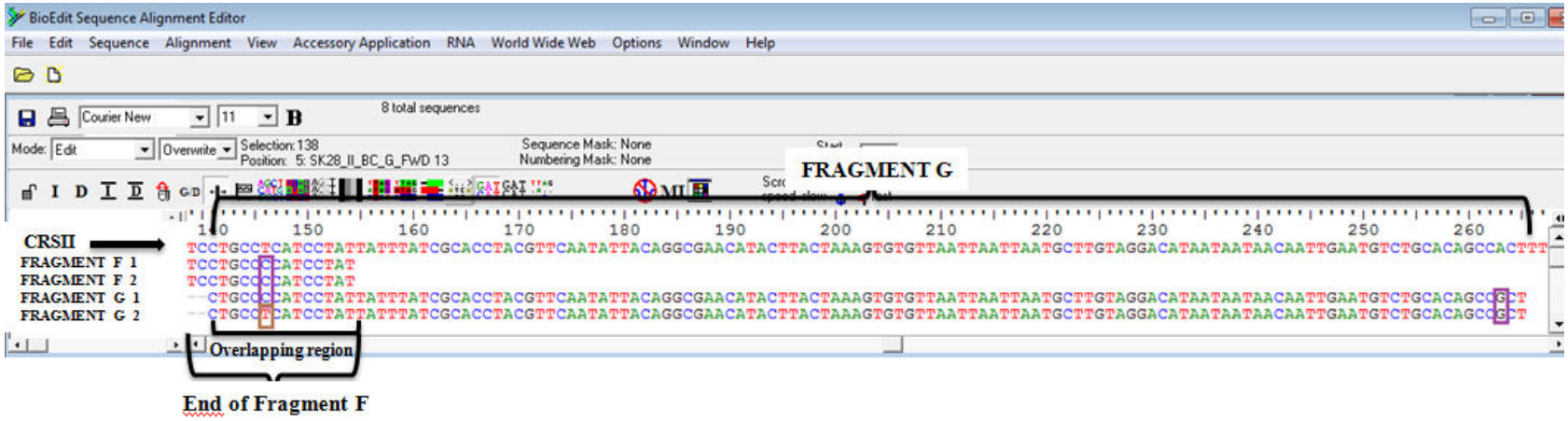
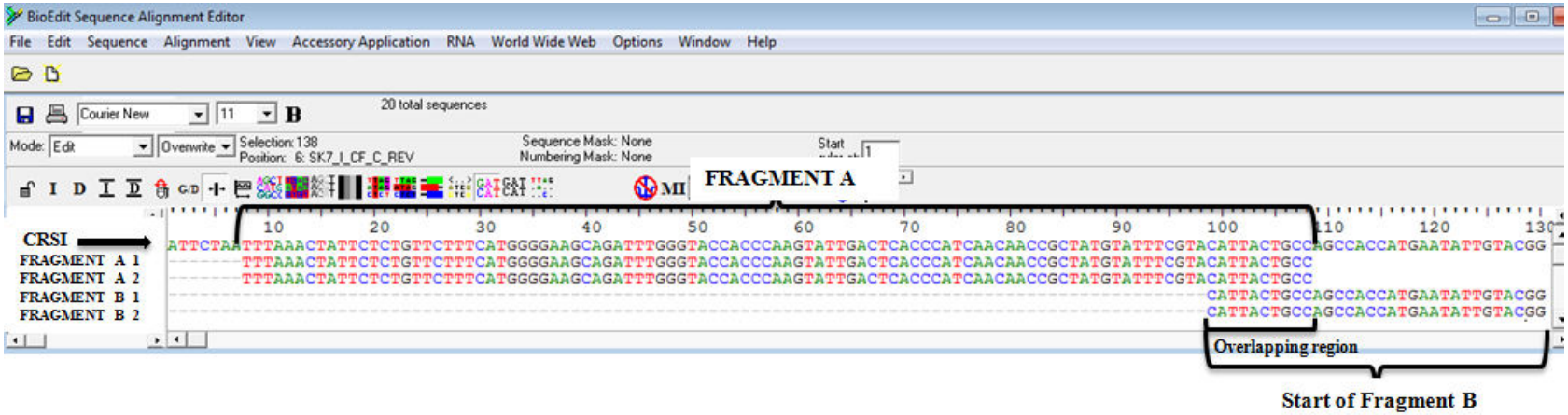


Figure 21 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment G which is the last fragment of HVII region and 166 bp in length. They are aligned with CRSII. A short part of fragment G which is 14 bp in length is overlapping with fragment F. Two sequences of fragment G belonged to two different extractions from the one tooth sample of individual SK28 and one independent PCR from each extraction. The purple boxes indicate mutation motifs that are from T to C at 146 np and A to G at 263 np in reference to CRSII. These np are actually 146th and 263rd positions in the given CRS when the sequence of whole mtDNA is used. T to C mutation motif is located at the overlapping part and one of the G fragment sequences had the same mutation at that position while both of the G fragment sequences had the same A to G mutations at another position. There were the "C" nucleotide on the 146th position of FRAGMENT F 1, FRAGMENT F 2 and FRAGMENT G 1, while it was "T" on the same position for FRAGMENT G 2 sequence (second G fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type2). This nucleotide misincorporation is located at the overlapping region of fragment F and G.

The mutation motifs (indicated by the purple boxes) observed on the 146 np and 263 np in reference to CRSII in Figure 21 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

SK7



150

Figure 22 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment A which is the first fragment of HVI region and 140 bp in length. They are aligned with CRSI. A short part of fragment A which is 10 bp in length is overlapping with fragment B. Two sequences of fragment A belonged to two different extractions from one tooth sample of individual SK7 and one PCR from each extraction. For this individual in both sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments A and B indicate that, with a high probability, these two sequences belong to the same individual.

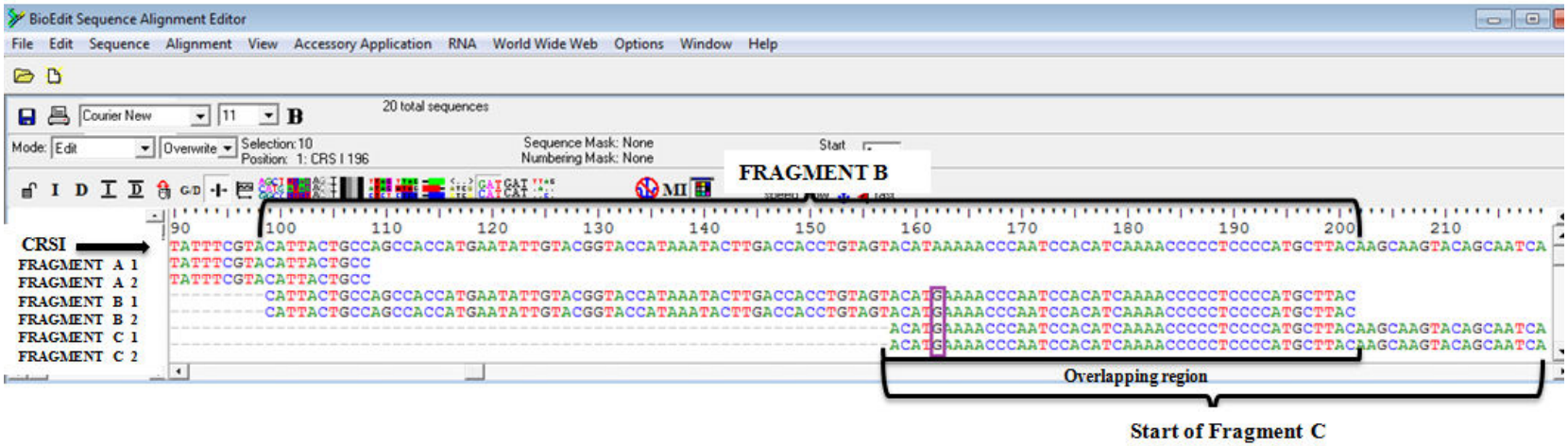


Figure 23 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment B which is the second fragment of HVI region and 139 bp in length. They are aligned with CRSI. A short part of fragment B which is 44 bp in length is overlapping with fragment C. Two sequences of fragment B belonged to two different extractions from one tooth sample of individual SK7 and one PCR from each extraction. The purple box indicates mutation motif that is from A to G at 162 np in reference to CRSI. This np is actually 16162nd position in the given CRS when the sequence of whole mtDNA is used. Both of the B fragment sequences had the same A to G mutations at this position.

The mutation motif (indicated by the purple box) observed on the 162 np in reference to CRSI in Figure 23 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments B and C indicate that, with a high probability, these two sequences belong to the same individual.

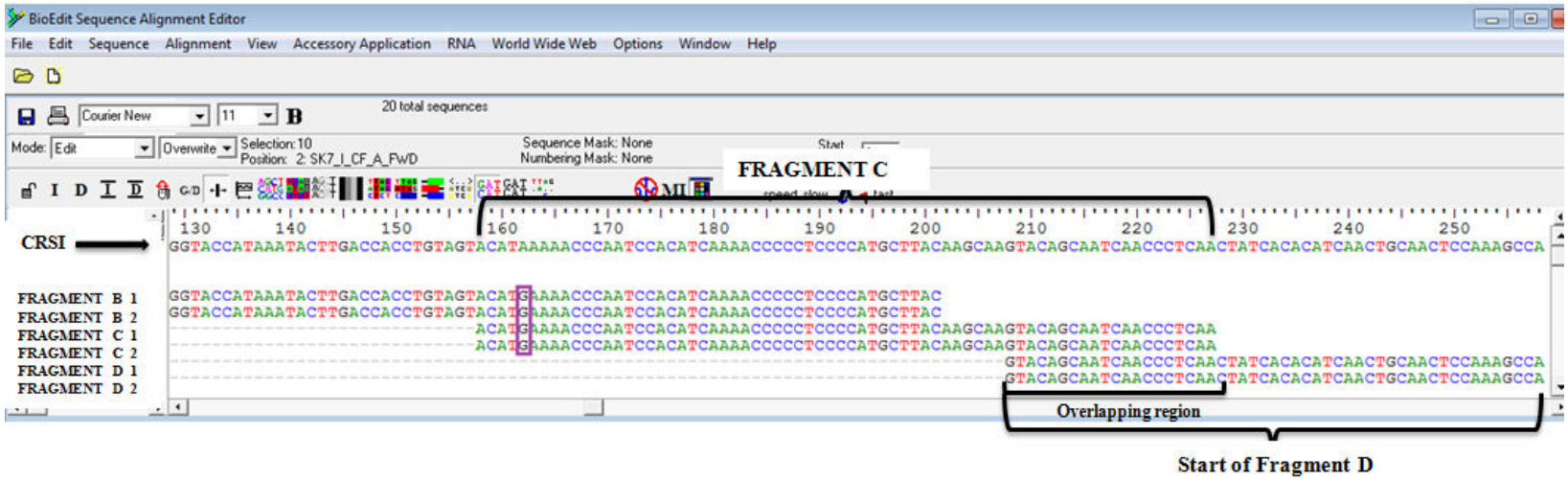


Figure 24 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment C which is the third fragment of HVI region and 109 bp in length. They are aligned with CRSI. A short part of fragment C which is 20 bp in length is overlapping with fragment D. Two sequences of fragment C belonged to two different extractions from one tooth sample of individual SK7 and one PCR from each extraction. The purple box indicates mutation motif that is from A to G at162 np in reference to CRSI. This np is actually 16162nd position in the given CRS when the sequence of whole mtDNA is used. Both of the C fragment sequences had the same A to G mutations at this position.

The mutation motif (indicated by the purple box) observed on the 162 np in reference to CRSI in Figure 24 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments C and D indicate that, with a high probability, these two sequences belong to the same individual.

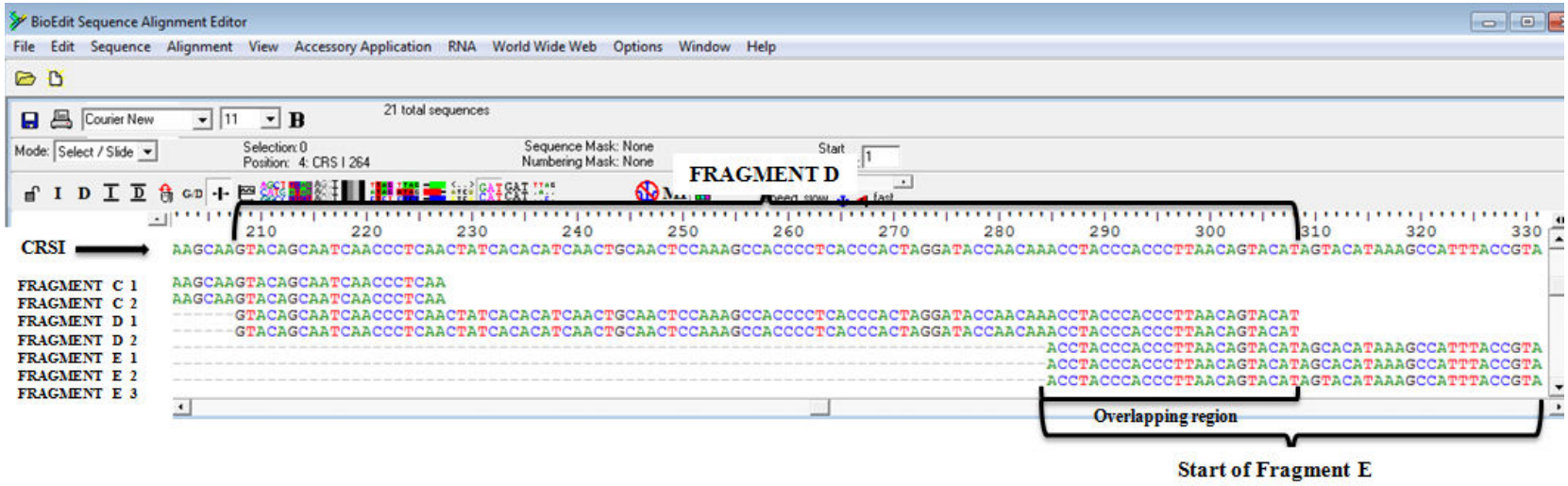


Figure 25 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment D which is the fourth fragment of HVI region and 133 bp in length. They are aligned with CRSI. A short part of fragment D which is 24 bp in length is overlapping with fragment E. Two sequences of fragment D belonged to two different extractions from one tooth sample of individual SK7 and one PCR from each extraction. For this individual in both of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these two sequences belong to the same individual.

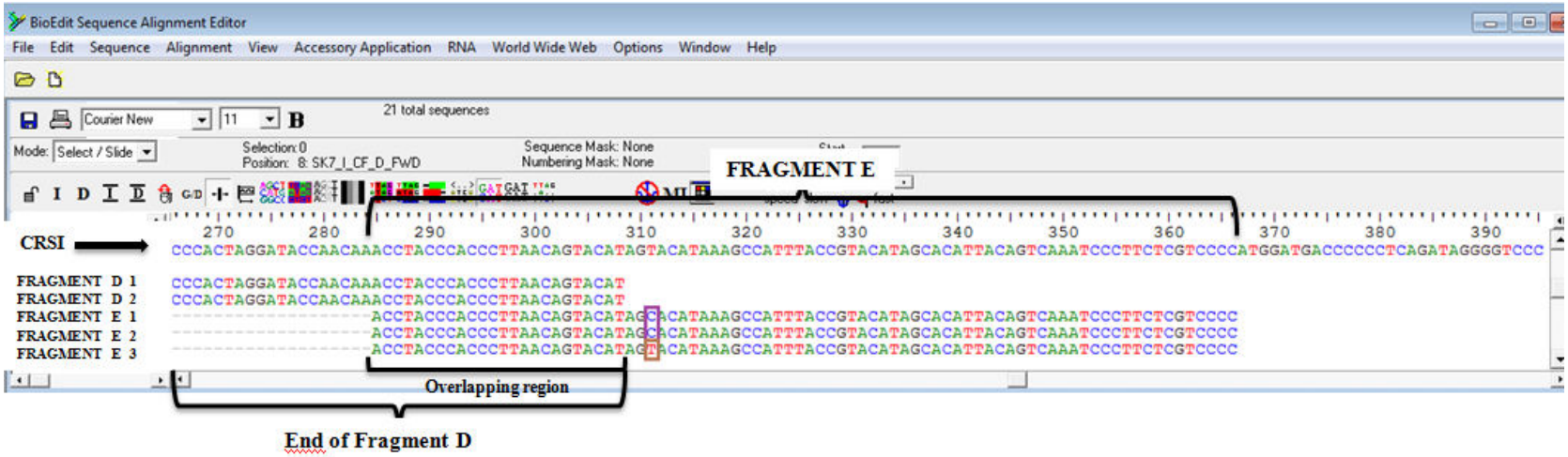


Figure 26 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment E which is the last fragment of HVI region and 118 bp in length. They are aligned with CRSI. A short part of fragment E which is 24 bp in length is overlapping with fragment D. Three sequences of fragment E belonged to two different extractions from one tooth sample of individual SK7 and two independent PCRs from one extraction, one PCR from another extraction. The purple box indicates mutation motifs that are from T to C at 311 np in reference to CRSI. This np are actually 16311th position in the given CRS when the sequence of whole mtDNA is used. There were the "C" nucleotide on the 311th position of FRAGMENT E 1 and FRAGMENT E 2, while it was "T" on the same position for FRAGMENT E 3 sequence (third E fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type2).

The mutation motif (indicated by the purple box) observed on the 311 np in reference to CRSI in Figure 26 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these three sequences belong to the same individual.

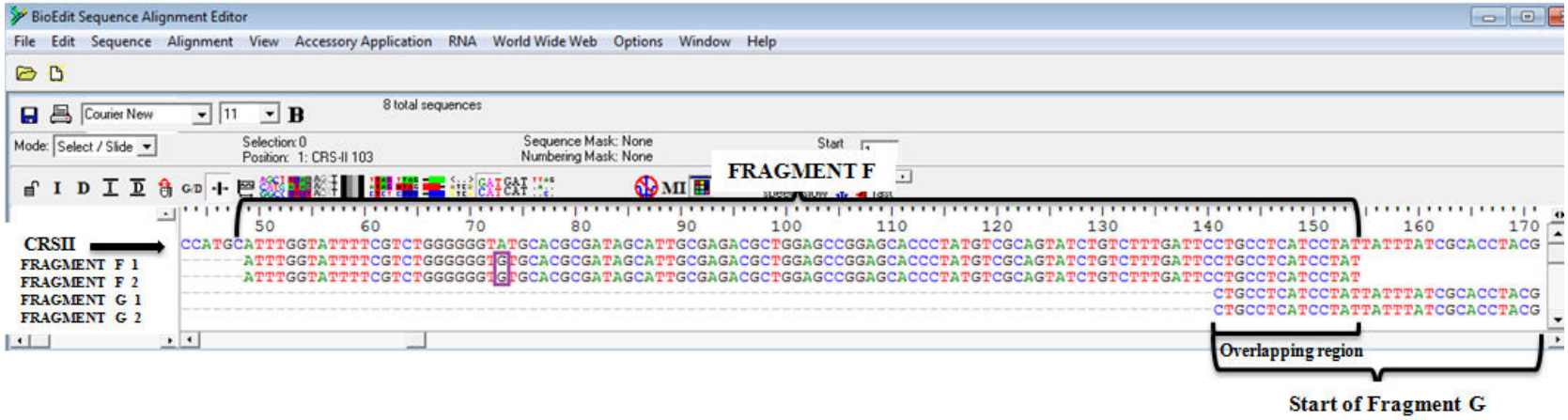


Figure 27 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment F which is the first fragment of HVII region and 147 bp in length. They are aligned with CRSII. A short part of fragment F which is 14 bp in length is overlapping with fragment G. Two sequences of fragment F belonged to two different extractions from one tooth sample of individual SK7 and one PCR from each extraction. The purple box indicates mutation motif that is from A to G at 73 np in reference to CRSII. This np are actually 73rd positions in the given CRS when the sequence of whole mtDNA is used. Both of the F fragment sequences had the same A to G mutations at this position.

The mutation motif (indicated by the purple box) observed on the 73 np in reference to CRSII in Figure 27 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

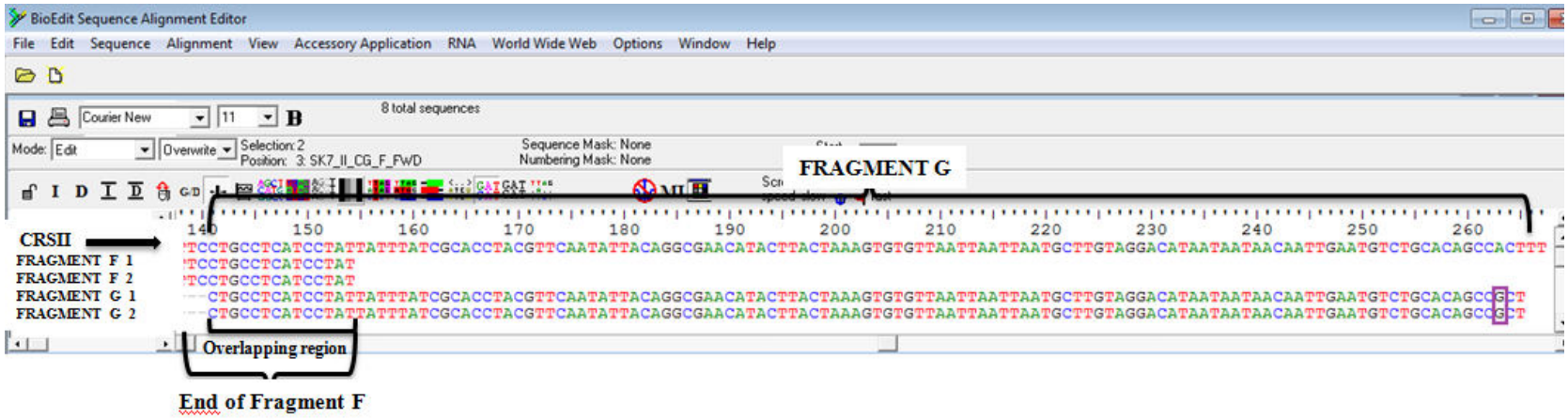


Figure 28 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment G which is the last fragment of HVII region and 166 bp in length. They are aligned with CRSII. A short part of fragment G which is 14 bp in length is overlapping with fragment F. Two sequences of fragment G belonged to two different extractions from one tooth sample of individual SK7 and one PCR from each extraction. The purple box indicates mutation motif that is from A to G at 263 np in reference to CRSII. This np are actually 263rd positions in the given CRS when the sequence of whole mtDNA is used. Both of the G fragment sequences had the same A to G mutations at this position.

The mutation motif (indicated by the purple box) observed on the 263 np in reference to CRSII in Figure 28 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments G and F indicate that, with a high probability, these two sequences belong to the same individual.

SK17

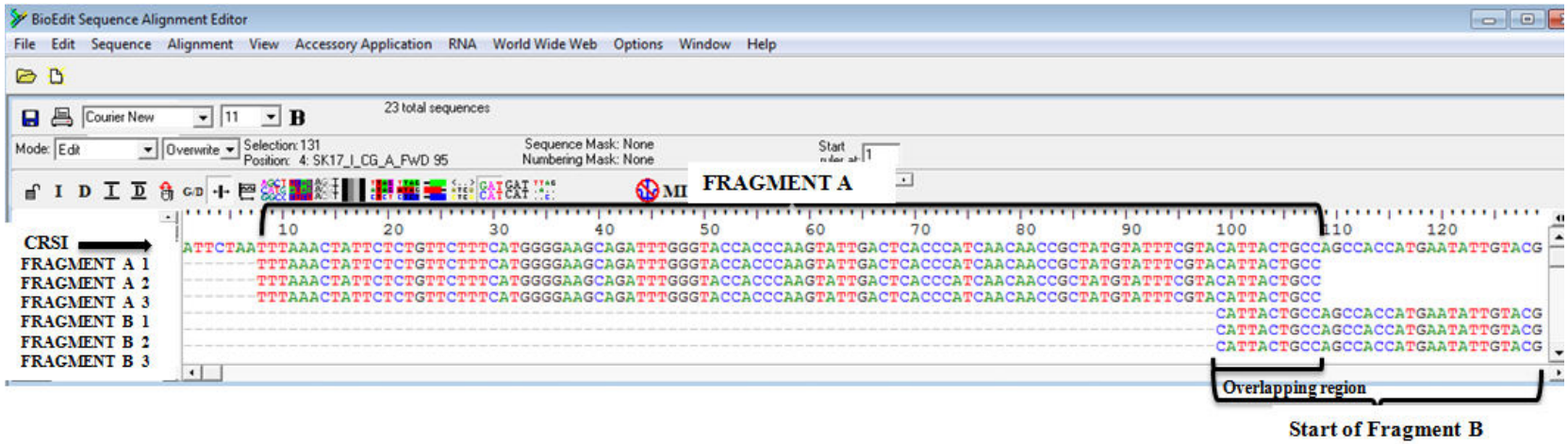


Figure 29 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment A which is the first fragment of HVI region and 140 bp in length. They are aligned with CRSI. A short part of fragment A which is 10 bp in length is overlapping with fragment B. Three sequences of fragment A belonged to two different extractions from one tooth sample of individual SK17 and two independent PCRs from one extraction, one PCR from another extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments A and B indicate that, with a high probability, these three sequences belong to the same individual.

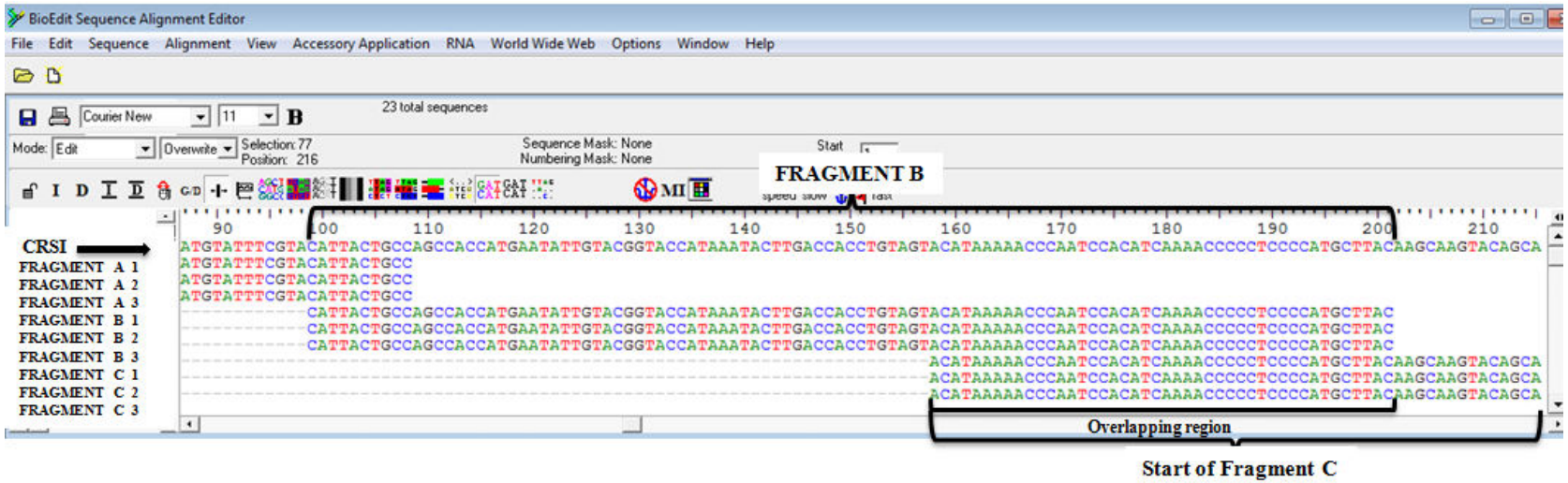


Figure 30 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment B which is the second fragment of HVI region and 139 bp in length. They are aligned with CRSI. A short part of fragment B which is 44 bp in length is overlapping with fragment C. Three sequences of fragment B belonged to two different extractions from one tooth sample of individual SK17 and two independent PCRs from one extraction, one PCR from another extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments B and C indicate that, with a high probability, these three sequences belong to the same individual.

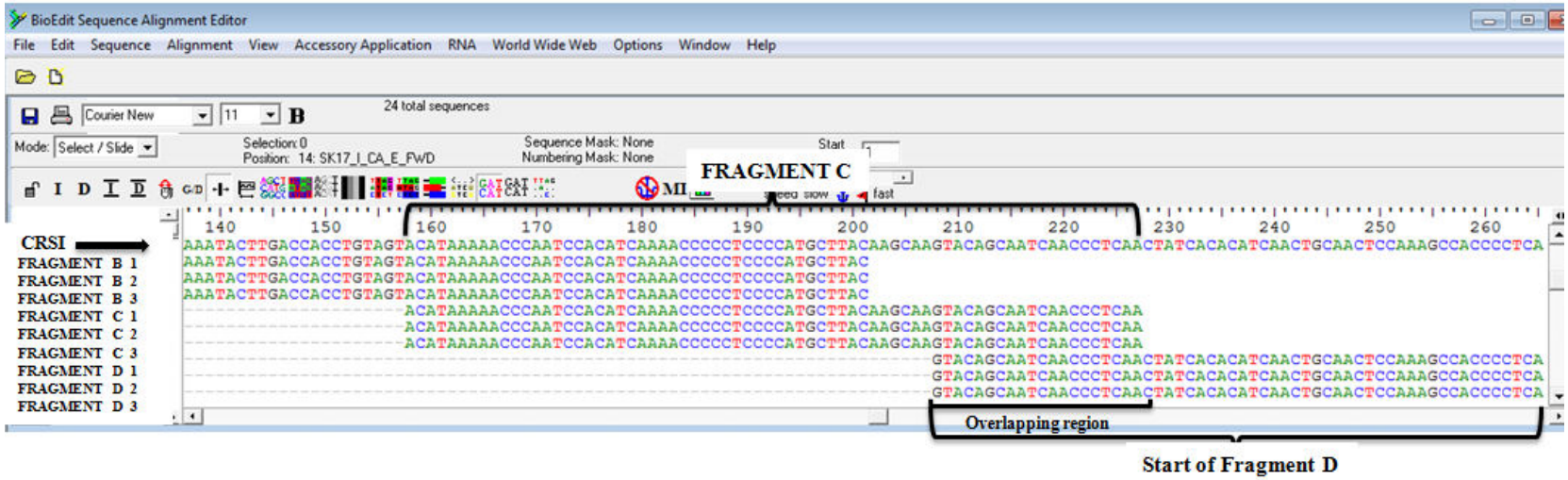


Figure 31 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment C which is the third fragment of HVI region and 109 bp in length. They are aligned with CRSI. A short part of fragment C which is 20 bp in length is overlapping with fragment D. Three sequences of fragment C belonged to two different extractions from one tooth sample of individual SK17 and two independent PCRs from one extraction, one PCR from another extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments C and D indicate that, with a high probability, these three sequences belong to the same individual.

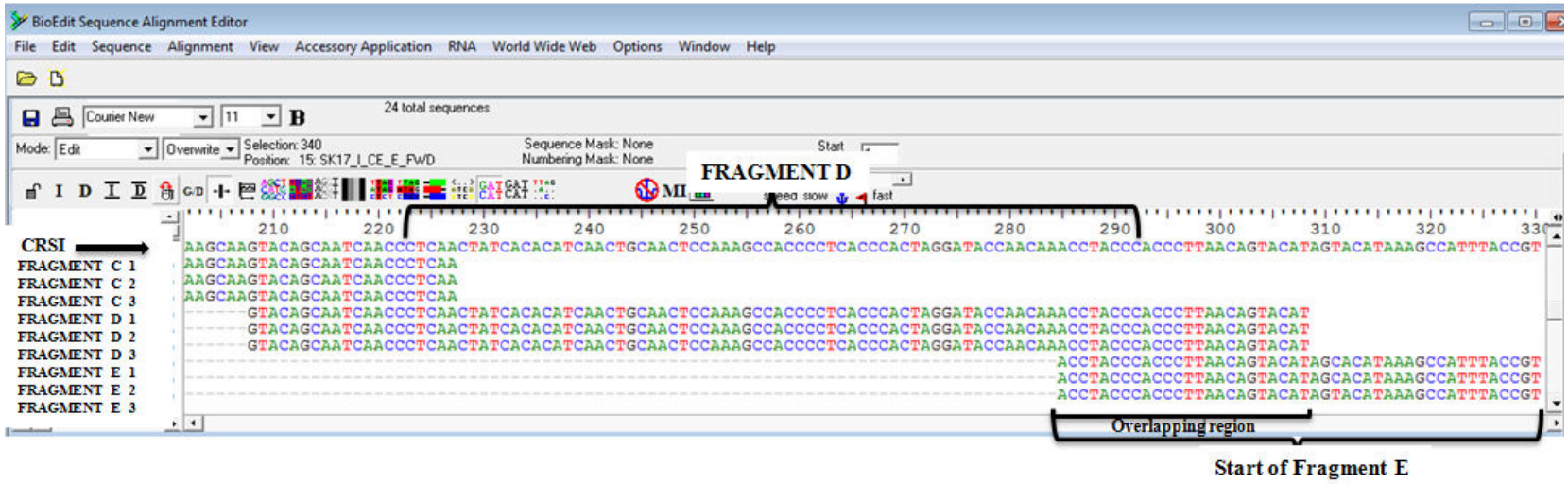


Figure 32 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment D which is the fourth fragment of HVI region and 133 bp in length. They are aligned with CRSI. A short part of fragment D which is 24 bp in length is overlapping with fragment E. Three sequences of fragment D belonged to two different extractions from one tooth sample of individual SK17 and two independent PCRs from one extraction, one PCR from another extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these three sequences belong to the same individual.

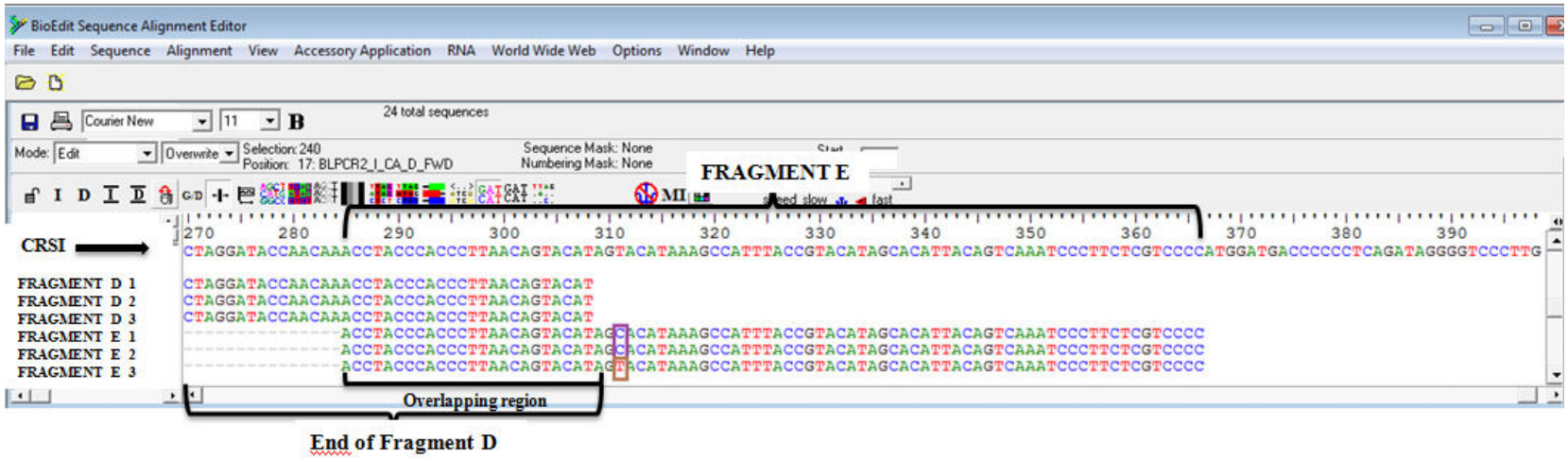


Figure 33 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment E which is the last fragment of HVI region and 118 bp in length. They are aligned with CRSI. A short part of fragment E which is 24 bp in length is overlapping with fragment D. Three sequences of fragment E belonged to two different extractions from one tooth sample of individual SK17 and two independent PCRs from one extraction, one PCR from another extraction. The purple box indicates mutation motifs that are from T to C at 311np in reference to CRSI. This np are actually 16311th position in the given CRS when the sequence of whole mtDNA is used. There were the "C" nucleotide on the 311th position of FRAGMENT E 1 and FRAGMENT E 2, while it was "T" on the same position for FRAGMENT E 3 sequence (third E fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type2).

The mutation motifs (indicated by the purple boxes) observed on the 311 np in reference to CRSI in Figure 33 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these three sequences belong to the same individual.

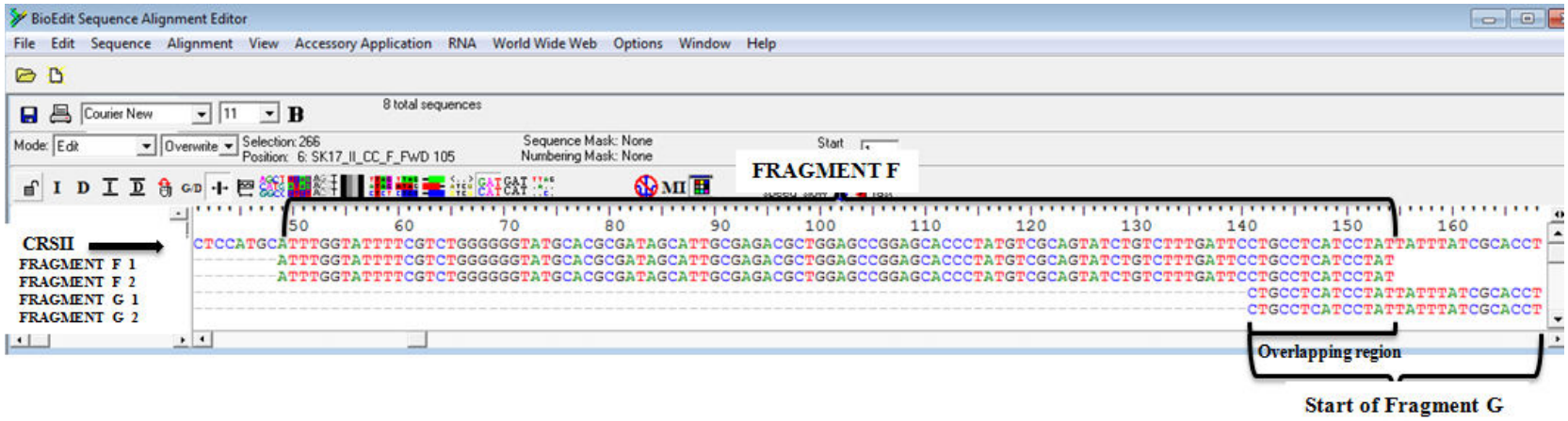


Figure 34 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment F which is the first fragment of HVII region and 147 bp in length. They are aligned with CRSII. A short part of fragment F which is 14 bp in length is overlapping with fragment G. Two sequences of fragment F belonged to two different extractions from one tooth sample of individual SK17 and one PCR from each extraction. For this individual in both of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSII.

Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

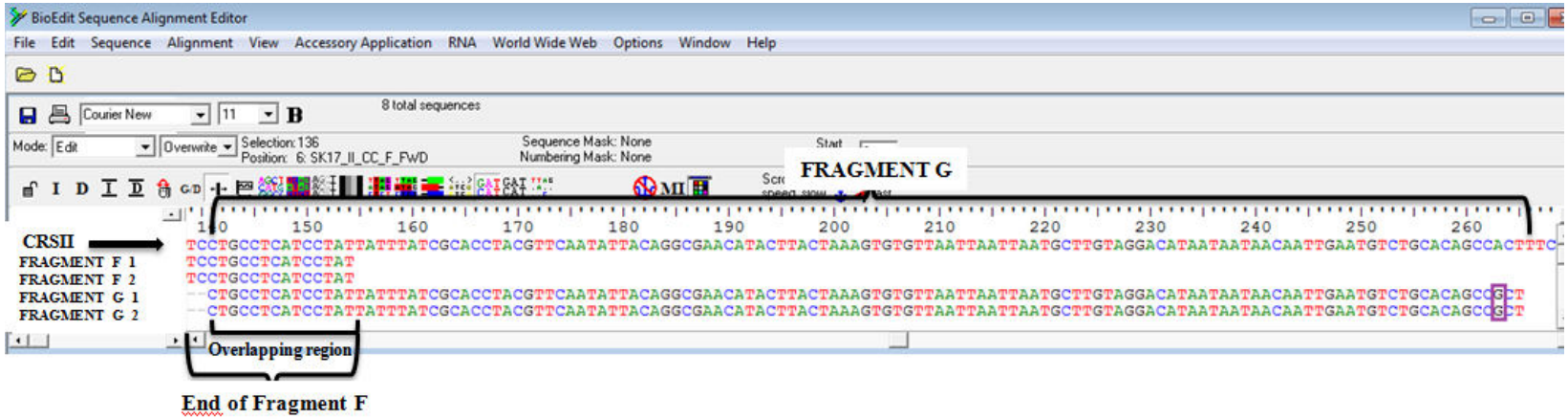


Figure 35 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment G which is the last fragment of HVII region and 166 bp in length. They are aligned with CRSII. A short part of fragment G which is 14 bp in length is overlapping with fragment F. Two sequences of fragment G belonged to two different extractions from one tooth sample of individual SK17 and one PCR from each extraction. The purple box indicates mutation motif that is from A to G at 263np in reference to CRSII. This np are actually 263rd positions in the given CRS when the sequence of whole mtDNA is used. Both of the G fragment sequences had the same A to G mutations at this position.

The mutation motif (indicated by the purple box) observed on the 263 np in reference to CRSII in Figure 35 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments G and F indicate that, with a high probability, these two sequences belong to the same individual.

SK21

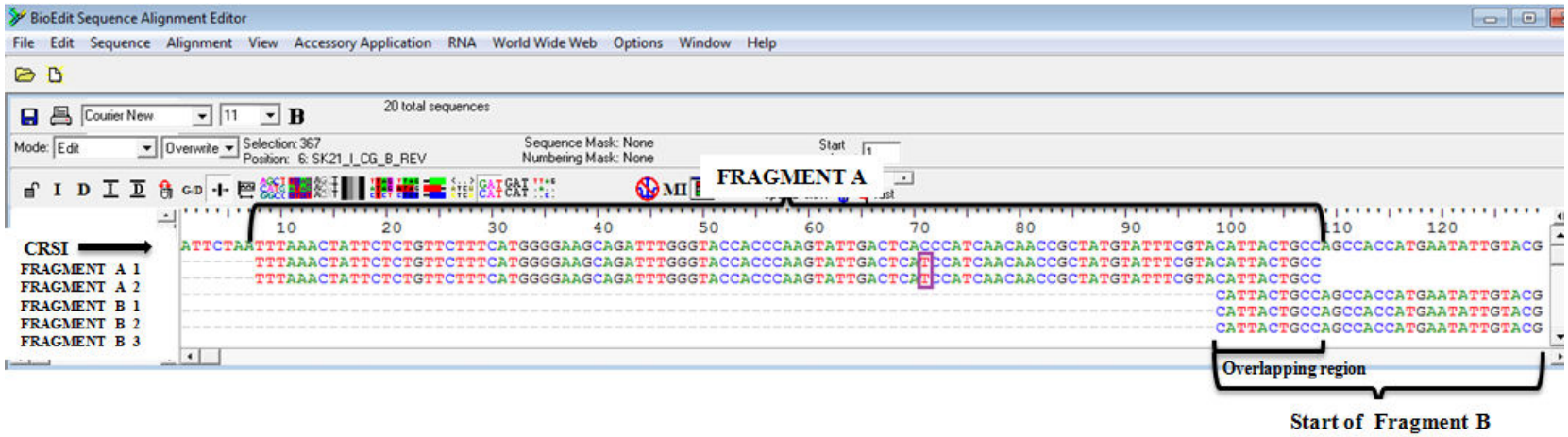


Figure 36 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment A which is the first fragment of HVI region and 140 bp in length. They are aligned with CRSI. A short part of fragment A which is 10 bp in length is overlapping with fragment B. Two sequences of fragment A belonged to two different extractions from one tooth sample of individual SK21 and one PCR from each extraction. The purple box indicates mutation motifs that are from C to T at 71 np in reference to CRSI. This np are actually 16071st positions in the given CRS when the sequence of whole mtDNA is used. Both of the A fragment sequences had the same C to T mutations at this position.

The mutation motif (indicated by the purple box) observed on the 71 np in reference to CRSI in Figure 36 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments A and B indicate that, with a high probability, these two sequences belong to the same individual.

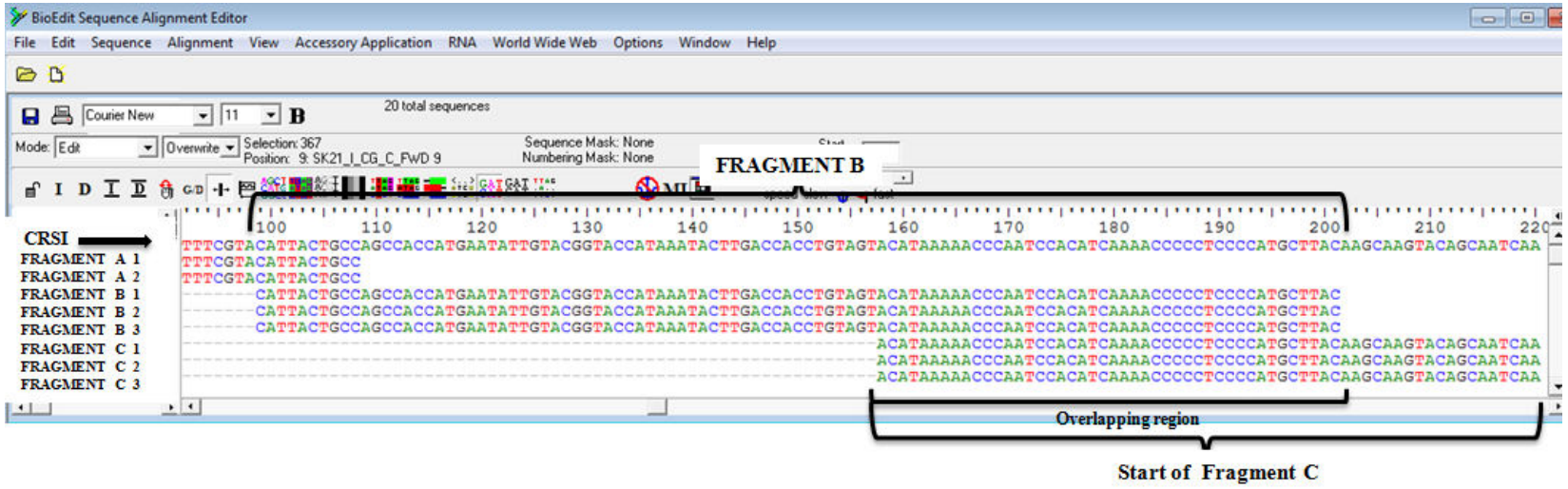


Figure 37 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment B which is the second fragment of HVI region and 139 bp in length. They are aligned with CRSI. A short part of fragment B which is 44 bp in length is overlapping with fragment C. Three sequences of fragment B belonged to two different extractions from one tooth sample of individual SK21 and two independent PCRs from one extraction, one PCR from another extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments B and C indicate that, with a high probability, these three sequences belong to the same individual.

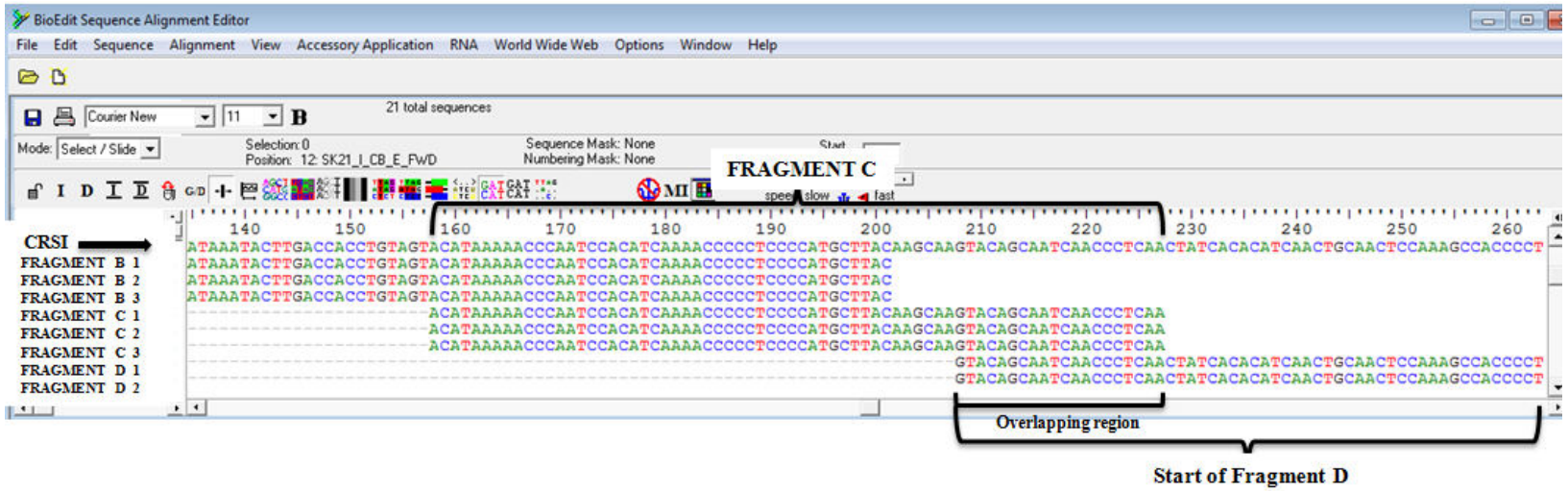


Figure 38 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment C which is the third fragment of HVI region and 109 bp in length. They are aligned with CRSI. A short part of fragment C which is 20 bp in length is overlapping with fragment D. Three sequences of fragment C belonged to two different extractions from one tooth sample of individual SK21 and two independent PCRs from one extraction, one PCR from another extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments C and D indicate that, with a high probability, these three sequences belong to the same individual.

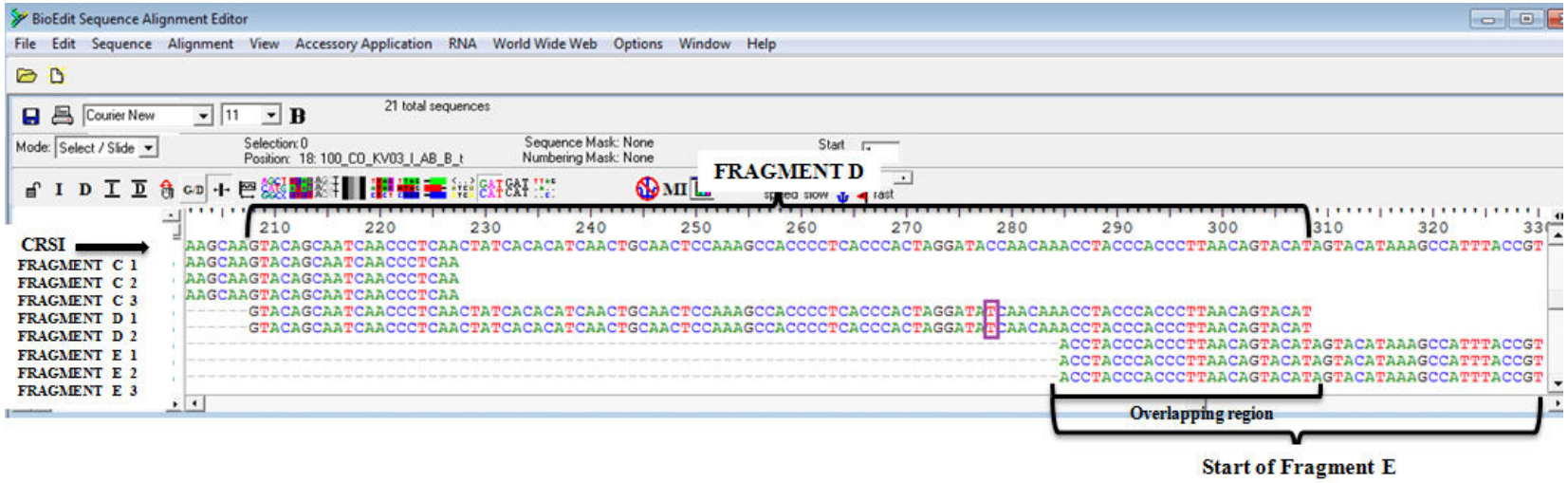


Figure 39 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment D which is the fourth fragment of HVI region and 133 bp in length. They are aligned with CRSI. A short part of fragment D which is 24 bp in length is overlapping with fragment E. Two sequences of fragment D belonged to two different extractions from one tooth sample of individual SK21 and one PCR from each extraction. The purple box indicates mutation motifs that are from C to T at 278 np in reference to CRSI. This np are actually 16278th positions in the given CRS when the sequence of whole mtDNA is used. Both of the D fragment sequences had the same C to T mutations at this position.

The mutation motif (indicated by the purple box) observed on the 278 np in reference to CRSI in Figure 39 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these two sequences belong to the same individual.

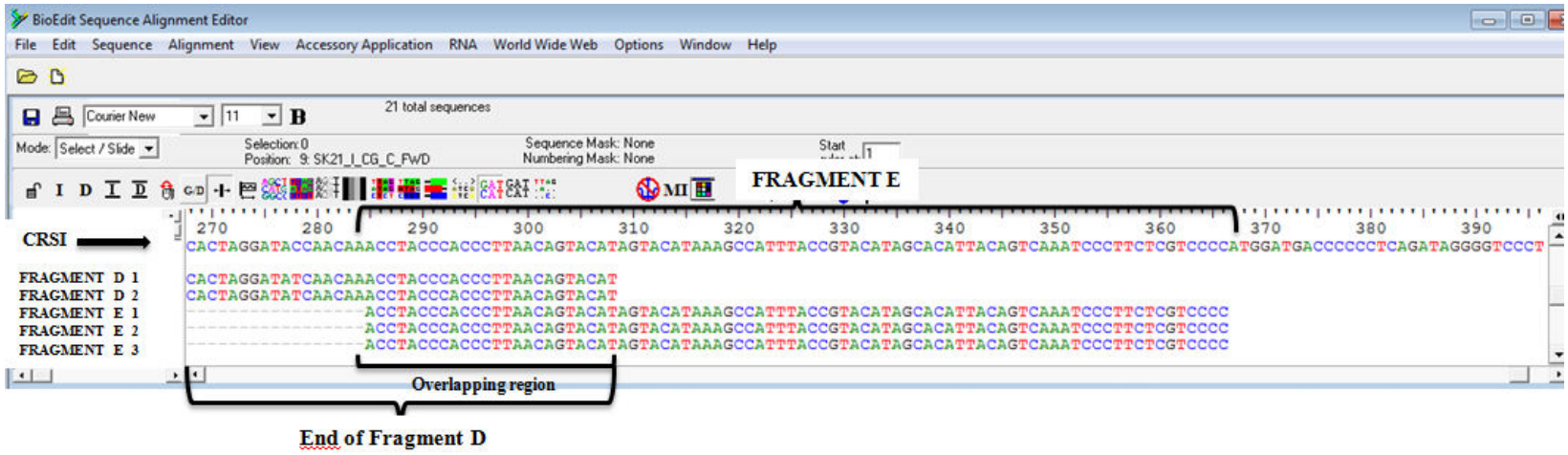


Figure 40 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment E which is the last fragment of HVI region and 118 bp in length. They are aligned with CRSI. A short part of fragment E which is 24 bp in length is overlapping with fragment D. Three sequences of fragment E belonged to two different extractions from one tooth sample of individual SK21 and two independent PCRs from one extraction, one PCR from another extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments E and D indicate that, with a high probability, these three sequences belong to the same individual.

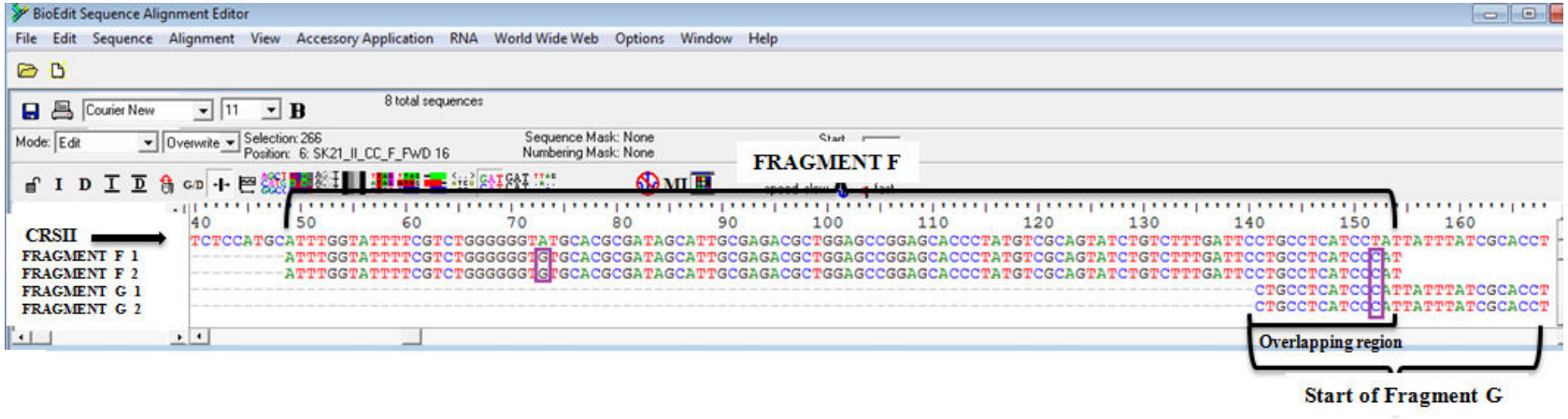


Figure 41 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment F which is the first fragment of HVII region and 147 bp in length. They are aligned with CRSII. A short part of fragment F which is 14 bp in length is overlapping with fragment G. Two sequences of fragment F belonged to two different extractions from the one tooth sample of individual SK21 and one independent PCR from each extraction. The purple boxes indicate mutation motifs that are from A to G at 73 np and T to C at 152 np in reference to CRSII. These np are actually 73rd and 152nd positions in the given CRS when the sequence of whole mtDNA is used. T to C mutation motif is located at the overlapping part and both of the F fragment sequences had the both mutation motifs at that two different positions.

The mutation motifs (indicated by the purple boxes) observed on the 73 np and 152 np in reference to CRSII in Figure 41 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

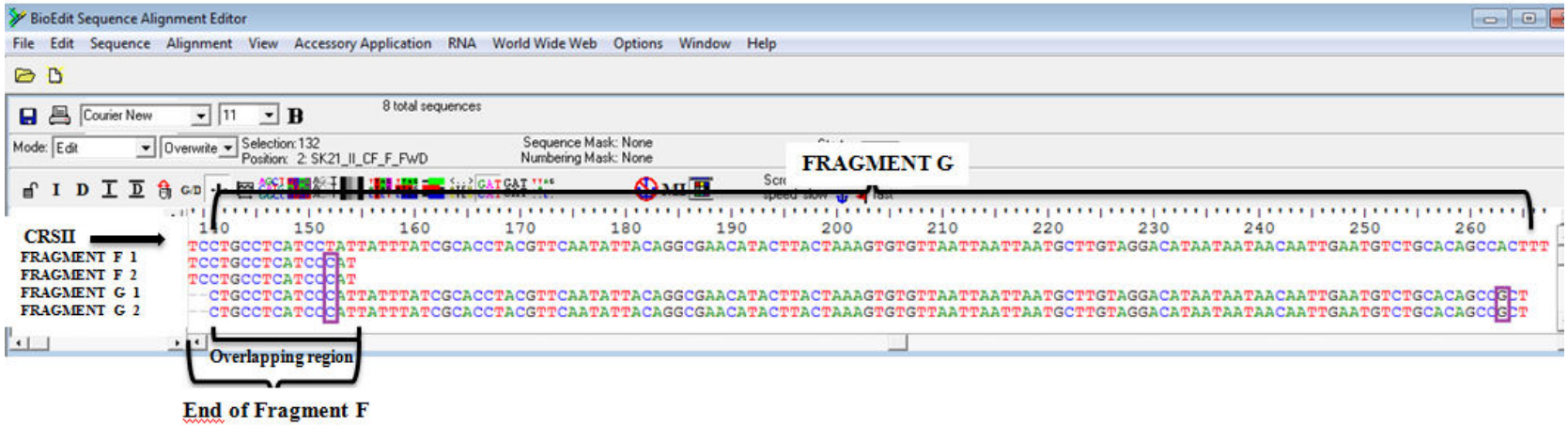


Figure 42 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment G which is the last fragment of HVII region and 166 bp in length. They are aligned with CRSII. A short part of fragment G which is 14 bp in length is overlapping with fragment F. Two sequences of fragment G belonged to two different extractions from the one tooth sample of individual SK21 and one independent PCR from each extraction. The purple boxes indicate mutation motifs that are from T to C at 152 np and A to G at 263 np in reference to CRSII. These np are actually 152nd and 263rd positions in the given CRS when the sequence of whole mtDNA is used. T to C mutation motif is located at the overlapping part and both of the G fragment sequences had the both mutation motifs at that two different positions.

The mutation motifs (indicated by the purple boxes) observed on the 152 np and 263 np in reference to CRSII in Figure 42 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

SK35

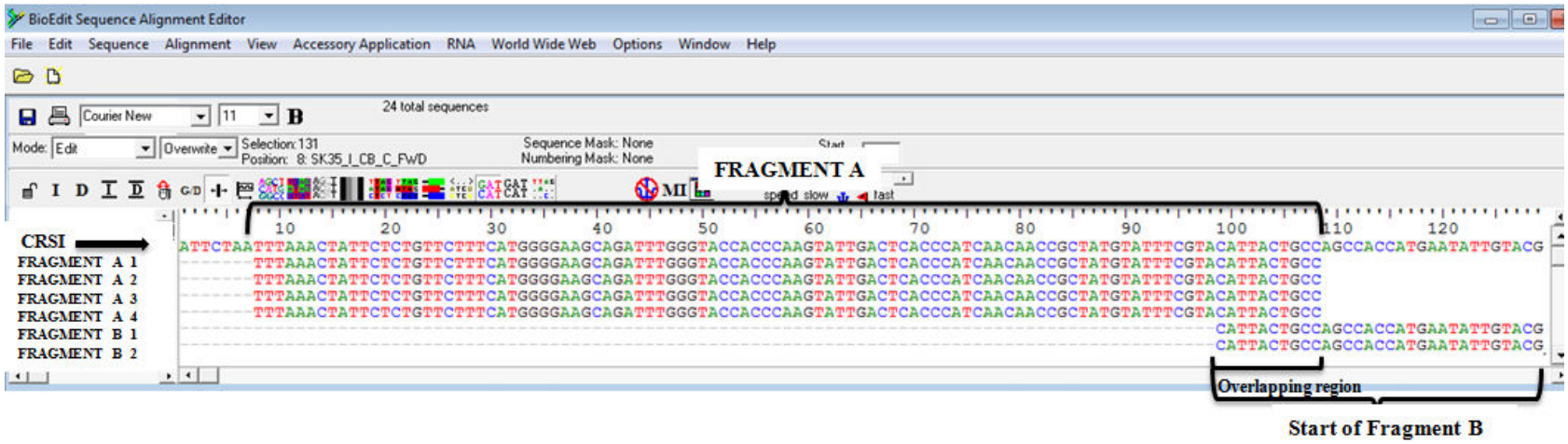


Figure 43 A screenshot from Bioedit 5.0.9 program shows the sequences (n=4) of fragment A which is the first fragment of HVI region and 140 bp in length. They are aligned with CRSI. A short part of fragment A which is 10 bp in length is overlapping with fragment B. Four sequences of fragment A belonged to two different extractions from one tooth sample of individual SK35 and two independent PCRs from each extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments A and B indicate that, with a high probability, these four sequences belong to the same individual.

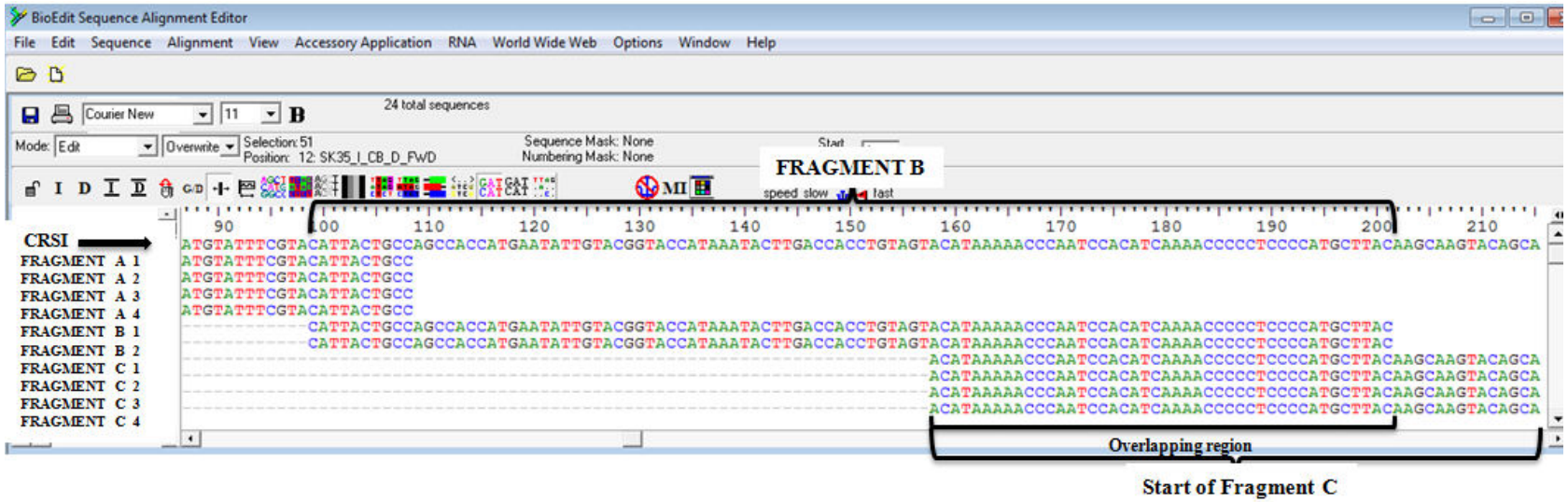


Figure 44 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment B which is the second fragment of HVI region and 139 bp in length. They are aligned with CRSI. A short part of fragment B which is 44 bp in length is overlapping with fragment C. Two sequences of fragment B belonged to two different extractions from one tooth sample of individual SK35 and one PCR from each extraction. For this individual in both of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments B and C indicate that, with a high probability, these two sequences belong to the same individual.

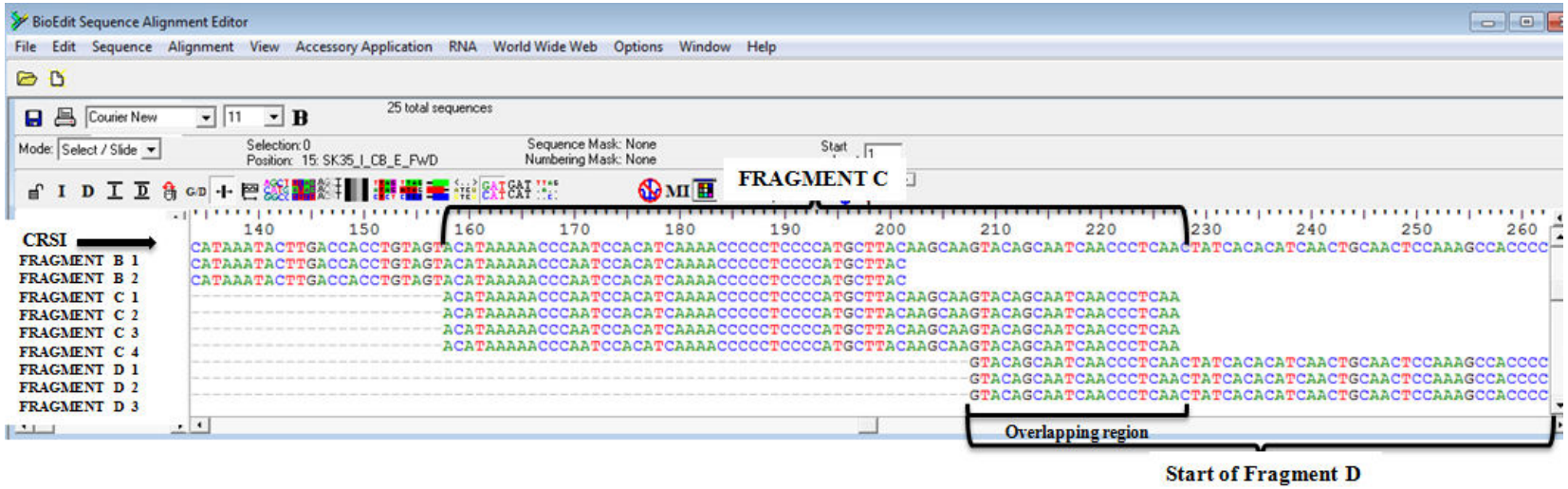


Figure 45 A screenshot from Bioedit 5.0.9 program shows the sequences (n=4) of fragment C which is the third fragment of HVI region and 109 bp in length. They are aligned with CRSI. A short part of fragment C which is 20 bp in length is overlapping with fragment D. Four sequences of fragment C belonged to two different extractions from one tooth sample of individual SK35 and two independent PCRs from each extraction. For this individual in all of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSI.

Once more, identical sequences observed for the overlapping parts of the fragments C and D indicate that, with a high probability, these four sequences belong to the same individual.

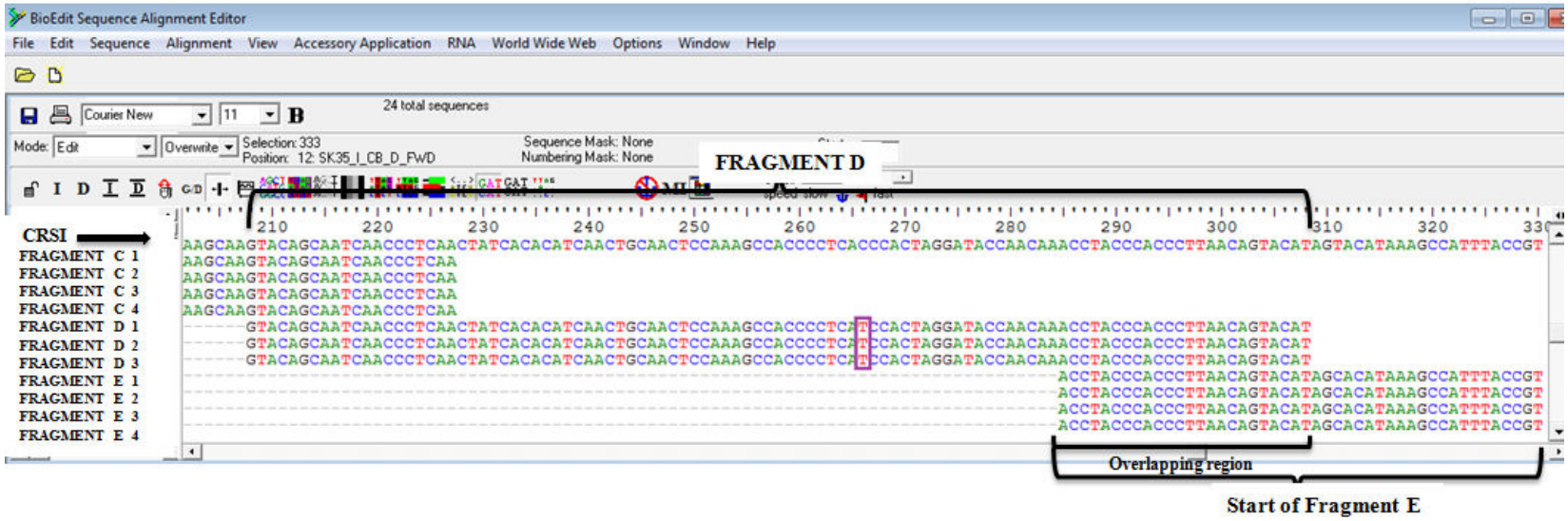


Figure 46 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment D which is the fourth fragment of HVI region and 133 bp in length. They are aligned with CRSI. A short part of fragment D which is 24 bp in length is overlapping with fragment E. Three sequences of fragment D belonged to two different extractions from one tooth sample of individual SK35 and two independent PCRs from one extraction, one PCR from another extraction. The purple box indicates mutation motifs that are from C to T at 266 np in reference to CRSI. This np are actually 16266th positions in the given CRS when the sequence of whole mtDNA is used. All of the D fragment sequences had the same C to T mutations at this position.

The mutation motif (indicated by the purple box) observed on the 266 np in reference to CRSI in Figure 46 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these three sequences belong to the same individual.

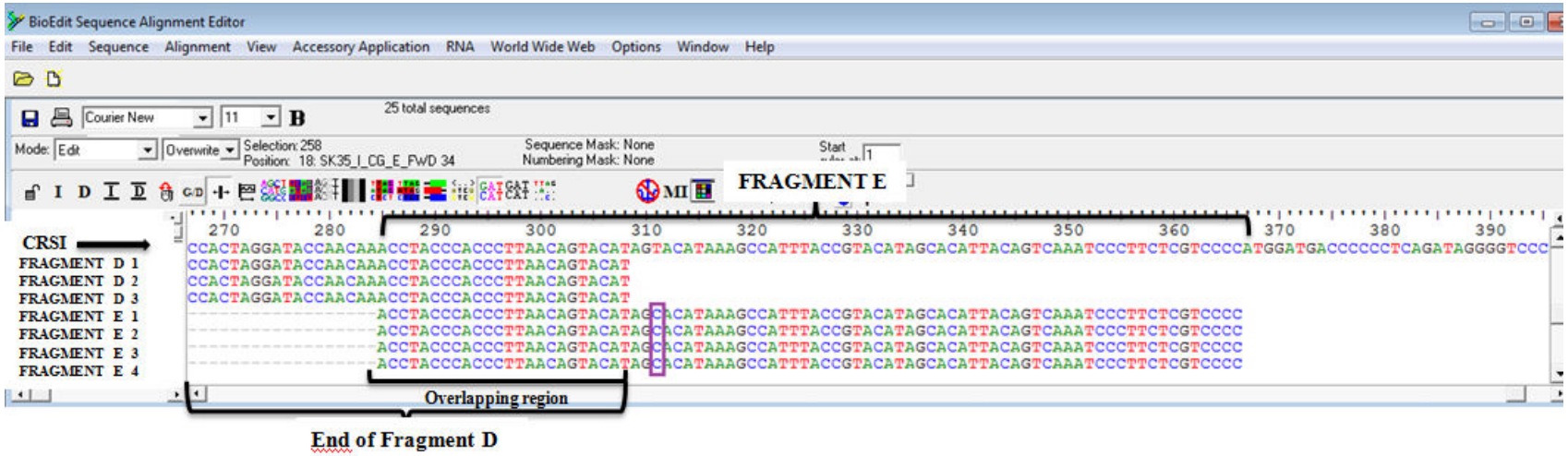


Figure 47 A screenshot from Bioedit 5.0.9 program shows the sequences (n=4) of fragment E which is the last fragment of HVI region and 118 bp in length. They are aligned with CRSI. A short part of fragment E which is 24 bp in length is overlapping with fragment D. Four sequences of fragment E belonged to two different extractions from one tooth sample of individual SK17 and two independent PCRs from each extraction. The purple box indicates mutation motifs that are from T to C at 311np in reference to CRSI. This np is actually 16311th position in the given CRS when the sequence of whole mtDNA is used. All of the E fragment sequences had the same T to C mutations at this position.

The mutation motifs (indicated by the purple boxes) observed on the 311 np in reference to CRSI in Figure 47 is going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments D and E indicate that, with a high probability, these four sequences belong to the same individual.

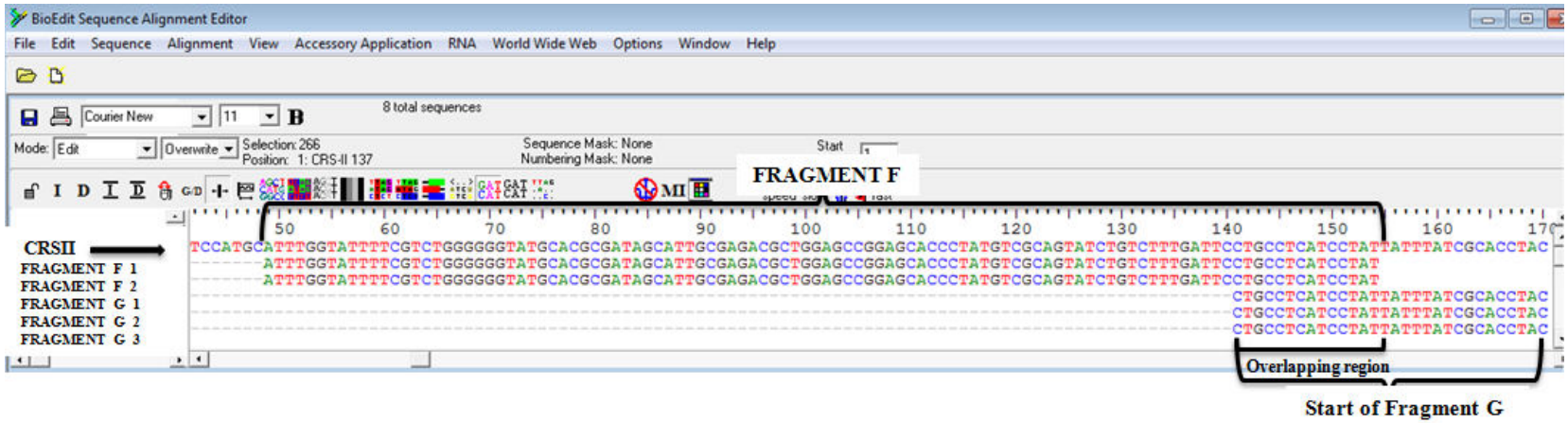


Figure 48 A screenshot from Bioedit 5.0.9 program shows the sequences (n=2) of fragment F which is the first fragment of HVII region and 147 bp in length. They are aligned with CRSII. A short part of fragment F which is 14 bp in length is overlapping with fragment G. Two sequences of fragment F belonged to two different extractions from one tooth sample of individual SK35 and one PCR from each extraction. For this individual in both of the sequences of this fragment, there was no mutation motif in relation to haplogroup determination in reference to CRSII.

Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these two sequences belong to the same individual.

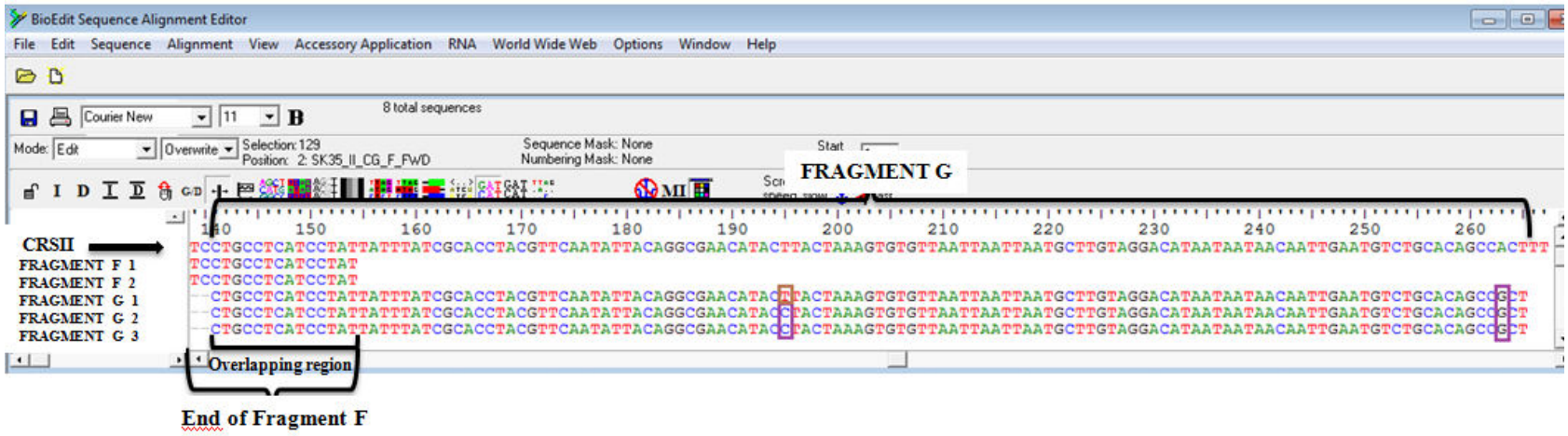


Figure 49 A screenshot from Bioedit 5.0.9 program shows the sequences (n=3) of fragment G which is the last fragment of HVII region and 166 bp in length. They are aligned with CRSII. A short part of fragment G which is 14 bp in length is overlapping with fragment F. Three sequences of fragment G belonged to two different extractions from the one tooth sample of individual SK35 and two independent PCRs from one extraction, one PCR from another extraction. The purple boxes indicate mutation motifs that are from T to C at 195 np and A to G at 263 np in reference to CRSII. These np are actually 195th and 263rd positions in the given CRS when the sequence of whole mtDNA is used. Also all of the G fragment sequences had the same mutation at 263rd position while one of the G fragment sequences did not have the T to C mutations at 195th position. There were the "C" nucleotide on the 195th position of FRAGMENT G 2 and FRAGMENT G 3, while it was "T" on k6the same position for FRAGMENT G 1 sequence (first G fragment sequence of the same sample) which was shown by brown box. Hence, it was considered as a nucleotide misincorporation (transition Type2).

The mutation motifs (indicated by the purple boxes) observed on the 195 np and 263 np in reference to CRSII in Figure 49 are going to be used in haplogroup determination. Once more, identical sequences observed for the overlapping parts of the fragments F and G indicate that, with a high probability, these three sequences belong to the same individual.

APPENDIX C: Detecting of Possible Contamination from Modern DNA

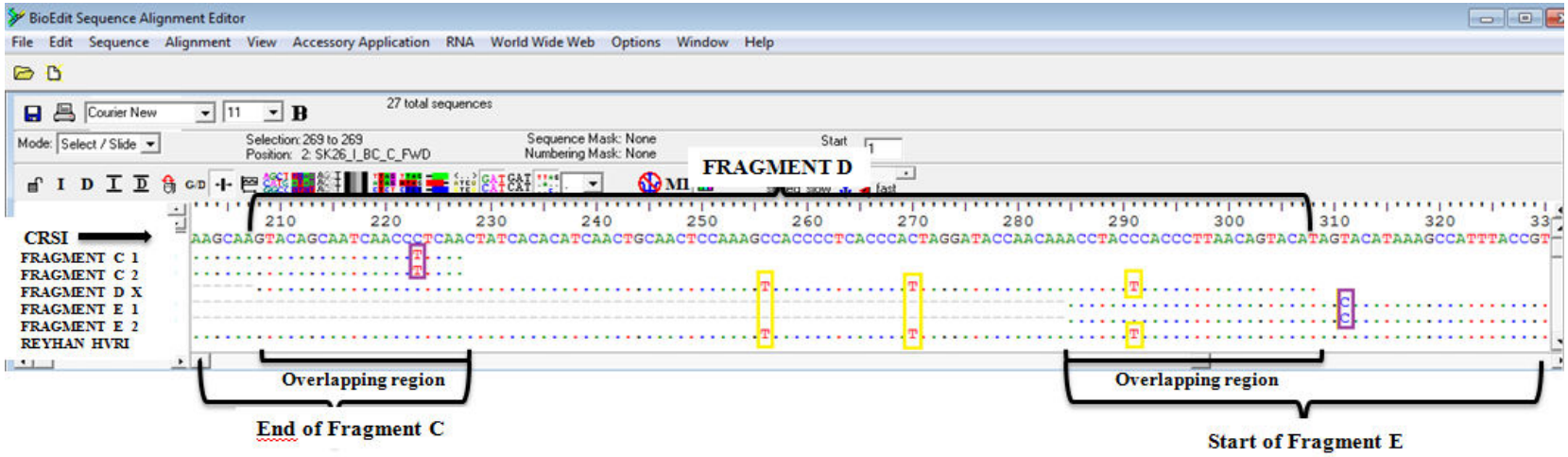


Figure 50 A screenhot shows the view of sequences by plotting identities to a standard as a dot in the BioEdit 5.0.9 software and can be used to visualize the mutation motifs.

A contamination from Reyhan Yaka (RY) was detected in only one D fragment sequence of sample SK26. In the screenshot from BioEdit 5.0.9 software, the purple boxes show the mutation motifs which were explained in Appendix B, used for the mtDNA haplogroup determination. Whereas the yellow boxes indicate contaminant D fragment sequence. It can be observed from the figure that FRAGMENT D X and REYHAN HVRI sequences have the same mutations from C to T at the 256 np, 270 np and 291 np in reference to CRSI. Thus this sequence of SK26 was determined as contaminant. Then all of the sequences obtained from the same PCR were excluded from analysis. For the rest of the aDNA sequences there were no contamination from the modern samples.

APPENDIX D: Pairwise population matrix of mean population haploid genetic distance

The table showing the population genetic (240 bp in length mtDNA HVI) distances between the Çemialo and 19 populations based on the genetic distance matrix of the PCA. Grec: Greece, Kazk: Kazakh, Tmen: Turkmen, Turk: Turkey, Uzbk: Uzbek, Cypr: Cyprus, Iran: Iran, Iraq: Iraq, Geor: Georgia, Azer: Azerbaijan, Arme: Armenia, Saga: Sagalassos, Cemi: Çemialo, Neol: Neareastern Neolithic population from Syria, Syra: Syria, Lbkt: LBKT (Linearbandkeramik culture in Transdanubia, 5800–4900 BC), LBK: LBK (Linear Pottery culture, 5500-4775 cal BC), RSC: RSC (Rössen culture, 4625-4250 cal BC), CWC: CWC (Corded Ware culture, 2800-2200/2050 cal BC), BBC: BBC (Bell Beaker culture, 2500-2200/2050 cal BC), UC:UC (Unetice culture, 2200-1550 cal BC),CRS: Cambridge Reference Sequence (Anderson et al., 1981).

	CRS	**Arme**	**Azer**	**Iran**	**Kazk**	**Geor**	**Tmen**	**Uzbk**	**Cypr**	**Grec**	**Turk**
CRS	0,000	2,867	3,275	2,560	3,477	2,554	3,300	3,250	2,747	2,559	2,656
Arme	2,867	5,221	5,694	4,916	5,829	5,039	5,580	5,613	5,170	4,969	5,085
Azer	3,275	5,694	6,142	5,398	6,286	5,475	6,125	6,083	5,589	5,463	5,551
Iran	2,560	4,916	5,398	4,561	5,325	4,693	5,228	5,307	4,811	4,702	4,764
Kazk	3,477	5,829	6,286	5,325	5,862	5,534	5,755	5,985	5,740	5,570	5,585
Geor	2,554	5,039	5,475	4,693	5,534	4,772	5,383	5,404	4,897	4,773	4,841
Tmen	3,300	5,580	6,125	5,228	5,755	5,383	5,516	5,803	5,594	5,373	5,378
Uzbk	3,250	5,613	6,083	5,307	5,985	5,404	5,803	5,947	5,529	5,375	5,432
Cypr	2,747	5,170	5,589	4,811	5,740	4,897	5,594	5,529	4,919	4,900	4,954
Grec	2,559	4,969	5,463	4,702	5,570	4,773	5,373	5,375	4,900	4,715	4,832
Turk	2,656	5,085	5,551	4,764	5,585	4,841	5,378	5,432	4,954	4,832	4,918
Cemi	2,000	4,490	4,801	4,000	5,023	4,124	5,000	4,693	4,009	4,141	4,163
Saga	2,755	5,101	5,629	4,822	5,649	4,951	5,414	5,530	5,074	4,907	4,985
Neol	1,400	4,003	4,339	3,544	4,559	3,598	4,510	4,260	3,500	3,663	3,696
Syra	2,708	5,126	5,606	4,876	5,807	4,945	5,624	5,600	5,023	4,900	4,991
Lbkt	2,615	4,994	5,430	4,725	5,661	4,783	5,426	5,540	4,905	4,784	4,884
UC	2,636	5,000	5,469	4,635	5,444	4,747	5,264	5,303	4,824	4,716	4,792
BBC	2,100	4,610	5,092	4,332	5,216	4,340	5,040	5,090	4,521	4,334	4,451
CWC	2,313	4,651	5,166	4,461	5,320	4,491	5,114	5,194	4,648	4,433	4,571
RSC	3,714	5,886	6,371	5,600	6,321	5,701	6,129	6,221	5,788	5,679	5,776
LBK	2,852	5,140	5,577	4,833	5,768	4,917	5,556	5,644	4,966	4,928	5,008

APPENDIX D (continued)

	Cemi	**Saga**	**Neol**	**Syra**	**Lbkt**	**UC**	**BBC**	**CWC**	**RSC**	**LBK**
CRS	2,000	2,755	1,400	2,708	2,615	2,636	2,100	2,313	3,714	2,852
Arme	4,490	5,101	4,003	5,126	4,994	5,000	4,610	4,651	5,886	5,140
Azer	4,801	5,629	4,339	5,606	5,430	5,469	5,092	5,166	6,371	5,577
Iran	4,000	4,822	3,544	4,876	4,725	4,635	4,332	4,461	5,600	4,833
Kazk	5,023	5,649	4,559	5,807	5,661	5,444	5,216	5,320	6,321	5,768
Geor	4,124	4,951	3,598	4,945	4,783	4,747	4,340	4,491	5,701	4,917
Tmen	5,000	5,414	4,510	5,624	5,426	5,264	5,040	5,114	6,129	5,556
Uzbk	4,693	5,530	4,260	5,600	5,540	5,303	5,090	5,194	6,221	5,644
Cypr	4,009	5,074	3,500	5,023	4,905	4,824	4,521	4,648	5,788	4,966
Grec	4,141	4,907	3,663	4,900	4,784	4,716	4,334	4,433	5,679	4,928
Turk	4,163	4,985	3,696	4,991	4,884	4,792	4,451	4,571	5,776	5,008
Cemi	2,571	4,431	2,343	4,232	4,308	4,045	3,943	4,080	5,020	4,302
Saga	4,431	4,994	3,894	5,065	4,903	4,884	4,521	4,565	5,863	5,053
Neol	2,343	3,894	1,867	3,763	3,703	3,583	3,310	3,494	4,600	3,674
Syra	4,232	5,065	3,763	5,059	4,893	4,900	4,488	4,575	5,857	5,047
Lbkt	4,308	4,903	3,703	4,893	4,435	4,805	4,326	4,349	5,407	4,522
UC	4,045	4,884	3,583	4,900	4,805	4,561	4,259	4,403	5,631	4,892
BBC	3,943	4,521	3,310	4,488	4,326	4,259	3,758	3,850	5,357	4,531
CWC	4,080	4,565	3,494	4,575	4,349	4,403	3,850	3,996	5,402	4,545
RSC	5,020	5,863	4,600	5,857	5,407	5,631	5,357	5,402	6,762	5,500
LBK	4,302	5,053	3,674	5,047	4,522	4,892	4,531	4,545	5,500	4,690