ASSOCIATION MAPPING FOR OIL AND PROTEIN CONTENT IN HAZELNUT (*Corylus avellana* L.)

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Molecular Biology and Genetics

by
Fatmanur SÖYLEMEZ

June 2020
İZMİR

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor Prof. Dr. Sami Doğanlar for giving me the opportunity to study in his research group, guidance and motivation. At the same time I would like thank Prof. Dr. Anne Frary for her patience, understanding and support while writing this thesis and I am grateful for her guidance during my master programme.

I would like to thank Hazelnut Research Institute members Dr. Hüseyin İrfan Balık for providing the plant materials. I would like to thank to The Scientific and Technological Research Council of Turkey (TUBITAK) (212T201). Also I would like to thank İYTE-TAM, in there especially Yekta Günay for her help in the oil extraction.

A special thanks to all the Frary/Doğanlar lab members for their help and kindness, especially two graduated lab members. Dr. Süleyman Can Öztürk, for his construction of hazelnut core collection and Dr. Mazen Ali Abuzayed for his help in the analysis.

Finally, I would like to thank very much to my family for always supporting and encouraging me. Without their support, I would not become the person I am now. Also I would like to give special thanks and love my fiancé for his love and support during my hard times.

ABSTRACT

ASSOCIATION MAPPING FOR OIL AND PROTEIN CONTENT IN HAZELNUT (*Corylus avellana* L.)

European hazelnut has an important place in terms of nutritional and economic value among tree nut species. Because of its nutritional content, the consumption of hazelnuts promotes human health in many ways. These nutritional components are controlled by multiple genes and affected by the environment; therefore, they are quantitative traits. The vast majority of world hazelnut production is provided by Turkey. So it is important to develop hazelnuts with higher nutritional quality for our country. In this work, we aimed to associate genetic diversity data and oil and protein content of hazelnuts to identify QTL. For this purpose, oil and protein content were measured in kernels of 96 accessions. Genotypic data were obtained with 30 SSR markers and resulted in 407 polymorphic alleles. According to allelic data, the mean dissimilarity value was 0.52 (52%) for the 96 accessions. Population structure analysis resulted in three clusters with 30, 30, and 16 accessions. Twenty accessions could not be assigned to any cluster and were considered admixed. Association mapping between allelic and phenotypic data indicated that five loci were significantly associated with oil content. The most significant result for oil content belonged to B628-307 loci ($p=0.0002$, $r^2=0.145$). Three loci were detected for protein content. Among them, A613-153 had the most significant effect ($p=0.003$, $r^2= 0.088$). We hope that our survey of germplasm and the identified loci associated with oil and protein amount can accelerate hazelnut breeding. In the future this study can contribute to develop new genotypes with enhanced nutritional value.

ÖZET

FINDIKTA (*Corylus avellana* L.) YAĞ VE PROTEİN İÇERİĞİ İÇİN İLİŞKİLENDİRME HARİTALAMASI

Avrupa fındığı (*Corylus avellana* L.) hem besin içeriği hem de ekonomik açıdan önemli bir türdür. Besin içeriği zengin olmasından dolayı, fındık tüketimi insan sağlığını birçok yönden destekler. Bu besin değerleri kantitatif özelliklerdir ve birden fazla gen tarafından kontrol edilmektedir. Dünya fındık üretiminin büyük çoğunluğu Türkiye tarafından sağlanmaktadır. Bu nedenle ülkemiz için daha yüksek besin kalitesine sahip fındıkların geliştirilmesi önemlidir. Bu çalışmadaki hedef; fındık genotiplerindeki genetik çeşitliliği, fındığın yağ ve protein içeriğindeki değişim ile ilişkilendirerek kantitatif lokusları bulmaktır. Bu amaçla 96 genotipin yağ ve protein içeriği ölçülmüştür. 30 SSR primeri kullanılarak genotipik veriler elde edilmiş ve bu analiz 407 polimorfik alel ile sonuçlanmıştır. Allelik verilere göre 96 genotipteki ortalama benzerlik değeri 0.52'dir. Popülasyon yapısı analizi sonucunda popülasyon üç kümeye ayrıldı. Küme 1 30, Küme 2 30 ve Küme 3, 16 bireye sahipti. Geriye kalan 20 genotip ise herhangi bir gruba atanamamıştır. Allelik ve fenotipik veriler arasında bir ilişki bulmak için TASSEL 2.1 yazılımı kullanılmıştır. Toplam beş lokus, yağ içeriği ile anlamlı derecede ilişkili bulunmuştur. Yağ içeriği için en önemli sonuç B628-307 lokusuna aittir ($p = 0.0002$, $r^2=0.145$). Protein içerik ilişkisi için dört lokus tespit edilmiştir. Bunlar arasında en anlamlı olanı A613-153'tür ($p = 0.003$, $r^2= 0.088$). Bu çalışmadan elde edilen veriler; fındıkta yağ ve protein miktarı ile ilişkili tanımlanmış lokusların fındık ıslahında kullanılması ve fındık ıslahını hızlandırması, sonuç olarak besin değeri daha yüksek fındık çeşitlerinin geliştirilmesine katkıda bulunabilecektir.

TABLE OF CONTENTS

LIST OF FIGURES...vii

LIST OF TABLES viii

CHAPTER 1. INTRODUCTION 1

1.1. European Hazelnut (*Corylus avellana* L.) 1

1.2. Nutritional Value and Health Benefits of Hazelnut 3

1.2.1 Lipids and Fatty Acids 4

1.2.2. Proteins and Amino Acids 5

1.2.3. Vitamins..6

1.2.4. Minerals...6

1.2.5. Carbohydrates and Dietary Fiber..7

1.3. DNA Markers..8

1.3.1. SSR (Simple Sequence Repeats) 8

1.4. Marker-Assisted Breeding (MAB) 9

1.5. QTL (Quantitative Trait Loci) 10

1.6. Association Mapping (Linkage Disequilibrium Mapping)..................11

1.7. Aim of the Study 14

CHAPTER 2. MATERIALS AND METHODS...15

2.1. Materials…………………………………………………………..….15

2.1.1. Plant Materials...15

2.2. Methods 18

2.2.1. Genotypic Evaluation of Population...18

2.2.1.1. DNA Extraction 18

2.2.1.2. SSR Analysis 18

2.2.1.3. Genetic Diversity Analysis...21

2.2.1.4. Population Structure Analysis 21

2.2.2. Phenotypic Evaluation of Population 21

2.2.2.1. Total Oil Extraction 22

2.2.2.2. Protein Extraction 22
2.2.3. Association Analysis 23

CHAPTER 3. RESULTS AND DISCUSSION 24

3.1. Phenotypic Evaluation of Population 24
3.1.1. Total Oil Content 24
3.1.2. Protein Content 25
3.2. Genotypic Evaluation of Population 29
3.2.1. SSR Analysis 29
3.2.2. Genetic Diversity Analysis 31
3.2.3. Population Structure Analysis 32
3.3. Association Analysis 40

CHAPTER 4. CONCLUSION 44

REFERENCES 45

LIST OF FIGURES

Figure **Page**

Figure 1.1. Average hazelnut production around the world (A) and in Turkey (B) from 1990 to 2018 (Source: FAO) 2
Figure 1.2. Steps for Association Mapping (Modified from Abdurakhmonov et al. 2008; Rahim et al. 2018) 13
Figure 3.1. Frequency distributions of oil and protein content 28
Figure 3.2. PCoA for 96 hazelnut accessions. 32
Figure 3.3. Dendrogram showing genetic diversity of 96 hazelnut accessions based on 407 SSR alleles. 33
Figure 3.5. Delta K values for each number of K 35
Figure 3.6. Population structure graph for K=3 36
Figure 3.7. Correspondence of genetic diversity and population structure analysis 41

LIST OF TABLES

Table **Page**

Table 2.1. Turkish hazelnut accessions used in this study 15

Table 2.2. SSR markers used in this study 19

Table 3.1. Total oil and protein content of accessions 26

Table 3.2. The number of total and polymorphic fragments, average GD value for each marker 29

Table 3.3. Estimation of ΔK value based on Ln probability 34

Table 3.4. Hazelnut accessions and their assigned subpopulations and dendrogram clusters 37

Table 3.5. SSR loci associated with oil and protein contents of hazelnuts in the GLM (Q) model.. 42

CHAPTER 1

INTRODUCTION

1.1. European Hazelnut (*Corylus avellana* L.)

Since ancient times hazelnut has been an important food for human beings. Nowadays the use of hazelnut in the cosmetics and pharmaceutical industry is becoming common. Hazelnut is spread over a wide area in the temperate climate zone of the northern hemisphere. The hazel tree belongs to the Betulaceae family and the genus *Corylus* L. This genus contains 9 to 25 species some of which are trees and others are shrubs. Two species are especially common in Europe (*C. avellana* L.) and Asia Minor (*C. colurna* L.). *C. avellana* L, also known as European hazelnut, has a high shrub form. *C. colurna* L. is called the Turkish hazel tree (Johnson 2019; İslam 2019).

C. avellana is wind pollinated, monoecious, and dichogamous; female and male flower exist on the same tree and bloom at different times. Male flowers, catkins, are cylindrical and produce more than 5 million pollen grains. Seven to 12 female flowers are found inside the vegetative bud. Female flowers appear during the winter, with fertilization of the ovule completed between mid-May and early June. Five months after pollination, nuts become fully mature in September or October. Hazelnut is 3 to 7 m tall and its leaves are wide and simple. Due to its shrub-like structure, it has multiple stems generated by the basal shoot. It is diploid with 11 basic chromosomes (2n = 22) and has a relatively small genome among trees with ~378 Mb of DNA (Johnson 2019).

The European hazelnut is one of the most economically important nut species in the world. In 2018, nearly one million tons of hazelnut were produced worldwide. Turkey is the primary producer and supplies around 70 % of total hazelnut production with 515.000 tons grown on 728.381 ha in 2018. Other important hazelnut producers are Italy, the USA, Spain, Georgia, Azerbaijan. The average hazelnut production by country and the hazelnut production of Turkey are given in Figure 1.1 for 1990 to 2018(Bhattarai, 2015; FAO, 2018). Among other crops, hazelnuts have a unique place for Turkey. In the Black Sea region about 400.000 families depend only on hazelnut

production and about four million people are included in production. Therefore hazelnut is strategically important as it is the only source of income for many lives (Anil et al., 2016).

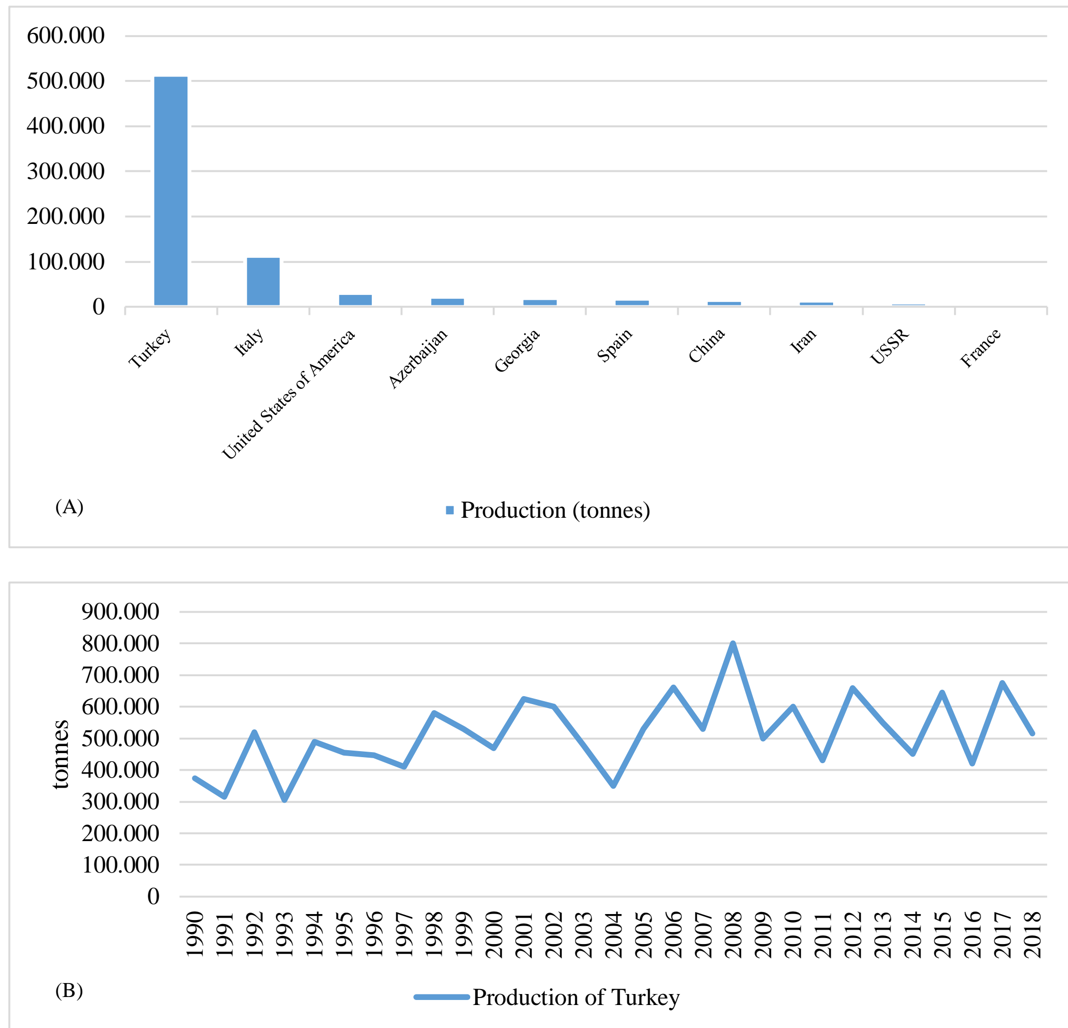


Figure 1.1. Average hazelnut production around the world (A) and in Turkey (B) from 1990 to 2018 (Source: FAO)

In Turkey, Tombul is the most important cultivar with 19 other cultivars developed in the country to date. Cultivars can be divided into three pomological groups. The first group has round nuts: Uzunmusa, Palaz, Okay28, Foşa, Kalınkara, İncekara, Tombul, Cavcava, Kargalak, Allahverdi, Çakıldak, Kan, Giresun Melezi, Kara and Mincane. The second group has pointed nuts: Sivri and İncekara. The third group has long nuts: Yuvarlak Badem and Yassı Badem. Turkish hazelnut cultivars are divided into two main quality classes: Giresun and Levant quality. Giresun quality

includes Tombul, with its thin shell and high fat. Levant quality contains mixed cultivars, and these cultivars are lower quality than Giresun (İslam, 2018).

1.2. Nutritional Value and Health Benefits of Hazelnut

Hazelnut kernels have an important role in human nutrition and health due to their nutritional value. Among all nut types, hazelnut is a highly energetic food because it contains high amounts of fat, carbohydrate, and protein and also contains other important nutrients like vitamins, minerals, phytosterols and antioxidants (Bacchetta et al. 2013).

It has been reported that chronic disease prevalence within the population can be lowered by consuming hazelnut oil which is rich in mono-unsaturated fatty acid (MUFA). Consumption of MUFA also prevents or heals cardiovascular disease, metabolic syndrome, and improves blood pressure, glucose levels, blood lipid composition, decreases obesity risk (Topkafa et al. 2019). Hazelnut also has a cardio-protective benefit due to its high arginine level which helps to relax blood vessels. In addition, hazelnut has a beneficial lysine/arginine value and atherosclerosis development has been reduced with the help of low lysine/arginine ratio diet in laboratory animals (Xu & Hanna, 2010).

Hazelnut is rich in both major and minor minerals (potassium, phosphorus, magnesium, calcium, manganese, iron, zinc, chromium, sodium, selenium). For this reason, the consumption of hazelnut is associated with improved health. Minerals are crucial for humans and each one has its own health benefits (Alasalvar et al. 2009). Potassium acts like an electrolyte and is important for the proper functioning of the heart, nervous system, kidney, and muscle tissues (Alaviani et al. 2012). Phosphorus plays a role in energy processes and builds up bones and cells. Magnesium reduces cardiovascular disease (Cosmulescu et al. 2013). The bones and the teeth contain a large amount of calcium which is also found in the circulatory system to prevent hemorrhages. Manganese can act as a coenzyme and has a role in the structure of proteins, regulates blood sugar, and is also required for healthy nerves and immune systems (Alaviani et al. 2012; Cosmulescu et al. 2013). Iron is another essential mineral that is part of hemoglobin, myoglobin, and many enzymes. Bones and muscles contain a high amount of zinc which also can play a role as a constituent of metabolic enzymes.

Chromium helps regulate glucose metabolism and is a cofactor of insulin. Sodium, an electrolyte, has an essential role in muscle contraction and enzyme. Additionally, sodium is important for fluid maintenance of the body, heart performance, glucose absorption, and regulation of osmosis (Alaviani et al. 2012). Selenium is crucial for human health. It decreases the risk of certain types of cancer, coronary heart disease (CHD), reduces oxidant species, plays an antioxidant role, preserves the elasticity of blood vessels, and protects the cell membrane by neutralizing free radicals (Alasalvar et al. 2010). Based on all these examples, it is clear that minerals are vital for maintaining a healthy and well balanced human body (Cosmulescu et al. 2013).

Hazelnut also contains dietary fiber which is not a nutrient but contributes to good health. Consuming high fiber can lower the body mass index, serum cholesterol, the risk of certain types of cancer and CHD, improve gastrointestinal function, and increase glycemic and weight control (Alasalvar et al. 2010).

1.2.1 Lipids and Fatty Acids

Lipid is the most abundant nutrient in hazelnut and makes up 40 to 70% of total weight (Topkafa et al. 2019). Hazelnut oil and its quality are important because oil is used for cooking, deep frying, salad dressing, and as an ingredient in skincare products (Alasalvar et al. 2003).

Many studies have been published about the lipid composition of hazelnut. These studies show that hazelnut lipid composition can be affected by variety, geographical origin and location, ecology, and growing conditions (Balta et al. 2006). Hazelnut lipid fraction consists of non-polar (98.8%) and polar (1.2%) components (Ciemniewska-Zytkiewicz et al. 2015). Triacylglycerol is the major non-polar lipid class of tree nut oils but also fatty acids, mono and diacylglycerols belong to major lipid constituents. Glycolipids and phospholipids are polar components (Miraliakbari and Shahidi 2008). In terms of fatty acids, hazelnut contains a higher amount of MUFA and lower amounts of polyunsaturated fatty acid (PUFA) and slightly saturated fatty acid (SFA) (Amaral et al. 2006). Oleic acid and linoleic acid, whose contents are conversely related, are the main fatty acids in hazelnut and represent 89% of total fatty acid, followed by palmitic and stearic acids (Bacchetta et al. 2013). The minor lipid fraction is composed of phytosterols and tocopherols. Among tree nut oils, hazelnut has the

highest β-sitosterol and α-tocopherol contents; α-tocopherol is also known as vitamin E (Miraliakbari and Shahidi 2008).

1.2.2. Proteins and Amino Acids

One of the characteristics that affects hazelnut kernel quality is the amount of protein (Balta et al. 2006). After lipid, protein is the second major constituent of hazelnut with content ranging from 10% to 24% (Köksal et al. 2006). Plant proteins are generally considered incomplete proteins because, compared to animal proteins, they do not have one or more essential amino acids (Iyanapathirana and Hshima 2003).

Hazelnut is abundant in both essential and non-essential amino acids. According to Iyanapathirana & Hshima (2003), three non-essential and essential amino acids contributed to 30.9% and 44.9% of total amino acids within the Turkish Tombul hazelnut. The most abundant amino acid is glutamic acid, followed by aspartic acid and arginine. Other essential amino acids are also found in Tombul but lysine and tryptophan occur at lower levels. In terms of free amino acids, a high amount of arginine and glutamic acid together with a lesser amount of alanine, asparagine, and aspartic acid are found in Tombul. Nearly 75% of the total free amino acids are comprised of these five amino acids. The distinct taste and flavor of hazelnut come from individual free amino acids. Tombul had the highest amount of free amino acids compared to other cultivars. The formation of color and aroma during roasting are also affected by free amino acids. Amino acid content varies among hazelnut genotypes. Growing conditions and geographical regions also affect the amount and composition of protein in hazelnut (Köksal et al. 2006).

A study performed by Köksal et al. (2006) with seventeen Turkish hazelnut cultivars found methionine in all varieties. According to Xu & Hanna (2010), a Nebraska hybrid hazelnut contains eight essential amino acids: histidine, isoleucine, threonine, leucine, phenylalanine, lysine, methionine and valine. Among these, lysine and threonine content were significantly lower compared to others. In addition to these essential amino acids, glutamic acid, tyrosine, alanine, serine, aspartic acid, glycine, cysteine and arginine were found in the hybrid. All of the studies show that hazelnut is a good source of protein and amino acids and is beneficial when included in the human diet.

1.2.3. Vitamins

One of the most important features of hazelnut is its vitamin content (Özdemir et al. 2001). Hazelnut is the best source of vitamin E among tree nuts. Vitamin E or α-tocopherol (the active form of vitamin E) is the most popular and powerful antioxidant in the body and protects cells via scavenging free radicals (Iyanapathirana and Hshima 2003).

According to Iyanapathirana & Hshima (2003), Turkish Tombul hazelnut contained a high amount of vitamin E; relatively less pantothenic acid, ascorbic acid and vitamin B_2; and small amounts of vitamin B_1, vitamin B_6, biotin, and folate. Among 11 tree nuts (hazelnut, cashew, almond, macadamia, walnut, chestnut, pistachio, pinenut, pecan, coconut, brazil nut) hazelnut has the highest amount of vitamin E, biotin and folate. Vitamin B_{12} and carotene were not detected in Tombul (Holland et al. 1995).

Köksal et al.(2006) found the highest levels of niacin, vitamin B_1, vitamin B_2, vitamin B_6, and ascorbic acid within the Mincane, Yuvarlak Badem, Tombul, Kan and Yassı Badem cultivars, respectively. The insoluble vitamins retinol, α-, δ-, and γ-tocopherol were all detected in the cultivars, however, α-tocopherol was highest in Tombul. Additionally, no cultivar contained biotin.

Özdemir et al. (2001) examined the storage stability index in commercial cultivars and hybrid varieties via α-tocopherol content and fatty acid composition. The equation, α-tocopherol x saturated fatty acids/unsaturated fatty acids, is used to estimate the storage stability of hazelnut. The storage stability index was higher when the ratio was lower. Commercial cultivars had a relatively higher stability index while Mincane had the lowest stability index due to its low α-tocopherol content. The authors concluded that soil composition, irrigation, and fertilization may affect the composition of varieties which then results in stability differences.

1.2.4. Minerals

Among tree nuts, hazelnut has a significant component in terms of mineral content. A total of 24 minerals have been found in hazelnut, 13 of which are essential. According to Cosmulescu et al. (2013) daily consumption of 100 g of hazelnut meets 94% of the Requirement Daily Allowance (RDA) for iron, 70% for magnesium, 55%

for phosphorus, 22% for zinc, 13% for potassium, 10% for calcium, 5.6% for chromium with the amounts of manganese and copper in 100 g hazelnut much higher than the daily requirement.

Seyhan et al. (2007) evaluated six essential minerals (potassium, zinc, phosphorus, iron, calcium, manganese) in three different Turkish hazelnut cultivars (Tombul, Palaz, Badem) from early development to maturity. They found that mineral content was significantly reduced as development passed through the early, middle, and harvest stages. They concluded that mineral content and composition of hazelnut were affected by variety, climate, soil composition, fertilizer use, harvest year, method of cultivation, irrigation, and geographical origin.

Alaviani et al. (2012) investigated the mineral content of Iranian hazelnut, almond, and peanut. Based on amounts, hazelnut and almond mineral contents were similar containing potassium, calcium, iron, zinc, sodium, manganese, and chromium. These researchers also compared the Turkish cultivar: Yassı Badem and the Iranian cultivar Alamut. Both contained a high amount of potassium and calcium and a low amount of manganese. However, the Turkish cultivar, Tombul, had more manganese than Alamut. Alasalvar et al. (2009) worked with five Turkish cultivars (Tombul, Yassı Badem, Karafındık, Sivri, Ham). Twelve minerals were determined; potassium was most abundant followed by phosphorus, calcium, and magnesium. Non-commercial Levant quality Ham variety was a great source of manganese. Ham had 8.7 fold more manganese than the Giresun quality cultivar Tombul and Levant quality cultivar Karafındık. Ham also had 3.2 and 2.0 fold higher manganese level than the Levant quality cultivars Badem and Sivri. The five cultivars contained all essential minerals. All were excellent sources of manganese and copper and good sources of chromium, phosphorus, iron, zinc, and magnesium.

1.2.5. Carbohydrates and Dietary Fiber

Carbohydrates constitute 10-22% of hazelnut depending on variety and dietary fiber corresponds to 85% of total carbohydrates (Köksal et al. 2006; Xu et al. 2012). Iyanapathirana & Hshima (2003) found that Tombul' s total carbohydrate content was 17.3% and its total dietary fiber content was 12.9%. Among dietary fiber, 9.7% was insoluble and 2.2% was soluble. According to a fresh weight basis, the highest amount

of dietary fiber was found in Tombul. Seyhan et al. (2007) evaluated that the different stages of three cultivars (early, middle, harvest), and found that there were no significant differences in terms of total carbohydrate content of Tombul, Badem and Palaz.

1.3. DNA Markers

A nucleotide sequence with a specific location is known as a DNA marker and these markers serve as ‘signs’ or ‘flags’ in the genome. The ideal DNA marker should meet several criteria: distribution across the genome, high polymorphism rate, co-dominance (the ability to distinguish heterozygotes from homozygotes), easy detection of a large number of genotypes with a large number of loci and lastly low cost (Jiang 2013).

Types of DNA markers include amplified fragment length polymorphism (AFLP), restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), simple sequence repeat or microsatellite (SSR), expressed sequence tag (EST) and single nucleotide polymorphism (SNP) markers. Currently, SSR and SNP markers are most widely used in plant breeding (Schulman 2007).

1.3.1. SSR (Simple Sequence Repeats)

SSRs are generally known as microsatellites. These are genereally 1 to 6 base pair tandem repeat motifs and distributed in the entire genomes of both eukaryotic and prokaryotic organisms (Mason 2015). They can be found in both non-coding or coding sequences and can appear in the mitochondria and chloroplast genomes (Kalia et al. 2011). Their role can be determined depending on where they appear. Activation or repression of a protein can be affected when SSRs are located in the coding region. Frameshift mutations due to SSRs can lead to the expression of truncated proteins. When they are located in non-coding regions, for example in 3’-UTR regions, alterations can cause replication slippage and gene silencing. In contrast, alteration to a SSR in a 5’-UTR region can induce gene transcription and translation because 5'-UTRs are known as protein binding sites. SSRs in introns also can lead to alternative splicing

and transport of mRNAs, gene silencing, transcription or translation of a gene (Kalia et al. 2011; Vieira et al. 2016).

SSRs are classified into six groups based on motif length: mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats. For example, $(AT)_n$ is the most common repeat motif in plants while animals have $(AC)_n$ motif. Tri- and tetranucleotide repeats are also frequently seen in plants, among them $(AAT)_n$ and $(AAG)_n$ motifs are prevalent (Gupta et al. 1996; Kalia et al. 2011). SSRs can be categorized as perfect, imperfect, and compound depending on the positioning of nucleotides in the repeat motifs. Mechanisms including recombination of double-stranded DNA, slippage of single-stranded DNA, transposition, mismatches, and double-stranded break repair may contribute to the formation and evolution of SSRs (Nadeem et al. 2018; Wang et al. 2009). For this reason, SSRs have a high mutation frequency with a rate of 1×10^{-2} to 10^{-3}mutations per locus, per gamete, per generation which is approximately 10^6 times higher than non-repetitive DNA (Mason 2015). The high mutation rate results in a high polymorphism rate. In addition, SSRs are co-dominant, multi-allelic and highly abundant. Therefore SSRs are ideal markers for the identification of plant varieties and for mapping (Vieira et al. 2016).

1.4. Marker-Assisted Breeding (MAB)

DNA markers have an essential role in plant breeding, ranging from quantitative trait locus (QTL) analysis, construction of linkage maps, marker-assisted breeding (MAB), to genetic diversity studies. Conventional breeding program are based upon identifying and selecting desirable traits or superior phenotypes within a segregating population derived from crosses. Such programs can be difficult because some phenotyping procedures are time-consuming and expensive, while others are unreliable for particular traits. MAB is integrated into conventional breeding but in comparison to conventional breeding, MAB requires DNA marker detection and selection, which requires more complex equipment and instruments (Jiang 2013). The most important aim of both conventional and MAB is to ensure sustainable crop improvement (Fu 2015).

MAB has three purposes: determining the most desired individuals in a segregating population based on allelic composition, identifying and collecting superior

allele(s) across the population, and breaking the linkage of superior alleles with undesirable loci. MAB is most frequently implemented for the improvement of quantitative or qualitative traits, increasing yield, disease and pest resistance, low-temperature stress tolerance, drought stress tolerance, and salinity stress tolerance (Francia et al. 2005).

In modern breeding, genetic diversity analysis is useful for parental selection and can be performed via screening numerous markers and genome sequencing. Genome-wide association studies (GWAS) can be implemented for QTL mapping in plants to identify marker-trait associations. In addition, genome sequencing enables the discovery of new allelic variants in cultivated populations, the detection of favorable mutations and targeted modifications of specific genes. These approaches are done with the help of bioinformatics and automated phenotyping, as well as Next Generation Sequencing (NGS) technology. Combinations of advanced technology can enhance and revolutionize the genetic improvement of crop species (Barabaschi et al. 2015).

1.5. QTL (Quantitative Trait Loci)

Numerous agriculturally important variations among plants such as environmental and stress tolerance, quality, disease resistance, and productivity are quantitative traits. These traits are regulated by the overall effect of multiple genes and the environment. The phenotypic variation of these traits does not represent discrete values, for this reason, these traits are called polygenic, multifactorial, complex, or continuous traits. Within the genome, polygenic traits are associated with quantitative trait loci known as QTL (Dreisigacke et al. 2016).

The identification of QTL is based upon correlation between phenotype and marker genotype. Depending on the genotype of a particular marker within the population, the mapping population can be divided into different genotypic groups. Trait (phenotypic) means for each genotypic group are then compared to identify significant differences which indicate that a locus controlling the trait is linked to that DNA markers. This process is repeated for numerous markers encompassing the entire genome and also allows estimation of the phenotypic effects of each QTL to the trait of interest (Mackay 2001). QTL analysis is used as a precursor to MAB in order to identify

markers that are linked to the most significant loci controlling the trait of interest (Mohan Jain and Brar 2009).

1.6. Association Mapping (Linkage Disequilibrium Mapping)

In general, genetic maps are created based on one of two strategies: (1) use of bi-parental populations which is called QTL mapping, gene tagging or linkage mapping and (2) use of natural populations, germplasm or diverse collections which is called association mapping or linkage disequilibrium (LD) mapping (Abdurakhmonov & Abdukarimov, 2008; Painter et al. 2011).

There are several steps for construction of a linkage map. First, an experimental population like backcross (BC), doubled haploid (DH), F_2, near-isogenic line (NIL), or recombinant inbred line (RIL) population must developed via hybridization of diverse parental genotypes. Then this population containing a great number of segregating lines should be measured for the trait of interest. In addition, genotypes of the parental lines and segregating population should be determined using a lot of polymorphic DNA markers. The genotypic data are used to calculate recombination rates between marker loci and to create a linkage map that represents the order and linkage of molecular markers. QTL regions affecting the traits of interest are then identified via statistically correlation of genotypes with the phenotypic values for the traits of interest (Abdurakhmonov and Abdukarimov 2008). The first QTL analysis in European hazelnut was conducted using a F1 population. A total of fifteen QTL associated with leaf budburst, vigor and sucker habit were identified with one of them responsible for 50% of the phenotypic variance for leaf budburst (Beltramo et al. 2016). Marinoni et al. (2018) used the same F1 population with a more saturated molecular map and found an additional 29 QTL associated with leaf budburst including the previously found major QTL. A full-sib line was used to detect QTL responsible for Eastern Filbert Blight disease resistance. One major QTL was determined accounting for 72.8% of phenotypic variance (Honig et al. 2019).

Another important genetic mapping method is association mapping which is also called linkage disequilibrium mapping. An understanding of this method requires an understanding of the concept of linkage. Physical linkage of loci means that alleles at the two loci have a tendency to be inherited together through a physical connection. In

contrast if the loci are unlinked and in the equilibrium state, by chance one allele can be found at one locus and the other allele can be found at another locus independently. In linkage disequilibrium, non-random associations are seen between alleles at different loci. Recombination events occur during meiosis and break these non-random associations. Eventually every possible region of a chromosome experiences recombination and, as a result, a pair of markers moves from linkage disequilibrium to linkage equilibrium (Bush & Moore, 2012; Gore et al. 2008).

In contrast to QTL mapping, association mapping has been accepted as a more efficient way to understand the genetic basis of a complex trait. Association mapping uses existent variation (historical recombination that happened during evolutionary history) in natural populations and does not require the generation of experimental populations. Therefore association mapping populations are much more diverse than QTL mapping populations (Fleury et al. 2012; Hall et al. 2010). The important factors that affect the resolution of association maps are markers and their non-random associations, marker density, population size and structure, and extent of LD across the genome. For example in a small population, LD increases because genetic drift causes the loss of rare combinations of alleles. However, LD decays more quickly in outcrossing species than selfing species and outcrossing individuals are more likely to be heterozygous. Population structure can be a limitation for association mapping. Factors like admixture, selection, and genetic drift can affect population structure by affecting LD between traits and markers. LD extent is also related to marker density. If LD is more extensive than marker density, high marker density is not needed because the marker-trait association can be detected. It can be assumed that if a pair of markers has strong LD, all markers between them have also strong LD. If marker density is low in the genome, some loci may not have strong LD with markers. This situation can decrease the power of association studies in detecting QTL (Fleury et al. 2012).

There are several steps to construct association maps. Firstly, a population representing a high degree of genetic diversity should be selected and measuring of the phenotypic trait of the selected population should be performed with multiple replicates. Then, genotyping of the population should be done with appropriate molecular markers. Using this genotypic data,the extent of LD, genetic differentiation within-population individuals (population structure) and kinship (coefficient of relatedness between pairs of each individual within a sample) should be measured. Finally, an appropriate statistical program should be used to correlate phenotypic and genotypic data based on

quantification of LD extent and population structure (Abdurakhmonov and Abdukarimov 2008). Steps for association mapping are summarized in Figure 1.2. In this way, QTLs controlling a trait of interest can be identified. Association maps facilitate superior allele mining for trait improvement. Although the genomic sequence of hazelnut has been determined, only a few studies have been published for association mapping of hazelnut.

Using a natural population from Slovenia, Ozturk et al. (2017) found 49 loci associated with nut and kernel quality traits. In Turkish germplasm including wild genotypes, landraces and cultivars, 78 loci were detected for nut, kernel, quality, yield and shell thickness traits (Frary et al. 2019a). In addition, the same material was used to detect 143 QTL associated with other agro-morphological traits such as plant habit, leaf, involucre and inflorescence characteristics (Frary et al. 2019b).

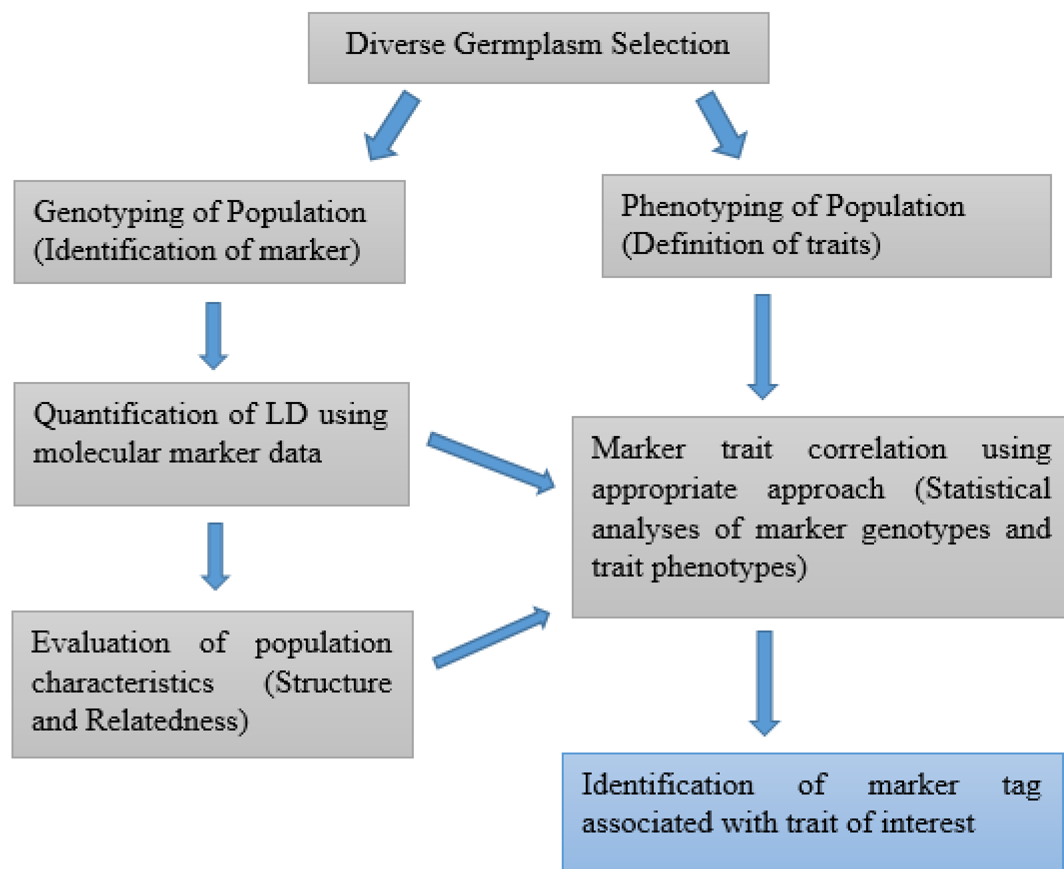


Figure 1.2. Steps for Association Mapping (Modified from Abdurakhmonov et al., 2008; Rahim et al. 2018)

1.7. Aim of the Study

Hazelnut is an agricultural product of Turkey with great economic importance because Turkey accounts for 70% of the total hazelnut production worldwide. The consumption of hazelnut is highly beneficial for human nutrition and health. This nutritional value primarily comes from the oil and lipid fraction, followed by its protein content, and other nutritional factors like carbohydrates, minerals, vitamins, and dietary fiber. There are a limited number of published studies about QTL mapping in hazelnut and association mapping for nutritional traits has not yet been done. We aimed to construct an association map for nutritional traits in hazelnut to provide information that can be used toward the improvement of nut quality. Phenotypic and genotypic analyses were carried out to find polymorphism within the hazelnut population for the detection of QTLs. Accessions originated from mainly the Black Sea region of Turkey and were comprised of 20 cultivars, 23 landraces and 53 wild genotypes. Total oil and protein content were measured and considered to be phenotypic traits. Thirty SSR markers were used to evaluate the genetic diversity of the population. In this way, we characterized germplasm and revealed their favorable traits, identified genomic regions associated with total oil, protein content and determined the number of QTLs controlling each trait and their relative importance. The resulting information can be used for the selection of germplasm and the breeding of nutritional traits in hazelnut.

CHAPTER 2

MATERIALS AND METHODS

2.1. Materials

2.1.1. Plant Materials

Plant materials were provided by the Hazelnut Research Institute, Giresun. A total of 78 accessions which were named as core set by Öztürk et al. (2017), were chosen based on their representing the maximum genetic diversity of Turkish hazelnut germplasm. Sixteen wild genotypes and four landraces were added the core set to expand the genetic diversity. This population contains 20 Turkish hazelnut cultivars, landraces, and wild genotypes. The individuals used in this study are listed in Table 2.1.

Table 2.1. Turkish hazelnut accessions used in this study

No	Accession Name	Type of Material	Origin
1	FAI105	Landrace	Giresun
2	FAI118	Wild	Ordu
3	FAI132	Wild	Unknown
4	FAI157	Landrace	Giresun
5	FAI166	Wild	Giresun
6	FAI181	Wild	Giresun
7	FAI183	Wild	Ordu
8	FAI195	Landrace	Giresun
9	FAI200	Landrace	Giresun
10	FAI213	Landrace	Ordu
11	FAI219	Landrace	Giresun
12	FAI233	Wild	Trabzon
13	FAI244	Landrace	Trabzon
14	FAI252	Landrace	Giresun

(cont. on next page)

Table 2.1 (cont.)

15	FAI253	Landrace	Giresun
16	FAI258	Wild	Trabzon
17	FAI270	Landrace	Rize
18	FAI281	Wild	Unknown
19	FAI283	Wild	Giresun
20	FAI286	Landrace	Giresun
21	FAI335	Wild	Giresun
22	FAI338	Landrace	Giresun
23	FAI351	Wild	Giresun
24	FAI359	Wild	Giresun
25	FAI377	Wild	Giresun
26	FAI391	Landrace	Giresun
27	FAI401	Wild	Unknown
28	FAI417	Wild	Unknown
29	FAI421	Landrace	Ordu
30	FAI431	Wild	Trabzon
31	FAI433	Landrace	Giresun
32	FAI434	Wild	Unknown
33	FAI436	Wild	Unknown
34	FAI444	Wild	Unknown
35	FAI453	Wild	Unknown
36	FAI469	Landrace	Giresun
37	FAI471	Wild	Unknown
38	FAI511	Wild	Unknown
39	FAI530	Wild	Unknown
40	FAI532	Wild	Unknown
41	FAI533	Wild	Unknown
42	FAI551	Wild	Unknown
43	FAI558	Wild	Unknown
44	FAI580	Wild	Unknown
45	FAI586	Wild	Unknown
46	FAI587	Wild	Unknown
47	FAI588	Wild	Unknown
48	FAI589	Landrace	Giresun
49	FAI590	Landrace	Giresun
50	FAI595	Wild	Unknown
51	FAI598	Wild	Unknown

(cont. on next page)

Table 2.1 (cont.)

52	FAI601	Wild	Unknown
53	FAI604	Landrace	Erzurum
54	FAI608	Wild	Unknown
55	FAI609	Wild	Unknown
56	FAI611	Wild	Unknown
57	FAI616	Wild	Unknown
58	FAI617	Wild	Unknown
59	Acı	Cultivar	Giresun
60	Allahverdi	Cultivar	Giresun
61	Cavcava	Cultivar	Giresun
62	Çakıldak	Cultivar	Giresun
63	Foşa	Cultivar	Giresun
64	Giresun Melezi	Cultivar	Giresun
65	Ince Kara	Cultivar	Giresun
66	Kalın Kara	Cultivar	Giresun
67	Kan	Cultivar	Giresun
68	Kara	Cultivar	Giresun
69	Kargalak	Cultivar	Giresun
70	Kuş	Cultivar	Giresun
71	Mincane	Cultivar	Giresun
72	Okay28	Cultivar	Giresun
73	Palaz	Cultivar	Giresun
74	Sivri	Cultivar	Giresun
75	Tombul	Cultivar	Giresun
76	Uzun Musa	Cultivar	Giresun
77	Yuvarlak Badem	Cultivar	Giresun
78	Yassı Badem	Cultivar	Giresun
79	FAI108	Landrace	Giresun
80	FAI129	Landrace	Giresun
81	FAI165	Wild	Giresun
82	FAI177	Landrace	Ordu
83	FAI190	Wild	Giresun
84	FAI201	Wild	Unknown
85	FAI221	Wild	Giresun
86	FAI250	Wild	Giresun
87	FAI266	Wild	Unknown
88	FAI284	Wild	Giresun

(cont. on next page)

Table 2.1 (cont.)

89	FAI287	Landrace	Giresun
90	FAI319	Wild	Unknown
91	FAI405	Wild	Unknown
92	FAI411	Wild	Unknown
93	FAI419	Wild	Unknown
94	FAI521	Wild	Unknown
95	FAI522	Wild	Unknown
96	FAI577	Wild	Unknown

2.2. Methods

2.2.1. Genotypic Evaluation of Population

For the genotypic evaluation of the population, genomic DNA (gDNA) was first obtained. Then SSR amplification was done and presence/absence, binary data were obtained. Using these data genetic diversity (clustering, PCoA, and dissimilarity), population structure analysis were performed.

2.2.1.1. DNA Extraction

Total genomic DNA was extracted from leaves and catkins according to Fulton et al.(1995) with minor modifications. The quantity and quality of gDNA were assessed by spectrophotometer (Thermo Scientific, Multiskan GO) and also the integrity of gDNA was checked with agarose gel electrophoresis. All gDNA samples were stored at -20°C.

2.2.1.2. SSR Analysis

A total of 30 SSR primers selected from Gürcan et al. (2010) were applied to the hazelnut population (Table 2.2). Polymerase Chain Reaction (PCR) was performed with

a total volume of 20 µl. According to their quality and quantity, all DNA samples were diluted to the concentration of 20 ng/µl. Amplification mixture contained: 2 µl 10X buffer (50 mM KCl, 10 mM Tris-HCl, 1.5 mM $MgCl_2$, pH: 8.3), 0.75 µl primer (10 pmol each), 0.5 µl dNTPs (0.2 mM), 1.5µl $MgCl_2$, 0.5 µl 0,6 U *Taq* polymerase, 1 µl 20 ng DNA and 13 µl double-distilled water.

PCR conditions were: one step of 5 min at 94°C for denaturation; followed by 30 cycles with 30 sec at 94°C, 30 sec for annealing at 55°C, 30 sec for the extension at 72°C; and final extension step of 5 min at 72°C in GeneAmp PCR system 9700 (Perkin Elmer Applied Biosystems). After PCR reactions, DNA fragments were separated using a Fragment Analyzer™ Automated CE System (Advanced Analytical Technologies). This capillary electrophoresis system has a high resolution to discriminate alleles with a minimum of 3 bp difference. DNA fragments were scored presence (1) or absence (0) because SSR markers yielded more than two fragments and it was not possible to identify allelic fragments.

Table 2.2. SSR primers used in this study

Marker Name	Forward Primer (5' to 3') Reverse Primer (5' to 3')	Motif
A601	TTACATGGTTCGGCAATGTG AGATGGGAGCAGAGTGAACTG	$(AC)_{26}$
A602	AAGAGTGGGGGTGCACTATG GGATTCATGCCTGCGATACT	$(AC)_{16}(AT)_{6}$
A604	GCTCCCGAGGACTTCCAG CCACGACATTTCCCTCTCAG	$(CT)_{16}(CA)_{14}$
A605	CACCCTCAAAACTGTGACGA TGGGTCGCATTCAATAACAC	$(TC)_{15}(CA)_{12}$
A606	CACCTAGCTTGTTGGTGAAGC TGACAATAATTAACCCTACACACTTTG	$(AC)_{12}$
A611	CACTAGCCAGCCCCTTTACA CTGATGCCACAAACACAAGG	$(AC)_{16}$
A613	CACACGCCTTGTCACTCTTT CCCCTTTCACATGTTTGCTT	$(TC)_{13}(CA)_{12}$
A616	CACTCATACCGCAAACTCCA ATGGCTTTTGCTTCGTTTTG	$(AC)_{11}$

(cont. on next page)

Table 2.2 (cont.)

A635	GGATCTGTGGTTGGCTTTTTGGTACTAT TTACCCAATGGATGATGGACTAGCATT	$(AC)_{11}$
A640	TGCCTCTGCAGTTAGTCATCAAATGTAGG CGCCATATAATTGGGATGCTTGTTG	$(CT)_{15}(CA)_{13}$
B602	AAGAGTGGGGGTGCACTATG GGATTCATGCCTGCGATACT	$(TC)_{15}(CA)_{10}$
B603	TGGTGGTGATAGGGAAGGAG TCTTTTCTTCTTCAATCAGACGA	$(CT)_{19}$
B606	TCTTGTGGTTTAGCATACTTCTCG GAAGAAAGCAAGAAGAGAGGAGA	$(ACAT)_6Ns(AG)_{16}$
B612	GCACCTCAAACTCCTTGGAC CCCAAACACACCCTTAGTGC	$(GA)_{14}$
B613	CGCGTTTTGAGTCCCTTTAG CTACCCGCCTGCGAGAAC	$(CT)_{16}$
B625	CGCAAGTCATTGCACATTTT GTGTGCTGTGCTCCTTTGAA	$(TC)_{13}$
B628	AATCCCCTCTAGCCCCATTA CACAGAATATTTGTAATTACCACCACA	$(TC)_7NN(CT)_6$
B631	AATCCCCTCTAGCCCCATTA TTGTGTCTCTTTGTCTTGTAAATCG	$(CT)_6NN(TC)_5Ns(TC)_{14}$
B635	GCATCGCCAAATTATCGTCT CTTCAACAAATCCAGGATGC	$(GA)_{20}$
B640	CTGCATTGATGGATTGGTTG TTAAGAAAGGTACAAGGGCTCTC	$(GA)_{27}$
B641a	CTCCCATGAAATGATTATTCTTAG CAAGCCATCTGTTTTGCTGA	$(AT)_5(TG)_7Ns$
B641b	ATATATATAGGCTGTGTGTGTGTGTG ACAAGCCATCTGTTTTGCTG	$(AG)_{17}Ns(GA)_{12}$
B648	TGAAAGCGCCCAAAACTTAT CTTGCGTCTTTTTGGAGAGC	$(TC)_{17}$
B651	TTTTCTGGAATGTCGCACAG TCTCCTCCTTCCAACAGTGG	$(CT)_{16}$
B652	AGGATGCGTGGTTGTGATTT TGGAGTAGGGTGATGAGAATGA	$(CT)_{20}$
B660	TGTTGTAGCACAACCCTTTCA TGCTAGCAGCAAATGGCTTA	$(TG)_{11}Ns(GA)_{15}$

(cont. on next page)

Table 2.2 (cont.)

B662	CGAAAGATGGACTTCCATGAC CAAGTTGAGATTCTTCCTGCAA	$(TC)_{15}$
B788	TCCCTTTCTCCGTCATCAAC TCGTCACCGTCACCAGATAA	$(CT)_{11}CCC(TCTT)_5(AG)_{16}$
B789	GCCACGTCCAGAATCAAAAT CCTCAGGGCTGAGAAGTTGA	$(AG)_{16}$
CAC-B753	AAGGGTTGTTACCCATGCAC GGTGCATTTAGTGCTTCTGG	$(GA)_{15}$

2.2.1.3. Genetic Diversity Analysis

The genetic diversity of the population was assessed using Darwin 6.0.14 software. The allelic data were used for calculation of the distance between individuals and cluster analysis. In this analysis, the Dice coefficient matrix and Unweighted Neighbor-Joining algorithm were carried out on the population.

2.2.1.4. Population Structure Analysis

Structure 2.3.4 software was used to divide the population into sub-populations and assign individuals to the sub-populations. The structure analysis was run with a burn-in period of 100.000 and 300.000 MCMC repeats. The number of subpopulations (K) was tested for 2 to 20 with 10 iterations for each group. The ad hoc statistic was used to detect the correct estimated number of clusters with Structure Harvester online program (Evanno et al. 2005).

2.2.2. Phenotypic Evaluation of Population

In the phenotypic evaluation of the population, total oil content was detected via n-hexane extraction and protein amount was determined with Bradford assay after Tris-Cl extraction.

2.2.2.1. Total Oil Extraction

Total oil extraction was performed based on Soxhlet solvent extraction using the Soxtherm Rapid Soxhlet machine (AOAC Official Method 948.22). With three replicates for each hazelnut accession, 1 g hazelnut powder was weighed and placed in a cellulose thimble. The cellulose thimble was placed in the extraction beaker, then n-hexane (≥96%, Isolab) was added until the powder was immersed. Extraction occurred with five steps. Hot extraction happened at 150°C with extraction beaker on the hotplate, in the first evaporation as the solvent level decreased, the excess solvent accumulated in the solvent tank. Then samples were rinsed with reflux and condensed solvent accumulated in the extraction beaker. In the second evaporation, the majority of solvent was collected into the rear tank for recovery. Finally extraction beakers were lifted from the hotplate, in this way residual solvent was removed (C. Gerhardt UK, Analytical Systems). When the extraction was completed, oil at the bottom of the extraction beaker was collected with a pipette and placed in a microtube. To minimize loss, n-hexane was added to the bottom of the extraction beaker and the remaining oil was collected. Microtubes were then placed in CentriVap Vacuum Concentrator (Labconco Corporation) at 30°C for 1 h to evaporate any remaining n-hexane. Finally, all samples were weighed and stored at -20°C.

2.2.2.2. Protein Extraction

Total protein was obtained by the method of Angelis et al. (2018) with three replicates per accession. Hazelnuts were ground with a grinder (Knife Mill GRINDOMIX GM 200). Then 10 ml 20 mM Tris-Cl (pH 8.2) extraction buffer was added to 500 mg powders and kept for 1 h at 25°C in shaking incubator (JSSI-100C, JS Research Inc.). Then, samples were placed in an ultrasonic water bath for 10 min (Elmasonic S 120 H, Elma Schmidbauer GmbH) and centrifuged for 15 min at 1734 g at 18°C (Allegra X-15R, Beckman Coulter). At this stage, extracts were separated into three layers: fat, protein, and debris. The protein fraction was collected and filtered with 0.45 μm PTFE filter. Until analysis, all protein extracts were kept at -80°C.

Total protein quantification was performed according to Bradford (1976). The Bradford assay is based on the color change from red to blue of Coomassie Brillant

Blue G-250 (CBBG-250) dye. This change indicates that the dye is bound to protein. This protein-dye complex absorbs maximum light at 590 nm. Dye reagent was prepared by dissolving 100 mg CBBG-250 (Sigma-Aldrich) in 50 ml 95% ethanol followed by addition of 100 ml 85% phosphoric acid. When the dye was completely dissolved, it was diluted to 1 L and filtered via Whatman paper. Afterward, the dye reagent was kept in the dark at 4°C. Protein standards were Bovine Serum Albumin (BSA) and obtained from Sigma-Aldrich. BSA concentrations ranged from 0 to 2 mg/ml. For spectrophotometric measurement dye reagent and diluted protein samples (1:20) were merged in 96 well plate. For every three biological replicates three reading was performed. The standard calibration curve was constructed measuring both absorbance at 590 and 450 (Zor & Selinger 1996).

2.2.3. Association Analysis

Association analysis between marker genotypes and individual trait phenotypes was carried out with TASSEL 2.1 (Trait Analysis by aSSociation, Evolution, and Linkage software) (Bradbury et al. 2007). Allelic data, total oil content, and two years of protein content (described in section 2.2.2) were used in this analysis. In the software, the general linear model (GLM) and a mixed linear model (MLM) model were tested. In addition to avoid false associations, Q-matrix (Q) was implemented with the GLM model and Kinship (K) and Q were implemented with the MLM model. Estimation of population membership (Q) obtained from the Structure program was used. A K matrix which shows the estimation of the relationship between population was calculated by TASSEL 2.1. Also, the same software was used to calculate P-values and LD values (r^2) between SSR loci. Loci with a p value less than 0.01 were considered to be associated to oil and protein content.

CHAPTER 3

RESULTS AND DISCUSSION

3.1. Phenotypic Evaluation of Population

Two biochemical characters, total oil and protein content were measured. For total oil, one season (second year-2017) of measurements was performed with 94 individuals. For protein content, two years of data were obtained with 85 accessions for the first year (2016) and 95 for the second year. Means, ranges and coefficient of variations (CV) were calculated.

3.1.1. Total Oil Content

Due to hazelnut's unsaturated fatty acids and healthy lipid composition, oil content is an important breeding trait. Since hazelnut oil can be used for cooking, the total oil content is also important (Miraliakbari and Shahidi 2008). The lipid composition of hazelnut oil and olive oil is similar and adulteration of virgin olive oil with hazelnut oil is difficult to detect. As a result hazelnut oil is the first choice of oil used to adulterate olive oil (Amaral et al. 2006).

In our work, total oil content ranged from 40.3% for Kargalak to 71.0% for FAI590 Table 3.1). Average oil content of hazelnut powders was 55.0 % for the 94 accessions. Means, ranges, and CV are listed at the bottom of Table 3.1 and a histogram of oil content is given in Figure 3.1. According to one-sample t-test, there were no statistically significant differences between the means of cultivars, landraces, and wild genotypes ($P \leq 0.05$). The highest group in terms of oil content was landraces with a value of 56.0%, followed by wild genotypes (55.0%) and cultivars (54.0%). Total oil contents of all accessions are listed in Table 3.1.

According to Taş & Gökmen (2015), 14 Turkish cultivars (Kargalak, Palaz, Incekara, Sivri, Yassı Badem, Fosa, Kalınkara, Yuvarlak Badem, Kus, Cakıldak, Kan,

Uzun Musa, Aci, Tombul) had higher oil content than was obtained in our work. Oil content of the cultivars ranged between 58.1% (Yuvarlak Badem) and 68.9% (Aci). They used the same organic solvent (n-hexane) and same extraction method (Soxhlet), extraction temperature was 55°C for ten hours. However, we used the rapid system of Soxhlet and our extraction temperature was 150°C for 1.5 hours. The increase in temperature may explain the decrease in oil yield because Tunç et al. (2014) indicated that increasing the temperature after a certain point had an adverse effect on the yield. Obviously, the amount of oil in hazelnuts collected from the same tree at different time periods can be shaped according to the climate conditions of the year in which it was collected. The oil content of 17 Turkish cultivars (Aci, Cavcava, Kara, Kus, İnce Kara, Kalin Kara, Kargalak, Mincane, Yuvarlak Badem, Yassı Badem, Palaz, Sivri, Tombul, Fosa, Uzun Musa, Cakildak, Kan) ranged between 56.1% (Cavcava) and 68.5% (Kalin Kara) according to Köksal (2002). Moreover, the lowest oil content was belong to Kargalak (40.3%) and highest was Aci (67%) in our 20 Turkish cultivars. Studies have also shown that the content of hazelnut can vary depending on where the tree is grown (Özdemir et al. 2001; Köksal et al. 2006; Matthäus and Özcan 2012).

3.1.2. Protein Content

Means, ranges, and CVs were calculated for the two years combined (Table 3.1.). Highest protein contents belonged to cultivars with a mean of 5.5% followed by wild genotypes (4.9%) and landraces (4.7%). Among cultivars, Kan and Palaz accessions had maximum (8.8%) and minimum (3.5%) protein contents, respectively. The first-year mean protein content was 5.3% and second-year protein content was 4.7%. Statistical analysis indicated that there was no significant difference in these values (P=0.001). Higher variation in protein content was seen in the second year (32.6%). A histogram of protein content for two years is given in Figure 3.1. Total protein contents of all accessions are listed in Table 3.1.

There have been limited number of studies about protein content in Turkish hazelnut accessions. According to Köksal (2002), 17 Turkish cultivars (Kargalak, Palaz, Incekara, Sivri, Yassı Badem, Fosa, Kalınkara, Yuvarlak Badem, Kus, Cakıldak, Kan, Uzun Musa, Aci, Tombul, Kara, Cavcava, Mincane) had protein content ranging between 11.7% and 20.8%. Protein content of Tombul, the most important hazelnut

accession of Turkey, was moderate with 15.3% (Iyanapathirana and Hshima 2003). On average our protein content of cultivars was 7.5%, this value was low compared to other studies.

Table 3.1. Total oil and protein content of accessions

	Oil Content	Protein Content
Accession Name	Mean ± SE (%)	Mean ± SE (%)
FAI105	61.3±0.8	3.9±0.9
FAI118	63.3±0.6	3.3±0.0
FAI132	43.0±0.0	6.4±0.0
FAI157	58.6±3.6	3.5±0.1
FAI166	53.6±1.7	6.7±0.1
FAI181	60.0±2.0	5.8±0.1
FAI183	52.6±2.4	3.9±0.2
FAI195	45.0±2.0	3.4±0.1
FAI200	60.3±4.0	4.7±0.4
FAI213	57.0±0.5	4.7±0.1
FAI219	59.6±0.3	3.2±1.4
FAI233	49.6±1.2	4.8±0.7
FAI244	64.0±6.0	5.5±0.0
FAI252	51.0±0.3	5.1±0.8
FAI253	65.0±3.4	3.4±0.6
FAI258	61.0±2.0	4.2±0.4
FAI270	62.3±0.8	5.2±1.1
FAI281	56.3±0.3	6.3±1.7
FAI283	53.0±4.1	5.6±0.5
FAI286	56.6±3.3	4.9±0.1
FAI335	56.3±1.3	4.4±0.0
FAI338	58.3±2.8	4.9±0.6
FAI351	56.3±2.8	3.9±0.1
FAI359	55.3±3.7	8.5±4.1
FAI377	50.3±2.4	6.4±1.5
FAI391	53.0±1.5	5.9±1.1
FAI401	46.0±2.5	3.6±0.2
FAI417	63.6±1.2	4.9±0.0
FAI421	54.0±0.5	4.9±0.8
FAI431	54.0±6.5	3.1±0.8
FAI433	51.0±8.5	4.5±0.0
FAI434	56.0±2.8	6.3±1.4
FAI436	51.6±0.8	5.4±0.6
FAI444	62.0±2.3	6.3±0.4
FAI453	57.3±0.8	4.2±0.3

(cont. on next page)

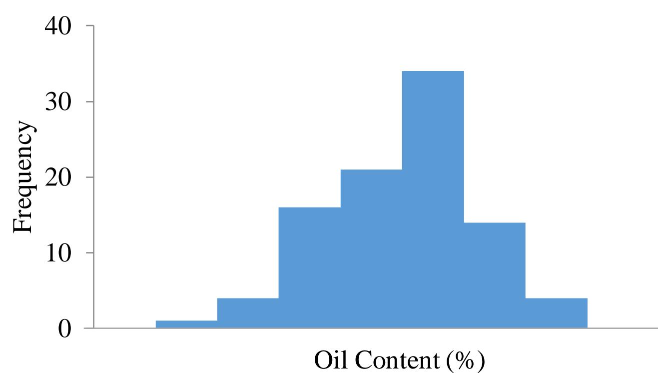
Table 3.1 (cont.)

FAI469	50.6±0.3	4.9±0.7
FAI471	62.6±3.6	4.0±1.2
FAI511	-	-
FAI530	67.0±3.0	3.9±0.0
FAI532	44.0±2.5	5.3±0.2
FAI533	60.0±1.0	3.7±0.3
FAI551	60.3±4.6	3.9±0.4
FAI558	50.0±0.5	6.4±0.4
FAI580	59.3±0.3	5.4±0.6
FAI586	59.6±0.8	4.4±0.6
FAI587	61.6±1.8	4.9±1.5
FAI588	61.6±2.0	3.4±0.4
FAI589	53.6±4.1	6.1±0.5
FAI590	71.0±3.0	3.6±0.1
FAI595	53.6±4.2	6.3±1.7
FAI598	60.3±1.8	3.7±1.2
FAI601	46.3±2.4	4.2±0.4
FAI604	58.6±2.4	4.8±0.4
FAI608	52.0±1.1	4.2±0.6
FAI609	52.3±1.8	4.3±0.4
FAI611	49.6±0.6	5.2±0.3
FAI616	68.3±3.3	5.9±1.8
FAI617	60.6±0.8	4.8±0.7
Acı	67.0±4.5	6.5±2.5
Allahverdi	62.6±2.3	6.9±0.2
Cavcava	54.6±3.6	6.6±0.6
Çakıldak	49.3±1.4	4.7±1.6
Foşa	53.0±6.1	5.9±1.4
Giresun Melezi	58.6±6.2	6.2±1.6
İnce Kara	55.3±0.8	5.2±1.5
Kalın Kara	60.3±2.0	5.0±1.4
Kan	52.0±1.5	8.8±0.3
Kara	49.0±2.0	4.5±0.3
Kargalak	40.3±2.0	6.4±0.2
Kuş	50.3±0.8	5.3±0.4
Mincane	60.6±0.3	5.2±1.1
Okay28	56.3±2.3	4.8±1.1
Palaz	58.3±4.0	3.5±1.6
Sivri	59.6±4.3	4.7±0.2
Tombul	54.6±1.4	5.4±0.1
Uzun Musa	58.6±4.3	4.1±0.1
Yassı Badem	57.3±1.4	5.7±0.5
Yuvarlak Badem	-	5.2±1.1
FAI108	57.3±1.4	6.2±1.1

(cont. on next page)

Table 3.1 (cont.)

FAI129	59.0±1.5	4.5±0.6
FAI165	53.0±1.0	6.8±0.1
FAI177	56.0±2.5	5.0±0.4
FAI190	56.3±4.9	5.5±0.0
FAI201	60.3±2.3	3.2±0.0
FAI221	50.6±2.1	2.9±0.0
FAI250	57.0±2.0	2.5±0.0
FAI266	49.0±2.5	4.2±0.0
FAI284	53.3±1.8	5.0±0.0
FAI287	41.0±1.5	4.9±0.0
FAI319	61.0±4.1	3.7±0.0
FAI405	52.3±3.1	6.1±0.9
FAI411	52.0±2.0	4.2±0.0
FAI419	55.3±1.6	7.8±0.0
FAI521	51.0±2.3	5.2±0.4
FAI522	58.0±0.5	5.6±1.1
FAI577	50.3±2.1	4.1±0.0
Mean ± SE (%)	55.0±0.6	5.4±0.1
Range (%)	40.0-70.0	25.0-89.9
CV (%)	9.1	23.5



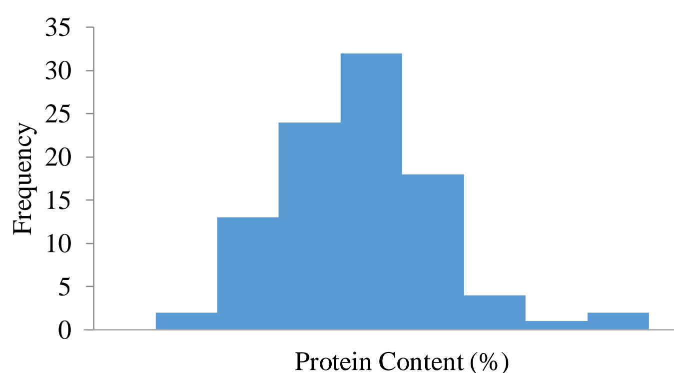


Figure 3.1. Frequency distributions of oil and protein content

3.2. Genotypic Evaluation of Population

The population was genotypically characterized with SSR markers and the data were used to examine genetic diversity and population structure.

3.2.1. SSR Analysis

Allelic polymorphism for the 96 individuals was obtained using SSR primers A total of 407 fragments were obtained with 30 primers. Except for primer B789, all primers were polymorphic for the population. On average, the number of polymorphic fragments was 13.5 for each primer. Primer B651 had the most polymorphic fragments (24) while B628 had the fewest (5). The maximum average gene diversity (using a scale of 0 to 0.50) was 0.46 for B788, the minimum was for B628 with a value of 0.14. The number of total and polymorphic fragments and the average GD value for each primer are listed in Table 3.2.

Table 3.2. The number of total and polymorphic fragments, average GD value for each primer.

Primer Name	Polymorphic fragments/ Total fragments	Average GD ± SE
A601	11/11	0.31 ± 0.05
A602	21/21	0.3 ± 0.02
A604	15/15	0.45 ± 0.01
A605	9/9	0.38 ± 0.03
A606	9/9	0.34 ± 0.04
A611	10/10	0.41 ± 0.02
A613	17/17	0.39 ± 0.02
A616	13/13	0.36 ± 0.03
A635	12/12	0.39 ± 0.02
A640	10/10	0.42 ± 0.02
B602	14/14	0.33 ± 0.03
B603	17/17	0.34 ± 0.02
B606	10/10	0.34 ± 0.03
B612	20/20	0.38 ± 0.01

(cont. on next page)

Table 3.2 (cont.)

B613	14/14	0.45 ± 0.01
B625	17/17	0.4 ± 0.01
B628	5/5	0.14 ± 0.06
B631	13/13	0.29 ± 0.03
B635	11/11	0.4 ± 0.03
B640	18/18	0.25 ± 0.03
B64a	9/9	0.34 ± 0.03
B64b	18/18	0.42 ± 0.01
B648	17/17	0.44 ± 0.01
B651	26/26	0.23 ± 0.03
B652	22/22	0.35 ± 0.02
B660	8/8	0.4 ± 0.04
B662	12/12	0.39 ± 0.03
B788	9/9	0.46 ± 0.009
B789	4/7	0.16 ± 0.06
CAC-B753	13/13	0.34 ± 0.02

The number of alleles (five to 24) obtained per SSR primer pair was similar to that reported in the literature. Gökirmak et al. (2005) investigated 272 cultivars with ten SSR markers, their allele number ranged from five to 20. The cultivars covered a broad geographical origin, such as Turkey, Italy, Spain and USA. Bassil et al. (2005) developed eight SSR loci for 20 accessions from worldwide locations. Amplification of eight loci resulted in 58 alleles with an average of seven alleles per marker. A total of 16 loci were amplified by Boccacci et al. (2006) in 78 cultivars. Their population was comprised of mainly European but also Turkish and American cultivars. Per locus, the number of alleles varied from six to 13 (mean 9.4). Gürcan et al. (2010a) analyzed 117 accessions with 12 loci, a total of 116 alleles were obtained from these loci. Accessions in this study mainly originated from Turkey but also cultivars from Georgia, Italy, Spain and Azerbaijcan were used. Gürcan et al. (2010b) implemented 86 newly developed loci enriched in (CA) and (GA) on 50 accessions and these loci were found to be highly polymorphic with from five to 21 alleles (mean 10.5). Our average polymorphic allele number was 13.5 which was higher than the previous studies indicating the usefulness of the selected markers on the tested germplasm.

3.2.2. Genetic Diversity Analysis

Darwin 6.0.14 software was utilized to detect genetic diversity within the population using the Dice coefficient and Unweighted Neighbor-Joining algorithm. According to the Dice coefficient dissimilarity matrix, the mean dissimilarity value was 0.52 (52%) within the 96 accessions. The lowest dissimilarity was between FAI319 and FAI287 (10%) and the highest between FAI419 and FAI444 (79%). Three clusters were obtained in the dendrogram produced by Unweighted Neighbor-Joining analysis (Figure 3.3). The cophenetic value generated by a Mantel test was 0.97 which demonstrated a high correlation between the dendrogram and dissimilarity matrix.

Cluster A had 38 accessions, Cluster B had 53 accessions and Cluster C had 5 accessions. All cultivars grouped in Cluster B which also contained 14 landraces and 19 wild genotypes. The remaining 34 wild genotypes and landraces fell into both Cluster A and Cluster C. Cluster C was the smallest group with 1 landrace and 4 wild genotypes. Cluster A had a great majority of wild genotypes, 30, with the remaining 8 accessions landraces. When we examined the dendrogram based on accession origin, no clustering was observed. Giresun was most commonly represented by the accessions and they fell into all three groups. Individuals with Erzurum and Rize origin fell into Cluster B. While individuals with Ordu origin were distributed in both Clusters A and B, individuals with Trabzon origin were present in all clusters.

According to PCoA analysis, all cultivars were clustered but not clearly separated from wild genotypes and landraces (Figure 3.2). Landraces and wild genotypes were scattered amongst the cultivars. The first two PCs explained 28.2 % of the total variation with 17.3 and 10.9% variance explained by Axes 1 and 2, respectively.

We observed partial separation in terms of cultivars, landraces, and wild genotypes as expected based on a previous study which examined all 402 accessions in the Turkish national collection (Öztürk et al., 2017). This separation can arise from the wild pollinated and dichogamous structure of hazelnuts. Over the years, selection, clonally propagation, and hybridization with known or unknown genotypes can explain the clustering of cultivars and their separation from wild genotypes.

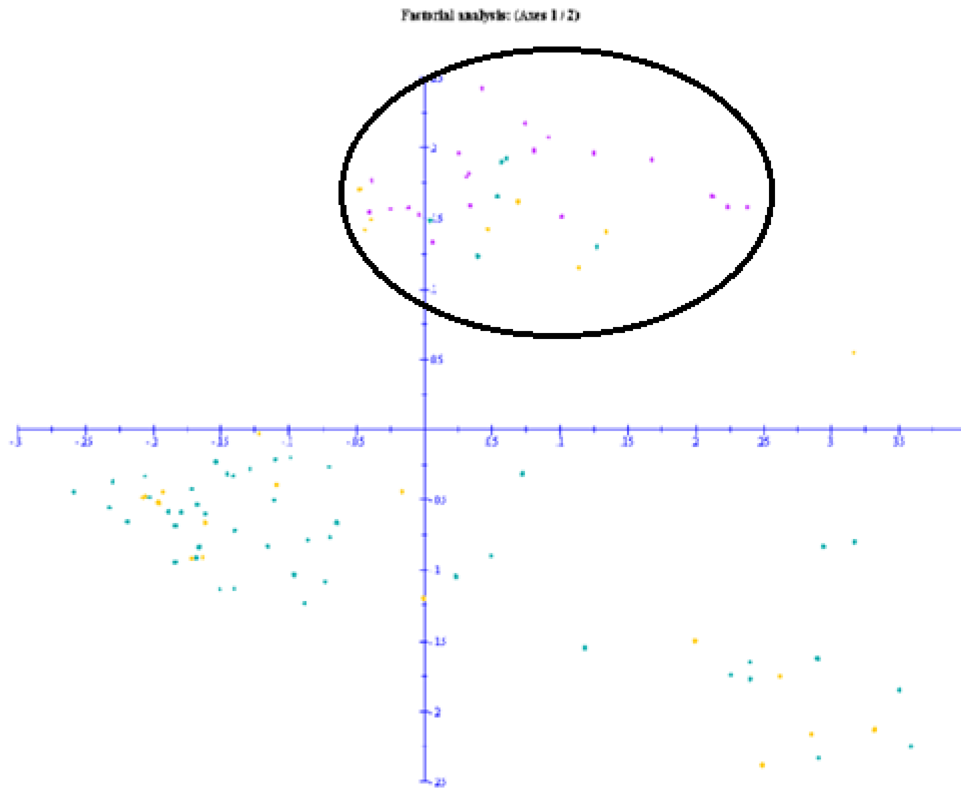


Figure 3.2. PCoA for 96 hazelnut accessions. Purple, orange, and cyan represent cultivars, landraces, and wild genotypes, respectively. Black circle is drawn to indicate the group containing cultivar.

3.2.3. Population Structure Analysis

Structure software was used to determine population structure. The population was analyzed for 2 to 20 subpopulations with 10 iterations for each group. Structure Harvester online program was utilized for evaluating these results. The second-order rate of change of likelihood Delta K (ΔK) was used to estimate the most correct number of subpopulations (Evanno et al. 2005). Standard deviation values for each K are shown in Figure 3.4 and Table 3.3 shows estimation of ΔK value based on Ln probability. According to the Harvester result, the highest ΔK represented the most likely number of subpopulations which occurred for K=3 (Figure 3.5).

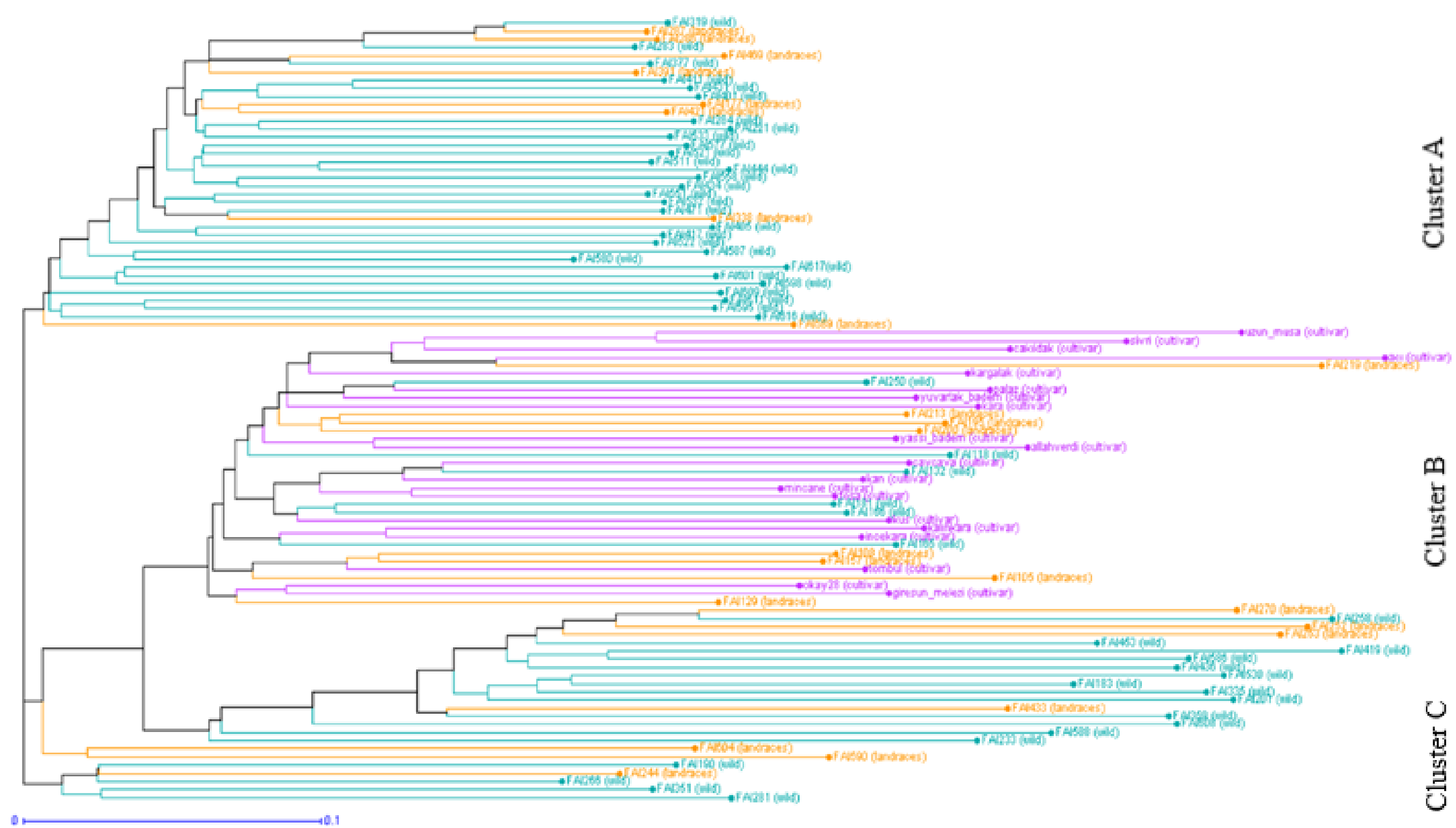


Figure 3.3. Dendrogram showing genetic diversity of 96 hazelnut accessions based on 407 SSR alleles. Purple, orange, and cyan represent cultivars, landraces, and wild genotypes, respectively.

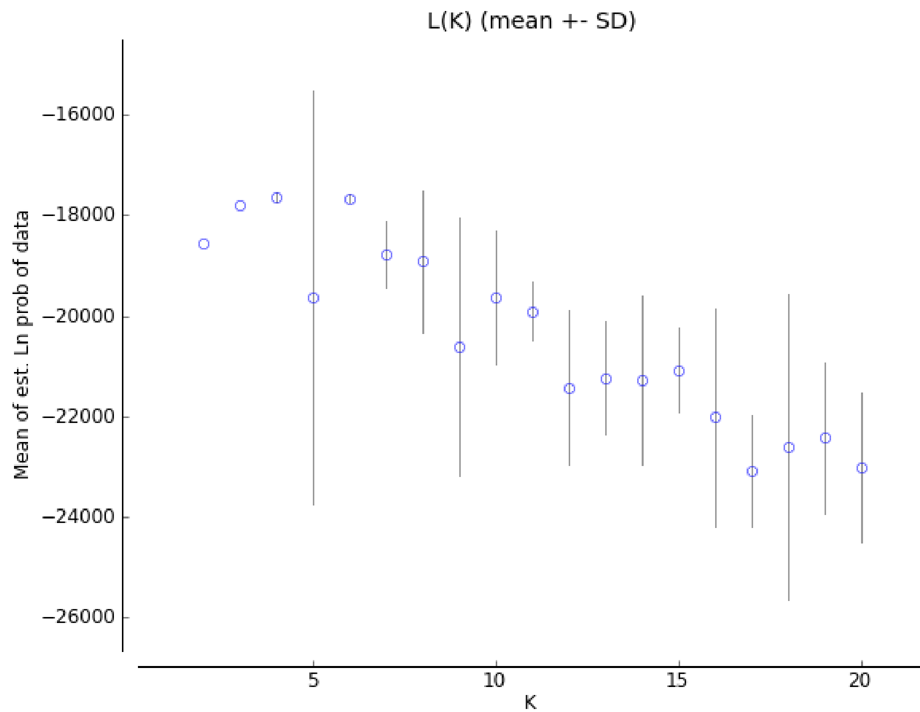


Figure 3.4. Standard deviation values of each number of K

Table 3.3. Estimation of ΔK value based on Ln probability

K	Mean LnP(K)	StDev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
2	-18562.480000	1.051771	—	—	—
3	-17784.580000	2.635569	777.900000	643.390000	244.118094
4	-17650.070000	95.550534	134.510000	2121.950000	22.207621
5	-19637.510000	4121.192506	-1987.440000	3968.960000	0.963061
6	-17655.990000	73.078245	1981.520000	3111.840000	42.582303
7	-18786.310000	661.180417	-1130.320000	1000.940000	1.513868
8	-18915.690000	1395.202987	-129.380000	1564.030000	1.121005
9	-20609.100000	2569.743445	-1693.410000	2662.120000	1.035948
10	-19640.390000	1333.196966	968.710000	1227.960000	0.921064
11	-19899.640000	580.441040	-259.250000	1267.960000	2.184477
12	-21426.850000	1532.203462	-1527.210000	1725.730000	1.126306
13	-21228.330000	1116.580351	198.520000	250.490000	0.224337

(cont. on next page)

Table 3.3 (cont.)

14	-21280.300000	1684.775391	-51.970000	252.280000	0.149741
15	-21079.990000	837.877620	200.310000	1128.030000	1.346294
16	-22007.710000	2163.502263	-927.720000	134.360000	0.062103
17	-23069.790000	1101.539061	-1062.080000	1523.580000	1.383138
18	-22608.290000	3033.275397	461.500000	267.740000	0.088268
19	-22414.530000	1503.039888	193.760000	791.870000	0.526846
20	-23012.640000	1492.555419	-598.110000	—	—

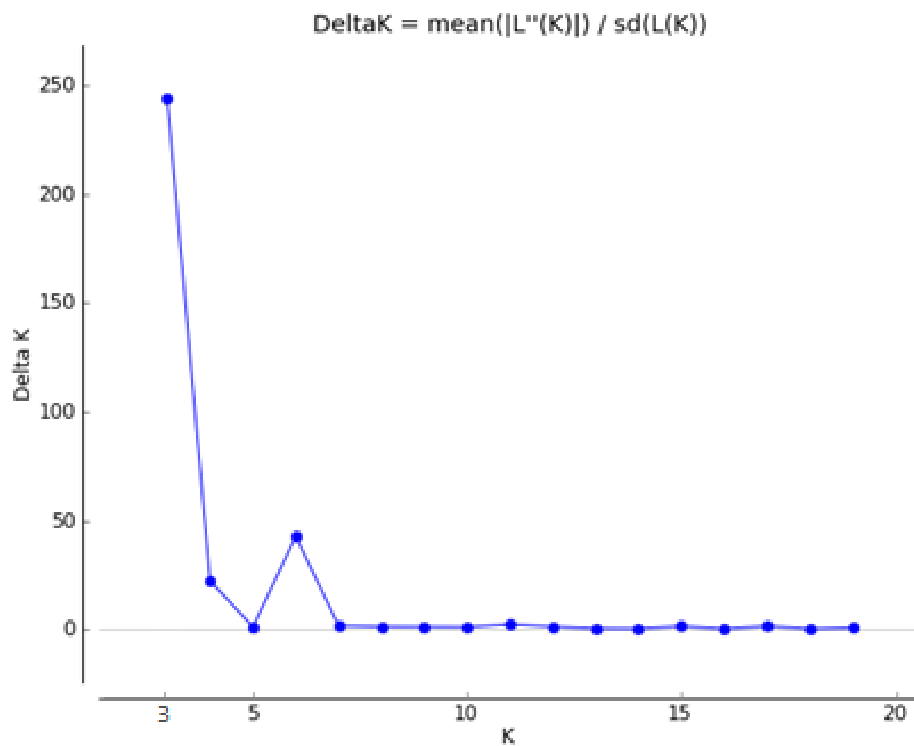


Figure 3.5. Delta K values for each number of K. The number of subpopulations within the hazelnut population was determined according to the highest ΔK value.

A bar plot graph representing the subpopulation structure of each accession was drawn showing membership in each subpopulation in a different color (Figure 3.6). In this graph, the height of each colored bar gives the proportion of the individual's genome that can be assigned to that cluster. Individuals are arranged along the x-axis. A membership coefficient threshold value (y-axis) of 0.70 was taken into consideration to assign

accessions to their clusters. Individuals which did not exceed the threshold for any of the three subpopulations were considered to be admixed.

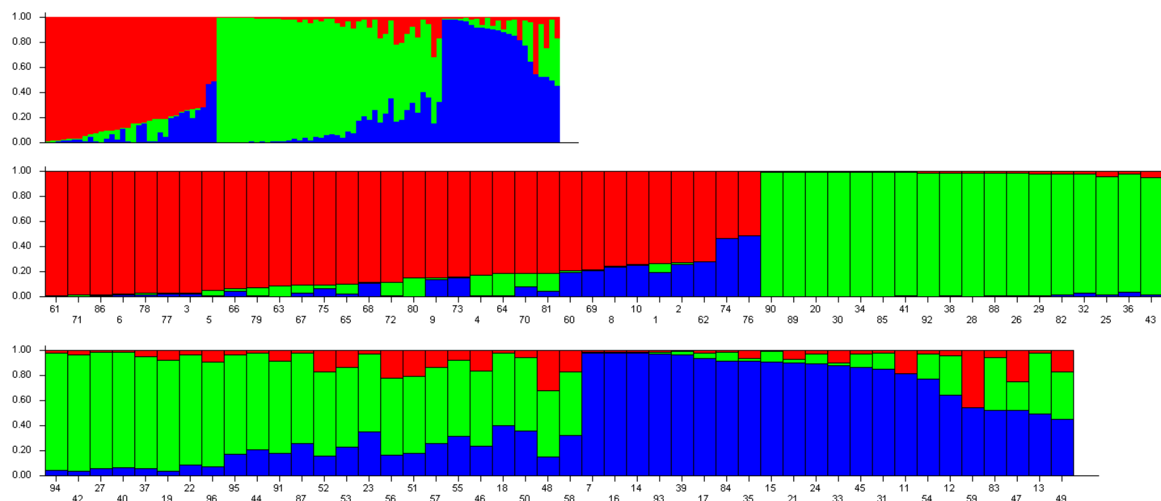


Figure 3.6. Population structure graph for K=3.The x-axis shows each accession, the y-axis represents the membership coefficient. Colors represent three different clusters; Cluster 1 is shown in red, Cluster 2 is shown in green, and Cluster 3 is shown in blue.

Nearly one third of the population was assigned to subpopulation 1 with 30 accessions. Subpopulation 2 also contained 30 accessions and subpopulation 3 had 16 accessions. The remaining 20 accessions did not belong to any subpopulation and were admixed. The average distances within each subpopulation were 0.27, 0.36, and 0.17, respectively. The cluster assignments based on diversity and structure analyses are given in Table 3.4, inferred subpopulation values of each accession are highlighted in yellow. Comparing the two methodologies in terms of subpopulation assignment, accessions in subpopulation 1 fell into Cluster B in the dendrogram. All accessions in subpopulation 2 belonged to Cluster A, except accession FAI266 which was in Cluster C in the dendrogram. Subpopulation 3 corresponded to Cluster B in the dendrogram. The admixed individuals were dispersed in all dendrogram clusters. Among the admixed group, nine accessions belonged to Cluster A, seven accessions, which included cultivars Acı, Sivri and Uzun Musa, belonged to Cluster B and four accessions belonged to Cluster C in the dendrogram.

Consequently, the evaluation of hazelnut population structure and diversity gave very similar results.

The fairly simple population structure of the hazelnut accessions was expected based on the results of Öztürk et al. (2017) who examined the entire Turkish national collection including 20 cultivars, 143 wild genotypes and 239 landraces and found two subpopulations One significant difference was that the cultivars Giresun Melezi, Kan, Ince Kara, Fosa, and Okay28 were considered admixed by Öztürk et al. (2017) but fell into subpopulation 1 in the current work. Similar to our results, cultivars gathered in one cluster which also contained wild genotypes and landraces. In other work which examined 47 cultivars from nine countries, the hazelnuts clustered into four subpopulations with Turkish cultivars falling into three of these clusters, Sivri was in the admixed group like our result (Öztürk et al. 2018). (Boccacci et al. 2013) analyzed 57 cultivars from Europe and Turkey, 77 landraces and 19 wild genotypes. According to their results, wild genotypes and cultivars clearly separated from each other. Although in our results, cultivars were grouped among themselves, they were not separated from wild genotypes and landraces clearly.

Table 3.4. Hazelnut accessions and their assigned subpopulations and dendrogram clusters

No	Accession Name	Subpopulation Identity Values			Assigned Subpopulation /Dendrogram Cluster
1	FAI105	0.731	0.074	0.195	1/B
2	FAI118	0.728	0.008	0.264	1/B
3	FAI132	0.967	0.011	0.022	1/B
4	FAI157	0.823	0.164	0.012	1/B
5	FAI166	0.948	0.039	0.013	1/B
6	FAI181	0.973	0.006	0.020	1/B
7	FAI183	0.008	0.008	0.984	3/B
8	FAI195	0.753	0.006	0.241	1/B
9	FAI200	0.843	0.016	0.140	1/B
10	FAI213	0.737	0.008	0.254	1/B
11	FAI219	0.180	0.004	0.816	3/B
12	FAI233	0.039	0.318	0.644	Admixed/B
13	FAI244	0.015	0.491	0.494	Admixed/C

(cont. on next page)

Table 3.4. (cont.)

14	FAI252	0.013	0.007	0.980	3/B
15	FAI253	0.006	0.082	0.912	3/B
16	FAI258	0.013	0.004	0.982	3/B
17	FAI270	0.018	0.045	0.937	3/B
18	FAI281	0.016	0.583	0.400	Admixed/C
19	FAI283	0.075	0.887	0.038	2/A
20	FAI286	0.006	0.990	0.004	2/A
21	FAI335	0.071	0.028	0.901	3/B
22	FAI338	0.035	0.874	0.091	2/A
23	FAI351	0.026	0.619	0.355	Admixed/C
24	FAI359	0.026	0.076	0.899	3/B
25	FAI377	0.037	0.947	0.016	2/A
26	FAI391	0.009	0.978	0.014	2/A
27	FAI401	0.011	0.929	0.060	2/A
28	FAI417	0.012	0.982	0.006	2/A
29	FAI421	0.016	0.975	0.009	2/A
30	FAI431	0.006	0.989	0.005	2/A
31	FAI433	0.019	0.131	0.850	3/B
32	FAI434	0.015	0.953	0.031	2/A
33	FAI436	0.100	0.021	0.879	3/B
34	FAI444	0.007	0.989	0.004	2/A
35	FAI453	0.063	0.018	0.919	3/B
36	FAI469	0.016	0.945	0.039	2/A
37	FAI471	0.043	0.899	0.058	2/A
38	FAI511	0.008	0.983	0.009	2/A
39	FAI530	0.006	0.029	0.965	3/B
40	FAI532	0.010	0.925	0.064	2/A
41	FAI533	0.006	0.985	0.009	2/A
42	FAI551	0.033	0.930	0.036	2/A
43	FAI558	0.049	0.934	0.017	2/A
44	FAI580	0.016	0.775	0.209	2/A
45	FAI586	0.026	0.107	0.867	3/B
46	FAI587	0.164	0.596	0.240	Admixed/A
47	FAI588	0.244	0.232	0.524	Admixed/B

(cont. on next page)

Table 3.4. (cont.)

48	FAI589	0.317	0.532	0.151	Admixed/A
49	FAI590	0.171	0.373	0.456	Admixed/B
50	FAI595	0.054	0.582	0.364	Admixed/A
51	FAI598	0.206	0.611	0.182	Admixed/A
52	FAI601	0.166	0.670	0.164	Admixed/A
53	FAI604	0.129	0.642	0.229	Admixed/B
54	FAI608	0.023	0.205	0.773	3/B
55	FAI609	0.073	0.607	0.320	Admixed/A
56	FAI611	0.219	0.615	0.167	Admixed/A
57	FAI616	0.131	0.610	0.260	Admixed/A
58	FAI617	0.168	0.504	0.328	Admixed/A
59	Aci	0.451	0.003	0.546	Admixed/B
60	Allahverdi	0.792	0.012	0.196	1/B
61	Cavcava	0.991	0.003	0.006	1/B
62	Cakıldak	0.716	0.003	0.281	1/B
63	Fosa	0.913	0.081	0.006	1/B
64	Giresun Melezi	0.812	0.176	0.011	1/B
65	Ince Kara	0.897	0.078	0.026	1/B
66	Kalin Kara	0.929	0.026	0.044	1/B
67	Kan	0.905	0.063	0.032	1/B
68	Kara	0.880	0.011	0.109	1/B
69	Kargalak	0.781	0.011	0.208	1/B
70	Kus	0.811	0.109	0.080	1/B
71	Mincane	0.982	0.011	0.007	1/B
72	Okay28	0.880	0.111	0.010	1/B
73	Palaz	0.841	0.005	0.154	1/B
74	Sivri	0.531	0.003	0.467	Admixed/B
75	Tombul	0.905	0.028	0.067	1/B
76	Uzun Musa	0.511	0.003	0.486	Admixed/B
77	Yuvarlak Badem	0.970	0.007	0.023	1/B
78	Yassi Badem	0.971	0.011	0.018	1/B
79	FAI108	0.924	0.067	0.009	1/B
80	FAI129	0.844	0.151	0.005	1/B
81	FAI165	0.811	0.141	0.048	1/B

(cont. on next page)

Table 3.4 (cont.)

82	FAI177	0.021	0.961	0.018	2/A
83	FAI190	0.052	0.421	0.526	Admixed/C
84	FAI201	0.014	0.066	0.921	3/B
85	FAI221	0.007	0.989	0.004	2/A
86	FAI250	0.981	0.005	0.014	1/B
87	FAI266	0.016	0.725	0.259	2/C
88	FAI284	0.011	0.980	0.009	2/A
89	FAI287	0.005	0.992	0.003	2/A
90	FAI319	0.005	0.993	0.003	2/A
91	FAI405	0.079	0.741	0.180	2/A
92	FAI411	0.011	0.984	0.005	2/A
93	FAI419	0.010	0.012	0.979	3/B
94	FAI521	0.020	0.934	0.046	2/A
95	FAI522	0.029	0.797	0.174	2/A
96	FAI577	0.090	0.837	0.073	2/A

3.3. Association Analysis

Total oil and protein contents were assessed for their association with the 407 SSR alleles. In the GLM (Q) model, the SSR loci associated with oil and protein content of hazelnuts are listed in Table 3.5. According to the results, five loci were significantly associated with oil content and three loci were associated with protein content ($P \leq 0.01$) and LD values ranged between 0.070 and 0.145.

As a result, this locus, B628-307, was considered to be the most significant locus for oil content. For protein content, the most significant result was for A613-153 with ($P=0.003$) and an LD value of 0.088 indicating that the QTL accounted for 8.8% of the phenotypic variance.

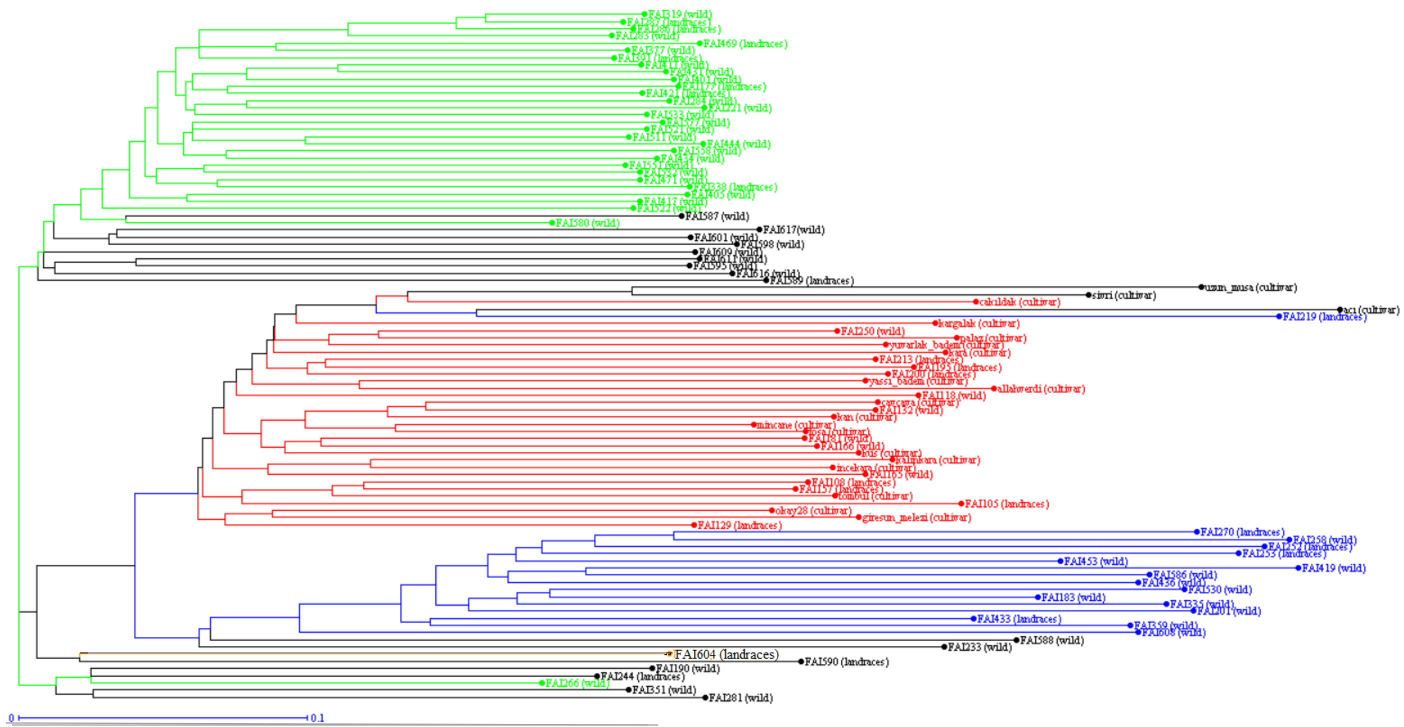


Figure 3.7. Correspondence of genetic diversity and population structure analysis. This neighbor joining dendrogram has been re-colored according to population structure analysis. Red color represents subpopulation 1, green color represents subpopulation.

Table 3.5. SSR loci associated with oil and protein contents of hazelnuts in the GLM (Q) model

Trait	Locus	LD value (r^2)	P-value
Oil	B628-307	0.145	0.0002
	B641b-444	0.083	0.0041
	A604-339	0.077	0.0056
	B652-292	0.071	0.0089
	B631-306	0.070	0.0094
Protein	A613-153	0.088	0.0030
	A602-342	0.084	0.0038
	B631-209	0.071	0.0082

The number of QTL mapping studies in hazelnut is limited and all of them have been recently published. In 2016, linkage mapping was conducted for vigor, sucker habit, and time of bud burst traits with three years of observations. A total of 163 F_1 individuals (Tonda Gentile delle Langhe × Hall's Giant) were scanned with 152 EST-SSR and after the construction of the linkage map, 151 markers segregated in 11 linkage groups. Fifteen QTL were detected in the linkage groups. For leaf budburst, one major QTL was located on linkage group 2 accounting for 50% of phenotypic variance (Beltramo et al. 2016). To saturate the linkage map of Beltramo et al. (2016), Marinoni et al. (2018) used 9.999 polymorphic SNP loci, 21 SSR loci and 50 additional individuals. Time of leaf budburst was associated with 29 QTL including the same major QTL which was detected by Beltramo et al. (2016) on linkage group 2. This research was the first time that fine saturated genetic mapping was performed in hazelnut.

Öztürk et al. (2017) performed association mapping in 54 wild and 48 cultivated Slovenian hazelnut accessions using 11 AFLP and 49 SSR markers. Their traits of interest examined nuts and kernels (length, shape uniformity, nut caliber, kernel weight, thickness). Among 17 traits, nine were associated with 49 SSR loci. They concluded that the remaining eight traits were not polymorphic enough in the population to detect association.

Most recently, 44 agro-morphological traits were associations with 30 SSR loci examined in 390 Turkish hazelnut accessions including cultivars, landraces and wild genotypes. A total of 406 polymorphic fragments were obtained and 145 alleles were associated with 41 traits. The same population was examined for nut, kernel and yield

traits. Seventy eight loci were significantly associated with these traits. Almost half of the 78 loci were associated with nut and kernel appearance traits. The remaining loci were associated with quality, shell thickness and yield related traits (Frary et al. 2019a, b).

To date, there have been no reports about association mapping for nutritional traits of hazelnuts. Therefore, it is not possible to compare our results with previous studies.

CHAPTER 4

CONCLUSION

European hazelnut has an important place among tree nuts species for both their nutritional composition and economic value. This was the first study aiming to find QTL for such traits by associating genetic and oil-protein content diversity of hazelnuts. To achieve this, we determined the variation in oil and protein content of 96 hazelnut accessions. Diversity of oil and protein content of hazelnut was seen with CVs of 23.5% and 9.1%, respectively. The presence of this diversity is important for breeding of these traits and also for the selection of parental lines for such breeding. This diversity also made it possible to identify QTLs for these traits with association mapping. We detected eight loci; five of them associated with oil content, three associated with protein content. The most significant results were obtained from markers B628-307 and A613-153 for oil and protein content, respectively. The identification of such QTL may accelerate the marker assisted breeding of hazelnut. In the future the association map can be saturated using more markers as genome wide coverage is needed to detect more associations. In addition, other nutritional characters of hazelnuts can be characterized and their QTLs mapped in the same way.

REFERENCES

Abdurakhmonov, Ibrokhim Y., and Abdusattor Abdukarimov. (2008). "Application of Association Mapping to Understanding the Genetic Diversity of Plant Germplasm Resources." *International Journal of Plant Genomics* 2008. https://doi.org/10.1155/2008/574927.

Alasalvar, Cesarettin, Joana S. Amaral, Gülçin Satir, and Fereidoon Shahidi. (2009). "Lipid Characteristics and Essential Minerals of Native Turkish Hazelnut Varieties (*Corylus avellana* L.)." *Food Chemistry* 113 (4): 919–25. https://doi.org/10.1016/j.foodchem.2008.08.019.

Alasalvar, Cesarettin, Fereidoon Shahidi, Joana S. Amaral, and Beatriz P.P. Oliveira. (2010). "Compositional Characteristics and Health Effects of Hazelnut (*Corylus avellana* L.)." In *Tree Nuts: Composition, Phytochemicals and Heath Effects*, edited by Cesarettin Alasalvar and Freeidoon Shadidi, 185–214. Nutraceutical Science and Technology. https://doi.org/https://doi.org/10.1365/s10337-010-1619-5.

Alasalvar, Cesarettin, Fereidoon Shahidi, Toshiaki Ohshima, Udaya Wanasundara, Hasan C. Yurttas, Chandrika M. Liyanapathirana, and Fabiana B. Rodrigues. (2003). "Turkish Tombul Hazelnut (*Corylus avellana* L.). 2. Lipid Characteristics and Oxidative Stability." *Journal of Agricultural and Food Chemistry* 51 (13): 3797–3805. https://doi.org/10.1021/jf021239x.

Alaviani, Shaghayegh, Farzaneh Mahmoudyar, Seyedeh Faranak Miraftabi, Mohammad Hossein Salehisormghi, and Mahnaz Qomi. (2012). "Determination of Iranian Almond , Peanut and Hazelnut Mineral Contents" 2 (7): 50–54.

Amaral, Joana S., Susana Casal, Ivana Citová, Alberto Santos, Rosa M. Seabra, and Beatriz P.P. Oliveira. (2006). "Characterization of Several Hazelnut (*Corylus avellana* L.) Cultivars Based in Chemical, Fatty Acid and Sterol Composition." *European Food Research and Technology* 222 (3–4): 274–80. https://doi.org/10.1007/s00217-005-0068-0.

Angelis, Elisabetta De, Simona Lucia Bavaro, Linda Monaci, and Rosa Pilolli. (2018). "Effects of the Varietal Diversity and the Thermal Treatment on the Protein Profile of Peanuts and Hazelnuts" 2018.

Bacchetta, L., M. Aramini, A. Zini, V. Di Giammatteo, D. Spera, P. Drogoudi, M. Rovira, A. P. Silva, A. Solar, and R. Botta. (2013). "Fatty Acids and Alpha-

Tocopherol Composition in Hazelnut (*Corylus avellana* L.): A Chemometric Approach to Emphasize the Quality of European Germplasm.” *Euphytica* 191 (1): 57–73. https://doi.org/10.1007/s10681-013-0861-y.

Balta, M. F., T. Yarilgaç, M. A. Aşkin, M. Kuçuk, F. Balta, and K. Özrenk. (2006). “Determination of Fatty Acid Compositions, Oil Contents and Some Quality Traits of Hazelnut Genetic Resources Grown in Eastern Anatolia of Turkey.” *Journal of Food Composition and Analysis* 19 (6–7): 681–86. https://doi.org/10.1016/j.jfca.2005.10.007.

Barabaschi, Delfina, Alessandro Tondelli, Francesca Desiderio, Andrea Volante, Patrizia Vaccino, Giampiero Valè, and Luigi Cattivelli. (2015). “Next Generation Breeding.” *Plant Science* 242: 3–13. https://doi.org/10.1016/j.plantsci.2015.07.010.

Bassil, N. V., S. A. Mehlenbacher, and R. Botta. (2005). “Additional Microsatellite Markers of the European Hazelnut.” *Acta Horticulturae* 686: 105–10. https://doi.org/10.17660/ActaHortic.2005.686.13.

Beltramo, Chiara, Nadia Valentini, Ezio Portis, Daniela Torello Marinoni, Paolo Boccacci, Maria Angelica Sandoval Prando, and Roberto Botta. (2016). “Genetic Mapping and QTL Analysis in European Hazelnut (*Corylus avellana* L.).” *Molecular Breeding* 36 (3): 1–17. https://doi.org/10.1007/s11032-016-0450-6.

Bhattarai, Gehendra. (2015). “Microsatellite Marker Development, Characterization and Mapping in European Hazelnut (*Corylus avellana* L.), and Investigation of Novel Sources of Eastern Filbert Blight Resistance in Corylus.” https://ir.library.oregonstate.edu/concern/graduate_thesis_or_dissertations/q524jt50r

Black, Alison E., and A. A. Paul. (1999). “McCance and Widdowson’s Tables of Food Composition: Origins and Clarification of the Fifth Edition.” *Journal of Human Nutrition and Dietetics* 12 (1): 1–5. https://doi.org/10.1046/j.1365-277x.1999.00140.x.

Boccacci, P., A. Akkak, and R. Botta. (2006). “DNA Typing and Genetic Relations among European Hazelnut (*Corylus avellana* L.) Cultivars Using Microsatellite Markers.” *Genome* 49 (6): 598–611. https://doi.org/10.1139/G06-017.

Boccacci, P., M. Aramini, N. Valentini, L. Bacchetta, M. Rovira, P. Drogoudi, A. P. Silva, et al. (2013). “Molecular and Morphological Diversity of On-Farm Hazelnut (*Corylus avellana* L.) Landraces from Southern Europe and Their Role in the Origin and Diffusion of Cultivated Germplasm.” *Tree Genetics and Genomes* 9 (6): 1465–80. https://doi.org/10.1007/s11295-013-0651-7.

Bradbury, Peter J., Zhiwu Zhang, Dallas E. Kroon, Terry M. Casstevens, Yogesh Ramdoss, and Edward S. Buckler. (2007). "TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples." *Bioinformatics* 23 (19): 2633–35. https://doi.org/10.1093/bioinformatics/btm308.

Bradford, Marion M. (1976). "A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding." *Analytical Biochemistry* 72: 248–54. https://doi.org/10.1016/j.cj.2017.04.003.

Bush, William S., and Jason H. Moore. (2012). "Chapter 11: Genome-Wide Association Studies." *PLoS Computational Biology* 8 (12). https://doi.org/10.1371/journal.pcbi.1002822.

Ciemniewska-Zytkiewicz, Hanna, Vito Verardo, Federica Pasini, Joanna Bryś, Piotr Koczoń, and Maria Fiorenza Caboni. (2015). "Determination of Lipid and Phenolic Fraction in Two Hazelnut (*Corylus avellana* L.) Cultivars Grown in Poland." *Food Chemistry* 168: 615–22. https://doi.org/10.1016/j.foodchem.2014.07.107.

Cosmulescu, Sina, Mihai Botu, and Ion Trandafir. (2013). "The Mineral Source for Human Nutrition of Nuts in Different Hazelnut (*Corylus avellana* L.) Cultivars." *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 41 (1): 250–56. https://doi.org/10.15835/nbha4118307.

Dreisigacke, Susanne, Sivakumar Sukumaran, Carlos Guzmán, Xinyao He, Caixa Lan, David Bonnett, and Jose Crossa. (2016). "Molecular Breeding for Sustainable Crop Improvement." *Molecular Breeding for Sustainable Crop Improvement* 11: 421–74. https://doi.org/10.1007/978-3-319-27090-6.

Evanno, G., S. Regnaut, and J. Goudet. (2005). "Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study." *Molecular Ecology* 14 (8): 2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

FGore, Michael, Edward S. Buckler, and Jianming Yu. (2008). "Status and Prospects of Association Mapping in Plants." *The Plant Genome Journal* 1 (1): 5. https://doi.org/10.3835/plantgenome2008.02.0089.

Fleury, Delphine, Ute Baumann, and Peter Langridge. (2012). *Plant Genome Sequencing: Models for Developing Synteny Maps and Association Mapping*. *Plant Biotechnology and Agriculture*. First Edit. Elsevier Inc. https://doi.org/10.1016/B978-0-12-381466-1.00006-7.

Francia, E., G. Tacconi, C. Crosatti, D. Barabaschi, D. Bulgarelli, E. Dall'Aglio, and G. Valè. (2005). "Marker Assisted Selection in Crop Plants." *Plant Cell, Tissue and Organ Culture* 82 (3): 317–42. https://doi.org/10.1007/s11240-005-2387-z.

Frary, Amy, Suleyman Can Öztürk, Hüseyin Irfan Balık, Selda Kayalak Balık, Gökhan Kızılcı, Sami Doğanlar, and Anne Frary. (2019). "Association Mapping of Agro-Morphological Traits in European Hazelnut (*Corylus avellana*)." *Euphytica* 215 (2). https://doi.org/10.1007/s10681-019-2352-2.

Frary, Amy, Süleyman Can Öztürk, Hüseyin Irfan Balık, Selda Kayalak Balık, Gökhan Kızılcı, Sami Doğanlar, and Anne Frary. (2019). "Analysis of European Hazelnut (*Corylus avellana*) Reveals Loci for Cultivar Improvement and the Effects of Domestication and Selection on Nut and Kernel Traits." *Molecular Genetics and Genomics* 294 (2): 519–27. https://doi.org/10.1007/s00438-018-1527-1.

Fu, Yong Bi. (2015). "Understanding Crop Genetic Diversity under Modern Plant Breeding." *Theoretical and Applied Genetics* 128 (11): 2131–42. https://doi.org/10.1007/s00122-015-2585-y.

Fulton, Theresa M., Julapark Chunwongse, and Steven D. Tanksley. (1995). "Microprep Protocol for Extraction of DNA from Tomato and Other Herbaceous Plants." *Plant Molecular Biology Reporter* 13 (3): 207–9. https://doi.org/10.1007/BF02670897.

Gökirmak, T., S. A. Mehlenbacher, and N. V. Bassil. (2005). "Investigation of Genetic Diversity among European Hazelnut (*Corylus avellana*) Cultivars Using SSR Markers." *Acta Horticulturae* 686 (April 2016): 141–48. https://doi.org/10.17660/ActaHortic.2005.686.18.

Gupta, P. K., H. S. Balyan, P. C. Sharma, and B. Ramesh. (1996). "Microsatellites in Plants: A New Class of Molecular Markers." *Current Science* 70 (1): 45–54.

Gürcan, Kahraman, Shawn A. Mehlenbacher, Roberto Botta, and Paolo Boccacci. (2010). "Development, Characterization, Segregation, and Mapping of Microsatellite Markers for European Hazelnut (*Corylus avellana* L.) from Enriched Genomic Libraries and Usefulness in Genetic Diversity Studies." *Tree Genetics and Genomes* 6 (4): 513–31. https://doi.org/10.1007/s11295-010-0269-y.

Hall, David, Carolina Tegström, and Pär K. Ingvarsson. (2010). "Using Association Mapping to Dissect the Genetic Basis of Complex Traits in Plants." *Briefings in Functional Genomics and Proteomics* 9 (2): 157–65. https://doi.org/10.1093/bfgp/elp048.

Honig, Josh A., Megan F. Muehlbauer, John M. Capik, Christine Kubik, Jennifer N. Vaiciunas, Shawn A. Mehlenbacher, and Thomas J. Molnar. (2019). "Identification and Mapping of Eastern Filbert Blight Resistance Quantitative Trait Loci in European Hazelnut Using Double Digestion Restriction Site Associated DNA Sequencing." *Journal of the American Society for Horticultural Science* 144 (5): 295–304. https://doi.org/10.21273/JASHS04694-19.

İslam, Ali. (2019). "Fındık Islahında Gelişmeler" 174: 167–74.

İslam, Ali. (2018). "Hazelnut Cultivation in Turkey." *Akademik Ziraat Dergisi* 7 (2): 251–58. https://doi.org/10.29278/azd.476665.

Iyanapathirana, C Handrika M L, and T Oshiaki O Hshima. (2003). "Turkish Tombul Hazelnut (*Corylus avellana* L.). 1 . Compositional Characteristics," 3790–96. https://doi.org/10.1021/jf0212385.

Jiang, Guo-Liang. (2013). "Molecular Markers and Marker-Assisted Breeding in Plants." *Plant Breeding from Laboratories to Fields*, no. September. https://doi.org/10.5772/52583.

Johnson, Dennis V. (2019). *Advances in Plant Breeding Strategies: Nut and Beverage Crops*. *Advances in Plant Breeding Strategies: Nut and Beverage Crops*. Vol. 4. https://doi.org/10.1007/978-3-030-23112-5.

Kalia, Rajwant K., Manoj K. Rai, Sanjay Kalia, Rohtas Singh, and A. K. Dhawan. (2011). "Microsatellite Markers: An Overview of the Recent Progress in Plants." *Euphytica* 177 (3): 309–34. https://doi.org/10.1007/s10681-010-0286-9.

Köksal, A. Ilhami, Nevzat Artik, Atilla Şimşek, and Nurdan Güneş. (2006). "Nutrient Composition of Hazelnut (*Corylus avellana* L.) Varieties Cultivated in Turkey." *Food Chemistry* 99 (3): 509–15. https://doi.org/10.1016/j.foodchem.2005.08.013.

M. Saba Rahim, Himanshu Sharma, Afsana Parveen, and Joy K. Roy. (2018). "Trait Mapping Approaches Through Association Analysis in Plants." *Springer,Cham*. https://doi.org/https://doi.org/10.1007/10_2017_49.

Mackay, Trudy F C. (2001). "Of Q Uantitative T Raits." *North* 35 (1): 303–39. https://doi.org/10.1146/annurev.genet.35.102401.090633.

Marinoni, Daniela Torello, Nadia Valentini, Ezio Portis, Alberto Acquadro, Chiara Beltramo, Shawn A. Mehlenbacher, Todd C. Mockler, Erik R. Rowley, and Roberto Botta. (2018). "High Density SNP Mapping and QTL Analysis for Time of Leaf Budburst in Corylus Avellana L." *PLoS ONE* 13 (4): 1–24. https://doi.org/10.1371/journal.pone.0195408.

Mason, Annaliese S. (2015). "SSR Genotyping." In *Plant Genotyping: Methods and Protocols*, 1245:77–89. https://doi.org/10.1007/978-1-4939-1966-6.

Matthäus, Bertrand, and Mehmet Musa Özcan. (2012). "The Comparison of Properties of the Oil and Kernels of Various Hazelnuts from Germany and Turkey." *European Journal of Lipid Science and Technology* 114 (7): 801–6. https://doi.org/10.1002/ejlt.201100299.

Miraliakbari, Homan, and Fereidoon Shahidi. (2008). "Lipid Class Compositions, Tocopherols and Sterols of Tree Nut Oils Extracted with Different Solvents." *Journal of Food Lipids* 15 (1): 81–96. https://doi.org/10.1111/j.1745-4522.2007.00104.x.

Mohan Jain, S., and D. S. Brar. (2009). "Molecular Techniques in Crop Improvement: 2nd Edition." *Molecular Techniques in Crop Improvement: 2nd Edition*, 1–772. https://doi.org/10.1007/978-90-481-2967-6.

Nadeem, Muhammad Azhar, Muhammad Amjad Nawaz, Muhammad Qasim Shahid, Yıldız Doğan, Gonul Comertpay, Mehtap Yıldız, Rüştü Hatipoğlu, et al. (2018). "DNA Molecular Markers in Plant Breeding: Current Status and Recent Advancements in Genomic Selection and Genome Editing." *Biotechnology and Biotechnological Equipment* 32 (2): 261–85. https://doi.org/10.1080/13102818.2017.1400401.

Özdemir, Murat, Filiz Açkurt, Muammer Kaplan, Mustaf Yildiz, Mahmut Löker, Türcan Gürcan, Gül Biringen, Ahmet Okay, and Ferda G. Seyhan. (2001). "Evaluation of New Turkish Hybrid Hazelnut (*Corylus avellana* L.) Varieties: Fatty Acid Composition, α-Tocopherol Content, Mineral Composition and Stability." *Food Chemistry* 73 (4): 411–15. https://doi.org/10.1016/S0308-8146(00)00315-0.

Öztürk, Süleyman Can, Hüseyin İrfan Balık, Selda Kayalak Balık, Gökhan Kızılcı, Ömür Duyar, Sami Doğanlar, and Anne Frary. (2017). "Molecular Genetic Diversity of the Turkish National Hazelnut Collection and Selection of a Core Set." *Tree Genetics and Genomes* 13 (5). https://doi.org/10.1007/s11295-017-1195-z.

Öztürk, Süleyman Can, Mehmet Göktay, Jens Allmer, Sami Doğanlar, and Anne Frary. (2018). “Development of Simple Sequence Repeat Markers in Hazelnut (*Corylus avellana* L.) by Next-Generation Sequencing and Discrimination of Turkish Hazelnut Cultivars.” *Plant Molecular Biology Reporter* 36 (5–6): 800–811. https://doi.org/10.1007/s11105-018-1120-0.

Ozturk, Süleyman Can, Saniye Elvan Ozturk, Ibrahim Celik, Franci Stampar, Robert Veberic, Sami Doganlar, Anita Solar, and Anne Frary. (2017). “Molecular Genetic Diversity and Association Mapping of Nut and Kernel Traits in Slovenian Hazelnut (*Corylus avellana*) Germplasm.” *Tree Genetics and Genomes* 13 (1). https://doi.org/10.1007/s11295-016-1098-4.

Painter, Jodie N., Dale R. Nyholt, and Grant W. Montgomery. (2011). *Association Mapping*. *Methods in Molecular Biology*. Vol. 760. https://doi.org/10.1007/978-1-61779-176-5_3.

Şahin Anıl, Haydar Kurt, Aysun Akar, Çiğdem Bulam Köse. (2016). “Hazelnut Culture in Turkey.” *Cronica Hortculture* 56 (4).

Schulman, Alan H. (2007). “Molecular Markers to Assess Genetic Diversity.” *Euphytica* 158 (3): 313–21. https://doi.org/10.1007/s10681-006-9282-5.

Seyhan, Ferda, Guner Ozay, Sena Saklar, Erdal Ertaş, Gülçin Satir, and Cesarettin Alasalvar. (2007). “Chemical Changes of Three Native Turkish Hazelnut Varieties (*Corylus avellana* L.) during Fruit Development.” *Food Chemistry* 105 (2): 590–96. https://doi.org/10.1016/j.foodchem.2007.04.016.

Taş, Neslihan Göncüoğlu, and Vural Gökmen. (2015). “Profiling Triacylglycerols, Fatty Acids and Tocopherols in Hazelnut Varieties Grown in Turkey.” *Journal of Food Composition and Analysis* 44: 115–21. https://doi.org/10.1016/j.jfca.2015.08.010.

Topkafa, Mustafa, Hamide Filiz Ayyildiz, and Huseyin Kara. (2019). “Fruit Oils: Chemistry and Functionality.” *Fruit Oils: Chemistry and Functionality*. https://doi.org/10.1007/978-3-030-12473-1.

Tunç, İlknur, Firdevs Çalışkan, Gülcan Özkan, and Erkan Karacabey. (2014). “Mikrodalga Destekli Soxhlet Cihazı Ile Fındık Yağı Ekstraksiyonunun Yanıt Yüzey Yöntemi Ile Optimizasyonu.” *Akademik Gıda* 12 (1): 20–28.

Vieira, Maria Lucia Carneiro, Luciane Santini, Augusto Lima Diniz, and Carla de Freitas Munhoz. (2016). “Microsatellite Markers: What They Mean and Why They Are so Useful.” *Genetics and Molecular Biology* 39 (3): 312–28. https://doi.org/10.1590/1678-4685-GMB-2016-0027.

Wang, Ming, Noelle Barkley, and Tracie Jenkins. (2009). “Microsatellite Markers in Plants and Insects. Part I: Applications of Biotechnology.” *Genes, Genomes and Genomics* 3 (1): 54–67.

Xu, Y. X., and M. A. Hanna. (2010). “Evaluation of Nebraska Hybrid Hazelnuts: Nut/Kernel Characteristics, Kernel Proximate Composition, and Oil and Protein Properties.” *Industrial Crops and Products* 31 (1): 84–91. https://doi.org/10.1016/j.indcrop.2009.09.005.

Xu, Yixiang, Edward N. Sismour, John Parry, Milford A. Hanna, and Haiwen Li. (2012). “Nutritional Composition and Antioxidant Activity in Hazelnut Shells from US-Grown Cultivars.” *International Journal of Food Science and Technology* 47 (5): 940–46. https://doi.org/10.1111/j.1365-2621.2011.02925.x.

Zor, Tsaffrir, and Zvi Selinger. (1996). “Linearization of the Bradford Protein Assay Increases Its Sensitivity: Theoretical and Experimental Studies.” *Analytical Biochemistry* 236 (2): 302–8. https://doi.org/10.1006/abio.1996.0171.