MOLECULAR GENETIC ANALYSIS IN SESAME (*Sesamum indicum 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
Ümmü SEYİTALİOĞLU

September 2010

We approve the thesis of **Ümmü SEYİTALİOĞLU**

_____
Assoc. Prof. Dr. Sami DOĞANLAR
Supervisor

_____
Prof. Dr. Anne FRARY
Committee Member

_____
Assoc. Prof. Dr. Bahattin TANYOLAÇ
Committee Member

28 September 2010

_____
Assoc. Prof. Dr. Sami DOĞANLAR
Head of the Department of
Molecular Biology and Genetics

_____
Assoc. Prof. Talat YALÇIN
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGEMENTS

I would like to express my sincere and deepest gratitude to my supervisor Assoc. Prof. Dr. Sami DOĞANLAR and Prof. Dr. Anne FRARY for their encouragement, understanding and excellent support during my MSc study.

I would like to express my thankfulness to the friends with whom I've worked in the Plant Molecular Genetics Lab and my colleagues for their kindness and help.

I also thank all the members of the department of Molecular Biology and Genetics, Institute of Technology for their kindness, friendship and support.

Finally, I gratefully thank my family for their excellent support, understanding, patience and encouragement.

ABSTRACT

MOLECULAR GENETIC ANALYSIS IN SESAME
(*Sesamum indicum L.*)

In this study, 156 Turkish sesame (*Sesamum indicum L.*) accessions were characterized morphologically and 161 accessions were characterized genotypically. A total of 15 criteria were used for morphological characterization. Simple Sequence Repeats (SSRs, microsatellites) were used to characterize genetic variability among Turkish accessions. A total of 318 new EST based SSRs were developed for genotypic analysis. Also, these microsatellites were screened in Korean-Japan and African sesame accessions (parents) for future map construction studies.

According to the morphological analysis, some characters showed variation whereas some characters had no variation among accessions. For example, height of first capsule, capsule number per plant, plant height, number of branches, number seeds/capsule, days to 1st flower and days to %50 flower characters showed variation among accessions. However, stem hair, leaf hairs, axil flower number, number of carpels/capsule and capsule hairs showed variation for only a few accessions. Also, growth habit, branching and capsule splitting characters had no variation.

In accordance with genetic characterization, survey results showed only five polymorphic markers among 318 EST based SSR markers. Also, in this study, genetic distance of Turkish sesame accessions was calculated using DICE matrix and UPGMA (Unweighted Pair Group Method) arithmetical averages via 19 EST based SSR fragments. Genetic analysis showed that Turkish sesame accessions have fairly low genetic diversity. As a result, genetic diversity among Turkish sesame accessions were defined and the genetic relationships of Turkish sesame accessions were established.

ÖZET

SUSAM (*Sesamum indicum L.*)' DA MOLEKÜLER GENETİK ANALİZLER

Bu çalışmada, 156 Türk susam hattı morfolojik, 161 Türk susam hattı ise genotipik olarak karakterize edilmiştir. Morfolojik karakterizasyon için 15 fenotipik özellik kullanılmıştır. PCR tabanlı moleküler markırlardan, Simple Sequence Repeats (SSRs, microsatellites) moleküker karakterizasyonda kullanlmıştır. Bu genotipik karakterizasyon için toplam 318 yeni EST temelli Simple Sequence Repeats (SSRs, microsatellites) geliştirilmiştir. Ayrıca bu SSR markır'ları kore-japon ve afrika hatlarında test edilmişlerdir. Bu hatların ileri çalışmalarda genetik haritalama populasyonu için anne ve baba olması düşünülmektedir.

Yapılan morfolojik analizlere göre bazı karakterler hatlar arasında varyete gösterirken, bazı karakterler açısından varyete görülmemiştir. Örneğin, ilk kapsül yükekliği, bitkide kapsül sayısı, bitki yüksekliği, dallanma sayısı, kapsülde tohum sayısı, ilk çiçeklenme tarihi, %50 çiçeklenme tarihi karakterleri hatlar arasında varyete göstermiştir. Ancak, gövde tüylülüğü, yaprak tüylülüğü, yaprak koltuğunda çiçek sayısı, kapsülde karpel sayısı ve kapsül tüylülüğü karakterleri yalnızca bir kaç hat arasında varyete göstermiştir. Ayrıca, büyüme tipi, dallanma durumu ve kapsül çatlama durumu karakterlerinde varyete görülmemiştir.

Genetik analizlere göre, anne ve baba da yapılan taramada 318 EST temelli SSR arasından yalnızca 5 tane polimorfik marker bulunmuştur. Ayrıca bu çalışmada Türk susam hatlarının genetik uzaklıkları 19 EST based SSR fragmenti aracılığıyla DICE matrix ve UPGMA algoritmaları kullanılarak analiz edilmiştir. Genetik analizler göstermiştir ki, Türk susam hatları genetik olarak çok az çeşitliliğe sahiptir. Sonuç olarak Türk susam hatlarının genetik çeşitliliği belirlenmiş ve filogenetik ağacı oluşturulmuştur.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. Morphology and Biology of Sesame

Sesame is one of the most ancient and important oil seed crops (Mabberley, et al. 1997). Sesame, *Sesamum orientale* L., is also known as *S. indicum* L. The diploid chromosome number of sesame is 2n=26 and it is usually self-pollinated (Pathirana 1994). Sesame belongs to the Pedaliaceae family. The Pedaliaceae family contains 16 genera and 60 species and is a small family. The most important genus of this family is Sesamum. Sesame is typically an annual species. There are lots of varieties of *Sesamum indicum L.* according to the size, form and colour of flowers, seed size, colour and composition. Another example of variation is that some varieties are highly branched whereas others are unbranched. (Peter 2004)

Sesame is generally 0.5-4.5 m tall depending on environmental conditions. Leaves of sesame plants are generally 7.5-2.5 cm long, flat and simple or tri-lobed. The first true leaves are usually small and undivided then they increase in size and the fourth or fifth leaves are largest. Flowers are white, pink or mauve-pink. In general, the unbranched varieties mature earlier. Mature leaves and stems change color from green to yellow. The bell-shaped white to pale rose flowers begin to develop in the leaf axils 6-8 weeks after planting and this continues for several weeks. The calyx is small, five parted, and the corolla is tubular, campanulate (bell-shaped), five lobed, and about 3 cm long. The lower corolla lobe is longer and forms a lower lip. The corolla exterior and interior are pigmented with colors ranging from white to pink, violet, red and maroon. There are five stamens in each flower which are inserted at the corolla base. Temperatures below 15 or above 40^0 C cause pollen strerility, reduced fertilization and lower seed set. Fruit is capsular, oblong and quadrangular with four deep grooves and is 1.5-5 cm long. Capsule length can range from 2.5 to 8 cm with a diameter of 0.5 to 2 cm and number of capsule loculi from 4-12. The seeds of sesame are black, brown or white; 2.5-3 mm long and approximately 1.5 mm wide (Wiley 2005).

Sesame is grown for its seeds, prized oil, and oil paste. The oil paste, tahini, is obtained by grinding the seeds. The seed is also used on breads and cakes. Sesame is useful as an extra rich source of protein in many developing countries (Uzun, et al. 2002). Sesame seeds contain 50-60% oil. Sesame is known as the queen of oil seeds because its oil not only has nutritive value but also is of high quality and quantity (Bedigian 2000). Lignan antioxidants are present in the oil and are unique for sesame. The lignans sesamin and sesamolin and their derivatives prevent oxidation of the oil and give it long shelf life and stability (Brar and Ahuja 1979). According to the examination of different cultivars worldwide, variation of the fatty acid composition of sesame oil is very high (Yermanos, et al. 1972, Brar 1982). The protein content of sesame seed is about 35% (Table 1) and it is rich in tryptophan, methionine and calcium (1.3%) and is a valuable source of minerals (Johnson, et al. 1979).

Table 1.1. Approximate composition of whole sesame seed
(Source: Peter 2004)

Constituent %	Joshi (1961)	Smih (1971)	Gopalan et al. (1982)	Weiss (1983)
Moisture	5.8	8.0	5.3	5.4
Protein	19.3	22.0	18.3	18.6
Fat	51.0	43.0	43.3	49.1
Carbohydrate	21.2	21.0	25.0	21.6
Ash	5.7	6.0	5.2	5.3

According to research, sesame has many beneficial effects for human health. For instance, scientists showed that sesame leads to reduction of total serum cholesterol and low density lipoprotein (LDL) cholesterol and improvement of antioxidant capacity in hypercholestrolemic patients (Chen, et al. 2005). Sesame also increases vitamin E concentrations in plasma (Frank 2005). Shahidi et al. (2006) showed significant levels of total phenolic compounds, total antioxidant capacity and free radical scavenging capacity of white and black sesame seeds. Phenolic compounds are very important antioxidants in plants because they can stabilize radical intermediates via donating hydrogen atoms or an electron and prevent the oxidation of various biological molecules (Cuvelier, et al. 1992). One of the phenolic compounds of sesame, sesamol,

is also reported to have strong antimutagenic (Kaur and Saini 2000) and chemopreventive (Kapadia, et al. 2002) effects.

In addition to its effects on animals, sesame has significant effects on microorganisms. When sesame and pearl millet were grown together, reduced numbers of the root parasite *Striga hermonthica L.* were observed (Hess and Dodo 2004). Striga is an important biotic constraint for pearl millet. Sesame also has an antimicrobial effect on a gram negative bacterium, *Klebsiella* sp. which causes human urinary infections (Costa, et al. 2007). Also, Laurentin et al. (2007) showed that sesame has a fungicidal effect on the fungal plant pathogens *Fusarium oxysporum f. sp.* and *Macrophomina phaseolina*.

Despite these important features of sesame, research on sesame is rare. The International CGIAR (Consultative Group on International Agricultural Research) agency is not mandated for research on sesame (Laurentin and Karlovsky 2006). Sesame also does not have a place in the crops selected for examination by The International Plant Genetic Resources Institute (now renamed Bioversity International). Growers and consumers of sesame are generally in poor nations and therefore, few researchers are interested in it (Bedigian 2003). Also, information about genetic diversity and germplasm collection in sesame is very limited. According to the available studies, India, China, Central Asia, Near East and Abyssinia (currently Ethiopia and Eritrea) are centers of diversity for sesame (Laurentin and Karlovsky 2006).

1.2. Ecology and Distribution

The growing period of sesame is 70 to 150 days. Sesame can be grown under high temperature and it does not need extra rainfall. Also it is a good crop in rotation. Sesame is normally self-pollinated but sometimes cross-pollination by insects is possible. There are approximately 50-100 or more seeds in each fruit. Seed maturation occurs 4-6 weeks after fertilization.

Although, sesame has been cultivated in the Harappan (northeast Pakistan), Mesopotamian, and Anatolian eras for over 5000 years, it was first cultivated and domesticated in India (Bedigian 2003). In Africa, sesame is a fundamental source of protein and there are many wild types. Sesame has also been dispersed to many places in the world. Currently, sesame is grown in tropical and subtropical areas (Ashri 1998).

According to FAO (Food and Agriculture Organization) more than 7 million hectares are used for sesame production in the world and every year approximately 3.5 million tons of product is obtained. India, Myanmar and China are the top producers, respectively. In Turkey, 38.000 hectares area were harvested in 2007 and Turkey ranks 24th according to the world production.

Sesame mostly has indeterminate growth and its seeds are dispersed via shattering. For these reasons, mechanized harvesting is inconvenient for sesame and its commercial value is limited. Sesame is generally harvested by traditional methods. In this method, plants are cut, bound and gathered to dry. After drying, bundles are inverted and shaken. This is a time consuming process and requires cheap labour. In developing countries, sesame is a small farmers' crop (Bhat, et al. 1999). As previously mentioned, sesame is not adapted for mechanized farming because of its indeterminate growth, uneven ripening of capsules and shattering. Like other plant crops, sesame is suseptible to biotic and abiotic stresses such as diseases, pests, and drought (Bhat, et al. 1999). Despite the fact that sesame is a cultivated plant, the shattering and indeterminate growth traits of sesame are like wild type. In addition to these wild type aspects, lack of improved cultivars, susceptibility to diseases, pests and environmental stresses lead to low yield in sesame cultivation.

1.3. Genetic Diversity

At the present time, classic agricultural techniques provide limited crop imrpovement and new agricultural methods are needed because of problems like the increase in world population; the decrease of available lands because of nonsustainable farming, erosion and soil degradation; water problems; and global climate changes. Also, in classical breeding methods, crossing and selection of superior recombinants from among many segregation products must be done and this process is time consuming.

Genetic similarity causes vulnerabilty of crops to epidemics and environmental disasters. The availability of plant genetic resources and genetic diversity allows the plant to adapt to changing environments such as new pests, diseases and climatic conditions. Therefore, protection of genetic diversity is very important. However, domestication, cultivation and breeding through the ages have resulted in more

productive accessions but the loss of genetic variability. Although many factors influence genetic diversity, level of polymorphism is the product of the effective population size (Ne), mutation rate and selection (which includes positive trait selection, balancing selection, line selection, and diversifying selection) according to the neutral theory of evolution. A relationship between recombination and nucleotide diversity was also shown in some organisms such as *Drosophila melanogaster, Lycopersicon spp.,* and *Beta vulagaris*. Background selection is also a fundamental factor affecting nucleotide diversity in plants. Background selection occurs when diversity is reduced at neutral sites because of selection against linked deleterious alleles that have arisen by mutation. Also selfing reduces the effective recombination rate. Therefore, outcrossing is a desired feature for maintaining a high level genetic diversity (Buckler and Thornsberry 2002).

Table 1.2. Factors that impact nucleotide diversity.
(Source: Buckler and Thornsberry 2002)

Factor	**Correlation with diversity**	**Scope**
Mutation rate	Positive	Often whole genome
Population size and structure	Positive	Whole genome
Outcrossing	Positive	Whole genome
Recombination	Positive	Whole genome
Positive trait selection	Negative	Individual genes
Line selection	Negative	Whole genome
Diversifying selection	Positive	Individual genes
Balancing selection	Positive	Individual genes
Background selection	Negative	Individual gene or whole genome
PCR problems	Negative	Individual genes

Living seed collections are sources of genetic variation for improving agricultural crops. Germplasm banks contain cultivated and wild type plants to protect genetic diversity. For maximum conservation of genetic diversity, saving all variant types is important. Although, germplasm collections are crucial to protect crops, some accessions can be unnecessary because they do not contribute to genetic variation and this causes loss of time and money. To avoid this, scientists previously classified accessions according to their morphological phenotypes. However, this approach is not efficient and is sometimes misleading. Therefore, a new approach was developed. This

approach is looking at the genotype of the collected plants (Tanksley and McCouch 1997). To create a germplasm collection, determination of genetic diversity is the first step. In this way, undesired accessions can be discarded and some good features that can not be determined by phenotype can be found.

1.4. Marker Systems

Measuring genetic variation is very useful for selective breeding, rapid domestication and/or conservation in populations or species and genetic markers are essential tools for quick detection and characterization of genetic variation. Genetic markers are specific locations on chromosomes and they are used for genome analysis. There are two types of genetic markers: morphological and molecular markers.

1.4.1. Morphological Markers

Morphological markers are traits that can be observed as visible phenotypic effects. Morphological markers depend only on phenotype and do not require any biochemical or molecular techniques. Although morphological markers can be useful, they have disadvantages because they affect phenotype, are limited in number and may be difficult or time-consuming to determine. Therefore, they are not adequate for determining genetic variation or for plant breeding.

1.4.2. Molecular Markers

Molecular markers are neutral sites of variation at the DNA sequence or protein level. Molecular markers are classified into two types: biochemical and DNA markers. These markers determine polymorphism between plant accessions or species for breeding, mapping or genetic variation surveys. **Biochemical markers** show polymorphism at the protein level. Isozymes are generally used as biochemical markers. Isozymes are differently charged proteins and can be determined using electrophoretic procedures. Isozymes were used as a molecular marker system in early plant diversity studies. Isozymes provide genetic information as codominant markers. However, their

numbers are very limited and also they can be altered by post-translational modifications (Hamrick, et al. 1990). **DNA markers** show polymorphism at the DNA level. There are many types of DNA marker systems, some of which are described below.

1.4.2.1. RFLP (Restriction Fragment Length Polymorphism)

In this method, genomic DNA is cut at restriction sites by restriction enzymes and then, southern blotting is performed to analyse polymorphism of the resulting DNA fragments. Filter-immobilized DNA is hybridized to radiolabeled probe DNA. Polymorphism derives from different nucleotides at restriction sites. Probes are generally 500 to 3000 base pairs. They are obtained by cloning of genomic or cDNA segments. RFLP is a codominant marker system. It was used extensively in the past, however, after the discovery of polymerase chain reaction and the development of PCR-based markers, RFLP is rarely used.

1.4.2.2. Random Amplified Polymorhic DNA (RAPD)

In the RAPD marker system, short random oligonucleotide sequences (approximately 10 bases) are used to amplify genomic DNA. Sequence information is not required. RAPD is simple, fast, relatively cheap and widely used for genetic diversity studies, construction of genetic maps, and marker assisted selection. However, this method is difficult to repeat and it is a dominant marker. Arbitrarily primed polymerase chain reaction (AP-PCR) and DNA amplification fingerprinting (DAF) techniques are new methodologies that are derived from RAPD.

1.4.2.3. SSR (Simple Sequence Repeats or Microsatellites)

SSRs are a group of repetitive DNA sequences. They occur abundantly and at random throughout most eukaryotic genomes. Microsatellite markers are permanent, highly informative resources for genotyping and management of germplasm collections and may also be useful for mapping of traits of interest. Also these markers can be used

for cultivar identification, determination of hybridity, analysis of genepool variation and as diagnostic markers for traits of economic value.

Plant genomes contain large numbers of simple sequence repeats (SSRs) or microsatellites. They can be at many different loci and they can be polyallelic. Dinucleotides are the most abundant type of SSRs (Morgante and Olivieri 1993). SSRs provide considerable polymorphism due to the variation in the number of repeating units.

Advantages of SSR markers are that they have good genome coverage and are multiallelic. They are codominantly inherited, reproducible, simple, PCR-based, extremely polymorphic and highly informative because of the number and frequency of detected alleles. SSR markers have the ability to distiguish between closely related individuals. However, they are expensive to develop, require sequence information and specific primers that usually do not work in other species (Morgante and Olivieri 1993) Nevertheless after collection of DNA sequence data which include expressed sequence tags (EST), this marker system has become more efficient for genomic analyses. Generally, these analyses include functional diversity, genome mapping and marker assisted selection in crop species (Varshney, et al. 2005). It is estimated that 2-5 % of all plant derived ESTs have SSRs (Kantety, et al. 2002). Also, one of the most important benefits of using EST-based SSRs is transferability of these markers among species because they are from more conserved regions of the genome. EST-based SSRs are generally polymorphic (Bandopadhyay, et al. 2004, Fraser, et al. 2004, Pashley, et al. 2006). EST-SSRs also are crucial for use in basic evolutionary applications (Ellis, et al. 2007).

1.4.2.3.1. Expressed Sequenced Taqs (ESTs)

Expressed sequence tags are sequenced nucleotide assemblies. They are produced from cDNAs which are small DNA molecules reverse transcibed from the cellular mRNA population. ESTs are essential because of their usage for gene discovery, genome annotation and comparative genomics. Since the first study, millions of ESTs have been sequenced from distinctly annotated species, representing a wide taxonomic variety of fungi, plants and animals. Now, the NCBI EST database (dbEST)

has over 36 million EST sequences from over 1100 taxa (Ellis and Burke 2007). ESTs are crucial resources for SSR marker development (Varshney, et al. 2005).

1.4.2.4. AFLP (Amplified Fragment Length Polymorphism)

The amplified fragment length polymorphism method uses restriction enzyme digestion of total genomic DNA with PCR amplification and electrophoresis of fragments. So, it determines fragment length polymorphism. AFLP markers are widely distributed throughout the genome. AFLP is a dominant marker system. It can be a cost-effective alternative to codominant markers such as SSRs and single nucleotide polymorphisms (SNPs). Individual AFLP loci are less informative than codominant loci but AFLP has significant statistical power from its ability to produce a large number of fragments for each primer combination.

AFLP is useful in a wide range of applications including linkage mapping, parentage analysis, measuring genetic diversity, identifying hybrids and cultivars, population genetics, reconstruction of phylogenies, population assignment, developing single-locus sequence characterized amplified region (SCAR) markers, high resolution mapping and marker-assisted selection. The advantages of the AFLP method are that we can detect a large number of loci and a great deal of polymorphism without knowing sequence information. However, AFLP is technically difficult and is expensive to set up (Meudt, et al. 2007).

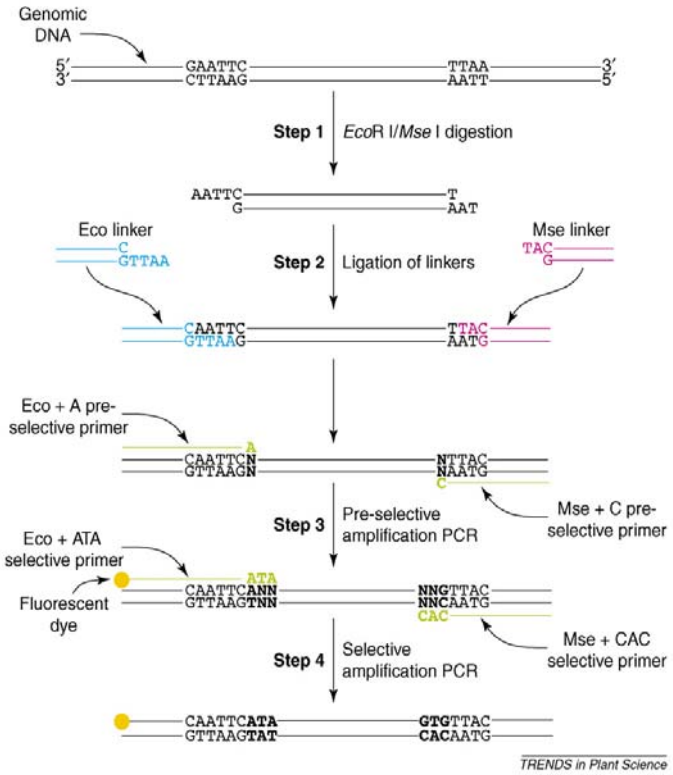

Figure.1.1. Basic steps of AFLP
(Source: Meudt, et al. 2007)

1.4.2.5. Sequence Related Amplified Polymorphism (SRAP)

Sequence related amplified polymorphism is a simple PCR-based technique. SRAP provides amplification of open reading frames. In this technique, only two primers are used. They are 17-21 bp length and their sequences are random. Primers contain AT-, GC- rich cores to amplify intragenic fragments. SRAP primers consist of two parts. The first part is called the core sequence (13-14 bp) and it has no specific constitution. Core sequence is followed by three selective nucleotides at 3' end. Core sequences must be different in forward and reverse primers (Agarwal, et al. 2008, Li and Quiros 2001).

Sequencing studies showed that SRAP polymorphisms are derived from two events. The first are fragment size changes because of insertions and deletions which result in codominant markers. The second are nucleotide changes which result in dominant markers (Agarwal, et al. 2008). SRAP markers are used for multiple aims in different crops such as map construction, gene tagging and genetic diversity studies. The application of this technique is simple and cheap. The SRAP marker system gives

reliable results. However, it gives multiple band results and therefore can be complex to analyse.

1.4.2.6. Cleaved Amplified Polymorphic Sequences (CAPS)

CAPS markers can also be called PCR-RFLP markers because they are very similar to RFLP in their use of restriction enzymes to detect polymorphism. In CAPS, PCR is performed before digestion with restriction enzymes and digested DNA is analyzed with agarose or polyacrylamide gels. Restriction fragment length polymorphism can be detect via CAPS even if there is only a single base change. The primers are obtained from sequence information in databases of genomic or cDNA sequences or cloned RAPD bands. The CAPS markers are codominant and locus specific. CAPS markers can be used for genotyping, positional or map based cloning and molecular identification studies (Agarwal, et al. 2008).

1.4.2.7. Single Nucleotide Polymorphism (SNPs)

SNPs are single nucleotide variations in the genome sequence of individuals. They are abundant, widely distributed in the genome and their distribution shows high variation among species. For instance, in maize there is 1 SNP per 60-120 bp whereas there is 1 SNP per 1000 bp in humans. They are generally widespread in non-coding regions. They can be identified by sequencing or denaturing high-performance liquid chromatography or *in silico*, aligning and comparing multiple sequences of the same region from public genome and expressed sequence (EST) databases. They can be genotyped either by allele specific hybridization, primer extension, oligonucleotide ligation or invasive cleavage. They are codominant markers and high output is obtained. Therefore, SNPs are an attractive genetic marker (Agarwal, et al. 2008).

1.4.3. Molecular Markers in Sesame

According to studies of morphological variation, sesame shows extensive variation (Bedigian, et al. 1986, Baydar, et al. 1999, Bisht, et al. 1998, Xiurong, et al.

2000). For example, diversity of an Indian sesame collection was determined for 100 accessions representing different agro-ecological zones for morphological and agronomic characters. The accessions were classified into seven clusters to create a core collection of sesame (Bisht, et al. 1998). A sesame germplasm collection in China was also established via morphological grouping (Xiurong, et al. 2000). Another morphological study was performed by Baydar in 2005. In this study, to improve the ideal sesame plant type, classic breeding techniques and examination of generations were applied based on eight features. Consequently, researchers showed that branching type is related with high yield and that plants with low yield contain high oil content. A similar study was performed by Sharmila et al. (2007). They found additive, dominant, and epistatic gene interactions for seven quantitative traits via generation mean analysis in different sesame plants.

In Turkey, Uzun et al. (2006) compared determinate and indeterminate types for agronomic traits and they showed that determinate mutant types have some disadvantages and they need further development. In previous work, Uzun et al. (2002) reported the oil content and fatty acid composition of determinate sesame types. They found that oil content of determinate types is close to their wild types but they have low seed yield. Other research was realized by Baydar et al. (1999) who classified Turkish sesame populations according to yield, oleic acid, and linoleic acid and determined line selections.

Genetic variability in sesame has also been researched by molecular techniques. Isozymes, RAPD, ISSR and AFLP have been used as molecular markers to date. The first molecular approach used to examine sesame genetic diversity was performed by Isshiki et al. in 1997. They used isozymes for determination of genetic variation in 68 accession of cultivated sesame (12 from Japan, 15 from Korea, and 41 from Thailand). As a result, only one enzyme system isocitrate dehydrogenase (IDH) showed variation. Bhat et al. (1999) evaluated genetic diversity of exotic sesame and Indian germplasm via RAPD markers. They found a high level of genetic diversity but they showed that Indian germplasm has more genetic variation than exotics. A similar study with RAPD was performed by Ercan et al. (2004) in Turkey and they showed important variation among populations. To determine genetic variation in sesame populations, another study was done using ISSR molecular markers (Kim, et al. 2002). They determined genetic diversity among 75 sesame accessions of Korean and exotic sesame. The

accessions clustered into seven groups and showed that different geographical origins are not completely distinct.

Recent studies about sesame genetic diversity were performed using the AFLP marker system. In 2006 Lurentin et al. performed AFLP analysis to examine genetic relationships and diversity in sesame germplasm. They used 32 sesame accessions from the Venezuelan germplasm collection which represents five diversity centers. Consequently, they tried eight primer combinations and recorded 457 AFLP marker that were 93% polymorphic. They found high genetic variability which was independent of geographical origin. Also, in 2007 Ali et al. used AFLP for determining the genetic diversity of 96 sesame accessions collected from different parts of the world and they found low (35%) genetic diversity.

Except for genetic diversity studies, there is only one molecular genetic study related with trait mapping in sesame. Uzun et al. (2003) identified a molecular marker linked to the closed capsule mutant trait via the AFLP method. The closed capsule mutant (induced via gamma ray irradiation) prevents shattering, a major problem for sesame production. Scientists used 72 primer combinations and one closely linked AFLP marker was found. This marker will be used for breeding to modify undesired side effects of the closed capsule mutation by marker assisted selection.

1.5. Goals

Our aim in this study was morphological and genotypical characterization of Turkish sesame accessions. For this, new EST- based SSR markers were developed and tested in Turkish sesame accessions. Genetic diversity of Turkish sesame accessions was determined EST based-SSRs (microsatellites). In this way, the genetic relationships and diversity of Turkey sesame accession were defined. After determination of genetic relationships among Turkish sesame accessions, redundant accessions can be eliminated and germplasm collection can be refined for Turkish sesame accessions. Also, EST - based SSR markers were tested in Korean-Japan and African sesame accessions for future map construction studies.

CHAPTER 2

MATERIALS AND METHODS

2.1. Materials

2.1.1. Plant Materials

The 160 national sesame accessions were obtained from the USDA-ARS Plant Germplasm Inspection Station, Beltsville, Maryland, USA. Also, 27 national sesame accessions, with known collection locations, were obtained from Aegean Agricultural Research Institute, Menemen, İzmir. Each accession was named with different pedigree number. They were grown, self pollinated and morphologically characterized in Antalya by Mutitarım Seeds Ltd.Co.

Ten seeds of each accession were planted in seedling plates containing a mixture of torf and perlite. Plants were germinated and grown in growth chamber (24-25 ^{0}C, approximately 33% humidity). All accessions, which were used in this study, with their pedigree numbers are shown in Table 2.1.

Table 2.1 Turkish sesame accessions, outgroups and their pedigree numbers, sources and locations.

Pedigree number	**Accession name**	**Source**	**Location**
170747 01 SD	*S. indicum* L.	USDA	Turkey
170745 01 SD	*S. indicum* L.	USDA	Turkey
170744 01 SD	*S. indicum* L.	USDA	Turkey
170743 01 SD	*S. indicum* L.	USDA	Turkey
170742 01 SD	*S. indicum* L.	USDA	Turkey
170739 01 SD	*S. indicum* L.	USDA	Turkey

(cont. on next page)

Table 2.1. (cont.)

Pedigree number	***Accession name***	**Source**	**Location**
170738 01 SD	*S. indicum* L.	USDA	Turkey
170737 01 SD	*S. indicum* L.	USDA	Turkey
170735 01 SD	*S. indicum* L.	USDA	Turkey
238487 01 SD	*S. indicum* L.	USDA	Turkey
238470 01 SD	*S. indicum* L.	USDA	Turkey
170722 01 SD	*S. indicum* L.	USDA	Turkey
170718 01 SD	*S. indicum* L.	USDA	Turkey
170717 01 SD	*S. indicum* L.	USDA	Turkey
170715 01 SD	*S. indicum* L.	USDA	Turkey
238469 01 SD	*S. indicum* L.	USDA	Turkey
238468 01 SD	*S. indicum* L.	USDA	Turkey
238466 01 SD	*S. indicum* L.	USDA	Turkey
238448 01 SD	*S. indicum* L.	USDA	Turkey
238447 01 SD	*S. indicum* L.	USDA	Turkey
238446 01 SD	*S. indicum* L.	USDA	Turkey
177071 01 SD	*S. indicum* L.	USDA	Turkey
177070 01 SD	*S. indicum* L.	USDA	Turkey
175908 01 SD	*S. indicum* L.	USDA	Turkey
175907 01 SD	*S. indicum* L.	USDA	Turkey
238449 01 SD	*S. indicum L.*	USDA	Turkey
238450 01 SD	*S. indicum L.*	USDA	Turkey
165021 01 SD	*S. indicum L.*	USDA	Turkey
167115 02 SD	*S. indicum L.*	USDA	Turkey
167224 01 SD	*S. indicum L.*	USDA	Turkey
167248 01 SD	*S. indicum L.*	USDA	Turkey
170733 01 SD	*S. indicum L.*	USDA	Turkey

(cont. on next page)

Table 2.1. (cont.)

Pedigree number	***Accession name***	**Source**	**Location**
238451 01 SD	*S. indicum L.*	USDA	Turkey
238453 01 SD	*S. indicum L.*	USDA	Turkey
238455 01 SD	*S. indicum L.*	USDA	Turkey
238456 01 SD	*S. indicum L.*	USDA	Turkey
238417 01 SD	*S. indicum L.*	USDA	Turkey
238416 01 SD	*S. indicum L.*	USDA	Turkey
179486 01 SD	*S. indicum L.*	USDA	Turkey
179484 01 SD	*S. indicum L.*	USDA	Turkey
179483 01 SD	*S. indicum L.*	USDA	Turkey
179482 01 SD	*S. indicum L.*	USDA	Turkey
238419 01 SD	*S. indicum L.*	USDA	Turkey
238420 01 SD	*S. indicum L.*	USDA	Turkey
238422 01 SD	*S. indicum L.*	USDA	Turkey
238435 01 SD	*S. indicum L.*	USDA	Turkey
238437 01 SD	*S. indicum L.*	USDA	Turkey
170711 01 SD	*S. indicum L.*	USDA	Turkey
170713 01 SD	*S. indicum L.*	USDA	Turkey
170714 01 SD	*S. indicum L.*	USDA	Turkey
238438 01 SD	*S. indicum L.*	USDA	Turkey
238439 01 SD	*S. indicum L.*	USDA	Turkey
238440 01 SD	*S. indicum L.*	USDA	Turkey
167343 01 SD	*S. indicum L.*	USDA	Turkey
238429 01 SD	*S. indicum L.*	USDA	Turkey
238428 01 SD	*S. indicum L.*	USDA	Turkey
238430 01 SD	*S. indicum L.*	USDA	Turkey
238431 01 SD	*S. indicum L.*	USDA	Turkey
238432 01 SD	*S. indicum L.*	USDA	Turkey

(cont. on next page)

Table 2.1. (cont.)

Pedigree number	***Accession name***	**Source**	**Location**
238433 01 SD	*S. indicum L.*	USDA	Turkey
238426 01 SD	*S. indicum L.*	USDA	Turkey
238423 01 SD	*S. indicum L.*	USDA	Turkey
238458 01 SD	*S. indicum L.*	USDA	Turkey
238434 01 SD	*S. indicum L.*	USDA	Turkey
170730 01 SD	*S. indicum L.*	USDA	Turkey
170729 01 SD	*S. indicum L.*	USDA	Turkey
205229 01 SD	*S. indicum L.*	USDA	Turkey
205225 01 SD	*S. indicum L.*	USDA	Turkey
205228 01 SD	*S. indicum L.*	USDA	Turkey
205227 01 SD	*S. indicum L.*	USDA	Turkey
238471 01 SD	*S. indicum L.*	USDA	Turkey
238473 01 SD	*S. indicum L.*	USDA	Turkey
238474 01 SD	*S. indicum L.*	USDA	Turkey
238475 01 SD	*S. indicum L.*	USDA	Turkey
238476 01 SD	*S. indicum L.*	USDA	Turkey
238477 01 SD	*S. indicum L.*	USDA	Turkey
238478 01 SD	*S. indicum L.*	USDA	Turkey
238479 01 SD	*S. indicum L.*	USDA	Turkey
238481 01 SD	*S. indicum L.*	USDA	Turkey
238482 01 SD	*S. indicum L.*	USDA	Turkey
238483 01 SD	*S. indicum L.*	USDA	Turkey
238485 01 SD	*S. indicum L.*	USDA	Turkey
238486 01 SD	*S. indicum L.*	USDA	Turkey
179481 01 SD	*S. indicum L.*	USDA	Turkey
240850 01 SD	*S. indicum L.*	USDA	Turkey
240848 01 SD	*S. indicum L.*	USDA	Turkey

(cont. on next page)

Table 2.1. (cont.)

Pedigree number	***Accession name***	**Source**	**Location**
240847 01 SD	*S. indicum L.*	USDA	Turkey
170726 01 SD	*S. indicum L.*	USDA	Turkey
170725 01 SD	*S. indicum L.*	USDA	Turkey
170724 01 SD	*S. indicum L.*	USDA	Turkey
170723 01 SD	*S. indicum L.*	USDA	Turkey
240846 01 SD	*S. indicum L.*	USDA	Turkey
240845 01 SD	*S. indicum L.*	USDA	Turkey
240844 01 SD	*S. indicum L.*	USDA	Turkey
238488 01 SD	*S. indicum L.*	USDA	Turkey
179035 01 SD	*S. indicum L.*	USDA	Turkey
179034 01 SD	*S. indicum L.*	USDA	Turkey
170710 01 SD	*S. indicum L.*	USDA	Turkey
170708 01 SD	*S. indicum L.*	USDA	Turkey
179033 01 SD	*S. indicum L.*	USDA	Turkey
179032 01 SD	*S. indicum L.*	USDA	Turkey
179031 01 SD	*S. indicum L.*	USDA	Turkey
177541 01 SD	*S. indicum L.*	USDA	Turkey
177540 01 SD	*S. indicum L.*	USDA	Turkey
170759 01 SD	*S. indicum L.*	USDA	Turkey
170758 01 SD	*S. indicum L.*	USDA	Turkey
170757 01 SD	*S. indicum L.*	USDA	Turkey
170755 01 SD	*S. indicum L.*	USDA	Turkey
170748 01 SD	*S. indicum L.*	USDA	Turkey
170749 01 SD	*S. indicum L.*	USDA	Turkey
170752 01 SD	*S. indicum L.*	USDA	Turkey
170753 01 SD	*S. indicum L.*	USDA	Turkey
170760 01 SD	*S. indicum L.*	USDA	Turkey

(cont. on next page)

Table 2.1. (cont.)

Pedigree number	***Accession name***	**Source**	**Location**
170761 01 SD	*S. indicum L.*	USDA	Turkey
170762 01 SD	*S. indicum L.*	USDA	Turkey
240852 01 SD	*S. indicum L.*	USDA	Turkey
240853 01 SD	*S. indicum L.*	USDA	Turkey
240854 01 SD	*S. indicum L.*	USDA	Turkey
240856 01 SD	*S. indicum L.*	USDA	Turkey
263373 01 SD	*S. indicum L.*	USDA	Turkey
263375 01 SD	*S. indicum L.*	USDA	Turkey
177072 01 SD	*S. indicum L.*	USDA	Turkey
204623 01 SD	*S. indicum L.*	USDA	Turkey
182295 01 SD	*S. indicum L.*	USDA	Turkey
182294 01 SD	*S. indicum L.*	USDA	Turkey
182293 01 SD	*S. indicum L.*	USDA	Turkey
179490 01 SD	*S. indicum L.*	USDA	Turkey
179489 01 SD	*S. indicum L.*	USDA	Turkey
179488 01 SD	*S. indicum L.*	USDA	Turkey
179487 01 SD	*S. indicum L.*	USDA	Turkey
238465 01 SD	*S. indicum L.*	USDA	Turkey
238464 01 SD	*S. indicum L.*	USDA	Turkey
238463 01 SD	*S. indicum L.*	USDA	Turkey
170727 01 SD	*S. indicum L.*	USDA	Turkey
238462 01 SD	*S. indicum L.*	USDA	Turkey
238461 01 SD	*S. indicum L.*	USDA	Turkey
238460 01 SD	*S. indicum L.*	USDA	Turkey
170728 01 SD	*S. indicum L.*	USDA	Turkey
238459 01 SD	*S. indicum L.*	USDA	Turkey
170732 01 SD	*S. indicum L.*	USDA	Turkey

(cont. on next page)

Table 2.1. (cont.)

Pedigree number	***Accession name***	**Source**	**Location**
175906 01 SD	*S. indicum L.*	USDA	Turkey
174355 01 SD	*S. indicum L.*	USDA	Turkey
174354 01 SD	*S. indicum L.*	USDA	Turkey
174353 01 SD	*S. indicum L.*	USDA	Turkey
173101 01 SD	*S. indicum L.*	USDA	Turkey
173100 01 SD	*S. indicum L.*	USDA	Turkey
170769 01 SD	*S. indicum L.*	USDA	Turkey
170768 01 SD	*S. indicum L.*	USDA	Turkey
170767 01 SD	*S. indicum L.*	USDA	Turkey
170765 01 SD	*S. indicum L.*	USDA	Turkey
170764 01 SD	*S. indicum L.*	USDA	Turkey
170763 01 SD	*S. indicum L.*	USDA	Turkey
238445 01 SD	*S. indicum L.*	USDA	Turkey
238444 01 SD	*S. indicum L.*	USDA	Turkey
238442 01 SD	*S. indicum L.*	USDA	Turkey
238441 01 SD	*S. indicum L.*	USDA	Turkey
ORHANGAZİ-99	*S. indicum L.*	AARI	Turkey
TAN-99	*S. indicum L.*	AARI	Turkey
KEPSUT-99	*S. indicum L.*	AARI	Turkey
OSMANLI-99	*S. indicum L.*	AARI	Turkey
CUMHURİYET-99	*S. indicum L.*	AARI	Turkey
TR 45524	*S. indicum L.*	AARI	Adana
TR 45572	*S. indicum L.*	AARI	Adıyaman
TR 39702	*S. indicum L.*	AARI	Antalya
TR 61609	*S. indicum L.*	AARI	Aydın
TR 38106	*S. indicum L.*	AARI	Balıkesir
TR 76589	*S. indicum L.*	AARI	Bilecik

(cont. on next page)

Table 2.1. (cont.)

Pedigree number	***Accession name***	**Source**	**Location**
TR 42870	*S. indicum L.*	AARI	Bursa
TR 68411	*S. indicum L.*	AARI	Çanakkale
TR 61927	*S. indicum L.*	AARI	Denizli
TR 45642	*S. indicum L.*	AARI	Diyarbakır
TR 38253	*S. indicum L.*	AARI	Edirne
TR 45599	*S. indicum L.*	AARI	Elazığ
TR 42145	*S. indicum L.*	AARI	Gaziantep
TR 39695	*S. indicum L.*	AARI	İçel
TR 52540	*S. indicum L.*	AARI	İzmir
TR 45543	*S. indicum L.*	AARI	Kahramanmaraş
TR 52533	*S. indicum L.*	AARI	Kars
TR 42635	*S. indicum L.*	AARI	Kırklareli
TR 50128	*S. indicum L.*	AARI	Kütahya
TR 45596	*S. indicum L.*	AARI	Malatya
TR 64094	*S. indicum L.*	AARI	Manisa
TR 45673	*S. indicum L.*	AARI	Mardin
TR 39716	*S. indicum L.*	AARI	Muğla
TR 37513	*S. indicum L.*	AARI	Siirt
TR 45707	*S. indicum L.*	AARI	Şanlıurfa
TR 38356	*S. indicum L.*	AARI	Tekirdağ
TR 68905	*S. indicum L.*	AARI	Uşak
	S. alatum L.		

2.2. Methods

2.2.1. Morphological Characterization

A total of 15 morphological features of 156 Turkish accessions of *S. indicum L.* were characterized in Antalya by Mutitarım Seeds Ltd.Co. Plant materials were planted in 22 June 2009 and 18 June 2010 seperately. Morphological analysis was performed in 2 years, then these datas were combined. These 15 traits are shown in Table 2.2.

Table. 2.2. Analysed morphological characters in Turkish sesame accessions

Trait Number	**Morphological Trait**
1	Growth Type
2	Stem hairs
3	Branching
4	Leaf Hairs
5	Axil flower number
6	Number of carpels/capsule
7	Capsule hairs
8	Capsule splitting
9	Days to 1st flower
10	Days to %50 flower
11	Height of 1st capsule
12	Plant height
13	Number of branches
14	Number of capsules
15	Number seeds/capsule

2.2.2. EST-Based SSR Marker Design

A total of 3662 transcript assemblies that belong to *Sesamum indicum L.* were downloaded from dbEST/GenBank (ftp://ftp.tigr.org/pub/data/plantta/Sesamum_indicum). These assemblies were used as data on PBC Public SSR Primer Discovery Input site (http://hornbill.cspp.latrobe.edu.au/cgi-binpub/ssrprimer/indexssr.pl) which was used to design SSR primers. Primer size (Min: 18-Max: 22), primer melting temperature Tm (Min: 50-Max: 70) and primer GC content % (Min: 50-Max: 70) parameters were considered basic criteria and at the end, 179 SSR primers were found.

Also, to design SSR makers, a total of 3328 transcript assemblies from dbEST/GenBank (http://www.ncbi.nlm.nih/entrez) was downloaded. However, some of the sequences could be repeated or so similar to each other that they could give the same SSR marker. Therefore, they were analysed by software program, and repetitive and close sequences were integrated to form contigs. In this way, the ESTs were transformed into unigenes.

The BatchPrimer3 Primer Design Input Site (http://probes.pw.usda.gov/cgi-bin/batchprimer3/batchprimer3.cgi) was used for determination of SSR markers. Unigene sequences for *Sesamum indicum L.* were used as the input file. For SSR marker design, primer size (Min: 18-Max: 22), primer melting temperature Tm (Min: 50-Max: 70), and primer GC content % (Min: 50-Max: 70) were taken into account. As a result, a total of 139 EST-based SSR markers were obtained.

All SSR markers designed from TIGR and NCBI EST databases are shown in Appendix A. with sequence ID, repeat type, repeat number, primer melting temperature and product size.

2.2.3. DNA Extraction

DNA extraction was performed from fresh leaf tissues of each seedling when they were at the 4-6 leaf stage. A CTAB-DNA extraction protocol modified according to Fulton et al. (1995) was used for DNA extraction. Quantification of DNAs was performed with a Nanodrop ND-1000 spectrophotometer and the DNA samples were stored at -20 0C in TE buffer.

2.2.4. SSR Analysis

In this study, 179 and 139 EST-SSR primer pairs were developed seperately using EST sequences of *Sesamum indicum* L.. These EST based SSR primers were screened in Korea- Japan and African sesame accessions which will be used as parents in future mapping studies. These accessions have high polymorphism according to the work of Laurentin et al. (2006).

Also, polymorphism among Turkish sesame accessions was screened via these primers. Each PCR reaction mixture was 25 µl which included 2,5 µl 10 X Buffer, 0,5 µl dNTP, 0,25 µl Taq, 18,75 µl dH2O, 2 µl DNA and 0,5 µl Forward, 0,5 µl Reverse primer. DNA concentration was approximately 50 ng/µl. PCR protocol for SSR analysis was 5 min initial denaturation at 94 ^{o}C, then 35 cycles of 45 sec of denaturing at 94 ^{o}C, 1 min at 50 ^{o}C for annealing, and extension at 72 ^{o}C for 1 min and final extension at 72 ^{o}C for 5 min. PCR products were run in 3 % agarose gels in 1X TAE buffer and they were visualized under UV light. The remaining PCR products were analysed in Qiaxcel capillary electrophoresis system which is enable to separate DNA fragments with high resolution (2-5 bp). The Qıaxcel DNA high resolution gel cartridge and QM500 method were preferred for efficient analysis.

2.2.5. Data Analysis

To determine genetic diversity among Turkish sesame accessions, SSR marker results were transformed into numerical data. For this, results were scored as present (1) and absent (0). NTSYS-pc version 2.2 (Numerical Taxonomy Multivariate Analysis System, Exeter Software, Setauket, N.Y.) software program was used for dendrogram construction. DICE matrix and Unweighted Pair Group Method with Arithmetic Averages (UPGMA) method were used for dendrogram construction.

The DICE similarity index calculates the similarity between two samples *i*,*j* with the formula GS(*i*,*j*)=2a(2a+b+c) where GS(*i*,*j*) is the similarity coefficient between samples *i* and *j*, *a* is number of polymorphic bands shared between *i* and *j*, *b* is number of bands present in *i* and absent in *j* and *c* is number of bands present in *j* and absent in *i*. The similarity dendrogram was produced by clustering the similarity data with the Unweighted Pair Group Method with Arithmetic Averages (UPGMA) and the SAHN

clustering program. Correlation between tree and data matrix was compared with the Mantel test.

CHAPTER 3

RESULT AND DISCUSSION

3.1. Morphological Analysis Results

In this study, fifteen phenotypic characters of 156 Turkish national sesame accessions were analyzed quantitatively in summer 2009 and 2010. These characters are important in term of agriculture. Means for the morphological criteria for each accession are shown in Table 3.1. Means for the Turkish sesame collection are shown in Table 3.2. According to the morphological analysis, some characters showed variation whereas some characters had no variation among accessions. For example, height of first capsule, capsule number per plant, plant height, number of branches, number seeds/capsule, days to 1st flower and days to %50 flower characters showed variation among accessions. However, stem hair, leaf hairs, axil flower number, number of carpels/capsule and capsule hairs showed variation for only a few accessions. Also, growth habit, branching and capsule splitting characters had no variation.

Table 3.1. 2009-2010 Morphological characterization for the national sesame accessions.

PI	Number	Growth Type	Stem hairs	Branching	Leaf Hairs	Leaf axil flower number	Carpel number for each capsule	Capsule Hair	Split status of capsules	First flowering time	50 % flowering time	Length of first capsule	Plant length	Branch number	Capsule number	Seed number in a capsule
PI 170747	S1	1	3	1	3	1	1	3	3	31.50	33.50	26	86.5	4	92.5	45
PI 170745	S2	1	3	1	3	1	1	3	3	33.50	37.00	27.5	91.5	3	74	45
PI 170744	S3	1	3	1	3	1	1	3	3	35.50	37.00	22.5	85	4	104.5	50
PI 170743	S4	1	3	1	3	1	1	3	3	37.50	40.00	23	77.5	4	96	42.5
PI 170742	S5	1	3	1	3	1	1	3	3	38.50	43.50	29	82.5	3.5	71	47.5
PI 170739	S6	1	3	1	3	1	1	3	3	37.00	43.50	36.5	115	3	68.5	47.5
PI 170738	S7	1	3	1	3	1	1	3	3	36.50	39.00	24	80	3.5	72	47.5
PI 170737	S8	1	3	1	3	1	1	3	3	34.50	37.00	21	70.5	3	64.5	45
PI 238487	S10	1	3	1	3	1	1	3	3	35.50	38.50	26.5	81	3.5	64	44.5
PI 238470	S11	1	3	1	3	1	1	3	3	37.50	40.50	32.5	92	3.5	93	43.5
PI 170718	S13	1	3	1	3	1	1	3	3	38.00	42.00	36.5	122	5	110.5	47.5
PI 170715	S15	1	3	1	3	1	1	3	3	33.00	37.00	40	109	5	120.5	47.5
PI 238469	S16	1	3	1	3	1	1	3	3	36.00	39.50	32	98.5	4	91	47
PI 238468	S17	1	3	1	3	1	1	3	3	40.50	47.00	34	107	3.5	82.5	45
PI 238466	S18	1	3	1	5	1	1	7	3	38.00	49.00	36.5	112	4	67.5	47
PI 238448	S19	1	3	1	3	1	1	3	3	37.00	46.00	40	113	5	94.5	45.5
PI 238447	S20	1	3	1	3	1	1	3	3	38.00	45.00	36	93.5	4.5	80.5	45.5
PI 238446	S21	1	3	1	3	1	1	3	3	32.00	35.00	34	99	3.5	73	45.5
PI 175908	S24	1	3	1	3	1	1	3	3	34.00	39.00	32	113	3.5	77	76
PI 238449	S26	1	3	1	3	1	1	3	3	32.00	35.00	26.5	81.5	3.5	63.5	47.5
PI 238450	S27	1	7	1	3	1	1	3	3	36.00	44.00	31.5	103	3.5	62.5	47
PI 167115	S29	1	7	1	3	1	1	3	3	42.00	47.00	39	85	4	78	47
PI 238451	S33	1	7	1	3	1	1	3	3	35.00	39.00	40	101	3.5	69.5	45

(cont. on next page)

Table 3.1. (cont.)

PI	Number	Growth Type	Stem hairs	Branching	Leaf Hairs	Leaf axil flower number	Carpel number for each capsule	Capsule Hair	Split status of capsules	First flowering time	50 % flowering time	Length of first capsule	Plant length	Branch number	Capsule number	Seed number in a capsule
PI 238453	S34	1	7	1	3	1	1	3	3	32.00	35.00	31	105	3	101.5	47.5
Muganlı-57	S36	1	7	1	3	1	1	3	3	35.00	39.00	39	112	3	89	47.5
PI 238417	S37	1	7	1	3	1	1	3	3	32.00	35.50	28.5	99.5	4	77.5	45
PI 179486	S39	1	7	1	3	1	1	3	3	42.00	48.00	46	108	3.5	71.5	47
PI 179484	S40	1	5	1	5	1	1	7	3	35.00	39.00	31.5	112	3	63	45
PI 238419	S43	1	3	1	3	1	1	3	3	42.00	44.50	43.5	114	4	96.5	45
PI 238420	S44	1	3	1	3	1	1	3	3	39.00	43.00	29	100	3.5	81	45
PI 238422	S45	1	3	1	3	1	1	3	3	39.00	43.00	36.5	124	3.5	70	47.5
PI 238435	S46	1	3	1	3	1	1	5	3	32.00	35.00	25	87.5	4	71.5	50
PI 238437	S47	1	3	1	3	1	1	3	3	35.00	39.00	31.5	106	3	72	47.5
PI 170711	S48	1	3	1	3	1	1	3	3	39.00	41.00	39	109	3	64.5	47.5
PI 170713	S49	1	3	1	3	1	1	3	3	39.00	41.00	40	124	4	95.5	45
PI 170714	S50	1	3	1	3	1	1	3	3	39.00	44.50	42.5	132	2.5	61	47.5
PI 238438	S51	1	3	1	3	1	1	5	3	38.00	43.50	39	122	4.5	85	50
PI 238439	S52	1	3	1	3	1	1	3	3	37.00	39.00	29	80	4	71	47.5
PI 238440	S53	1	3	1	3	1	1	3	3	37.00	39.00	28.5	98.5	3.5	81	42.5
PI 167343	S54	1	3	1	3	1	1	3	3	36.00	39.00	32.5	98.5	4.5	80.5	42.5
PI 238429	S55	1	3	1	3	1	1	5	3	43.00	46.50	45.5	111	4	88	42.5
PI 238428	S56	1	3	1	3	1	1	3	3	41.00	44.50	39	109	4.5	83	42.5
PI 238430	S57	1	3	1	3	1	1	3	3	32.00	35.50	51	126	4	86.5	45
PI 238431	S58	1	3	1	3	1	1	3	3	38.00	41.00	41.5	120	3.5	67	47.5
PI 238432	S59	1	3	1	3	1	1	3	3	42.00	44.50	44	123	4.5	110.5	42.5
PI 238433	S60	1	3	1	3	1	1	3	3	31.00	33.00	28.5	99	3	87.5	45

(cont. on next page)

Table 3.1. (cont.)

PI	Number	Growth Type	Stem hairs	Branching	Leaf Hairs	Leaf axil flower number	Carpel number for each capsule	Capsule Hair	Split status of capsules	First flowering time	50 % flowering time	Length of first capsule	Plant length	Branch number	Capsule number	Seed number in a capsule
PI 238426	S61	1	3	1	3	1	1	5	3	32.00	34.00	36.5	110	4.5	100.5	45
PI 238423	S62	1	3	1	3	1	1	3	3	33.00	36.00	36	122	3.5	85	45
PI 238458	S63	1	3	1	3	1	1	3	3	32.00	35.00	36.5	116	3	81.5	45
PI 238434	S64	1	3	1	3	1	1	3	3	33.00	36.00	36.5	111	4	114	50
PI 170730	S65	1	3	1	3	1	1	5	3	42.00	44.50	47.5	106	4.5	76.5	45
PI 170729	S66	1	3	1	3	1	1	3	3	35.00	39.00	34.5	101	3.5	88	47.5
PI 205229	S67	1	3	1	3	1	1	3	3	32.50	37.50	26	93.5	4	65	45
PI 205225	S68	1	3	1	3	1	1	3	3	34.00	37.00	38.5	132	3.5	98	45
PI 205228	S69	1	3	1	3	1	1	3	3	35.00	39.00	46.5	117	3	70	42.5
PI 205227	S70	1	3	1	3	1	1	3	3	36.00	39.00	39	121	3.5	79.5	52.5
PI 238471	S71	1	3	1	3	2	1	3	3	36.00	39.00	36.5	107	3.5	87	47.5
PI 238473	S72	1	3	1	3	1	1	3	3	43.00	46.00	42.5	128	3.5	77	45
PI 238474	S73	1	3	1	3	1	1	3	3	32.00	35.00	25	102	4.5	114	50
PI 238475	S74	1	3	1	3	1	1	3	3	34.00	39.00	45	121	3.5	67.5	47.5
PI 238477	S76	1	3	1	3	1	1	3	3	32.50	35.50	40.5	109	5	104.5	47.5
PI 238478	S77	1	3	1	3	1	1	3	3	35.50	39.00	28	88.5	4	100.5	72
PI 238479	S78	1	3	1	3	1	1	3	3	39.00	42.00	48.5	113	5	103	72
PI 238481	S79	1	3	1	3	2	1	3	3	38.50	42.00	35.5	88.5	3.5	77.5	76
PI 238482	S80	1	3	1	3	1	1	3	3	36.00	41.50	22.5	91.5	3.5	62	80
PI 238483	S81	1	5	1	5	1	1	7	3	32.00	34.50	34	106	4.5	70	76
PI 238485	S82	1	3	1	3	1	1	3	3	33.00	35.50	28	97	3.5	92	76
PI 238486	S83	1	3	1	3	1	1	3	3	37.50	41.50	32.5	99	3.5	96.5	72

(cont. on next page)

Table 3.1. (cont.)

PI	Number	Growth Type	Stem hairs	Branching	Leaf Hairs	Leaf axil flower number	Carpel number for each capsule	Capsule Hair	Split status of capsules	First flowering time	50 % flowering time	Length of first capsule	Plant length	Branch number	Capsule number	Seed number in a capsule
PI 179481	S84	1	3	1	3	1	1	3	3	34.00	37.00	34.5	103	4	119.5	76
PI 240850	S85	1	3	1	3	1	1	3	3	39.00	42.50	46.5	107	3	67.5	80
PI 240848	S86	1	3	1	3	1	1	3	3	34.00	37.00	27.5	95	3	64.5	76
PI 240847	S87	1	3	1	3	1	1	3	3	35.00	39.00	37.5	93.5	3.5	84	68
PI 170726	S88	1	3	1	3	1	1	3	3	34.50	38.50	39	110	3.5	78.5	68
PI 170725	S89	1	3	1	3	1	1	3	3	36.00	39.50	49	114	3.5	71	68
PI 170724	S90	1	3	1	3	1	1	3	3	33.00	36.00	28.5	99	3	89	68
PI 170723	S91	1	3	1	3	1	1	3	3	31.50	35.00	33	95	4.5	93.5	72
PI 240846	S92	1	3	1	3	1	1	3	3	43.00	45.00	47.5	113	4	98.5	76
PI 240845	S93	1	3	1	3	1	1	3	3	43.00	45.00	35.5	123	3.5	51	68
PI 240844	S94	1	3	1	3	1	1	3	3	45.50	48.00	37	122	4	87	72
PI 238488	S95	1	3	1	3	1	1	3	3	35.00	39.00	36.5	101	3.5	58	72
PI 179034	S97	1	3	1	3	1	1	3	3	36.00	39.00	37	96.5	3.5	197	72
PI 170710	S98	1	3	1	3	1	1	3	3	34.00	38.00	34	92	2.5	83	80
PI 170708	S99	1	3	1	3	1	1	3	3	33.00	35.00	32.5	98.5	3.5	71.5	72
PI 179032	S101	1	3	1	3	1	1	3	3	39.00	44.00	30	96	4	88.5	72
PI 177540	S104	1	3	1	3	1	1	3	3	38.00	43.00	39	94	4	81.5	84
PI 170759	S105	1	3	1	3	1	1	3	3	31.00	32.00	46.5	109	4.5	86	72
PI 170758	S106	1	3	1	3	1	1	3	3	32.00	35.00	32.5	95	3.5	83	76
PI 170757	S107	1	3	1	3	1	1	3	3	35.00	39.00	42.5	91.5	4.5	73.5	72
PI 170748	S109	1	3	1	3	1	1	3	3	39.00	45.00	42.5	99	4	66	68
PI 170749	S110	1	3	1	3	1	1	3	3	36.00	39.00	34	87.5	3.5	65.5	84

(cont. on next page)

Table 3.1. (cont.)

PI	Number	Growth Type	Stem hairs	Branching	Leaf Hairs	Leaf axil flower number	Carpel number for each capsule	Capsule Hair	Split status of capsules	First flowering time	50 % flowering time	Length of first capsule	Plant length	Branch number	Capsule number	Seed number in a capsule
PI 170752	S111	1	3	1	3	1	1	3	3	35.00	39.00	36.5	103	3.5	63	72
PI 170760	S113	1	3	1	3	1	1	3	3	34.00	38.00	36.5	90	4	80.5	72
PI 170762	S115	1	3	1	3	1	1	3	3	38.00	41.00	34	84.5	5	67.5	72
PI 240852	S116	1	3	1	3	1	1	3	3	33.50	37.50	27.5	80	3.5	55.5	76
PI 240853	S117	1	3	1	3	1	1	3	3	34.50	37.00	34	96	3.5	73	42.5
PI 240854	S118	1	3	1	3	1	1	3	3	33.00	37.00	36	83.5	3.5	68.5	45
PI 240856	S119	1	3	1	3	2	1	3	3	34.50	38.00	32.5	86.5	3.5	47.5	45
PI 263373	S120	1	3	1	3	1	1	3	3	39.00	41.50	35.5	111	3	57	45
PI 263375	S121	1	3	1	3	1	1	3	3	38.00	42.50	40	106	3.5	84	47.5
PI 177072	S122	1	3	1	3	1	1	3	3	33.00	36.00	39	100	3	68.5	45
PI 204623	S123	1	3	1	3	1	1	3	3	32.00	35.00	41.5	114	4.5	89.5	50
PI 179489	S128	1	3	1	3	1	1	3	3	32.00	35.00	45	116	3.5	80	45
PI 238465	S131	1	3	1	3	1	1	3	3	35.00	39.00	38.5	114	4.5	76.5	47.5
PI 238464	S132	1	3	1	3	1	1	3	3	34.00	38.00	38	95	4	69.5	47.5
PI 238463	S133	1	3	1	3	1	1	3	3	43.00	45.00	56.5	105	4	62	45
PI 170727	S134	1	3	1	3	1	1	3	3	34.00	39.00	38.5	88	3.5	63.5	50
PI 238462	S135	1	3	1	3	1	1	3	3	32.00	35.00	41.5	96	4.5	73.5	42.5
PI 238461	S136	1	3	1	3	1	1	3	3	32.00	36.00	32.5	90	3	69.5	42.5
PI 238460	S137	1	3	1	3	1	1	3	3	31.00	32.00	32.5	92	4.5	79.5	45
PI 238459	S139	1	3	1	3	1	1	3	3	31.00	32.00	31.5	87.5	4	63.5	47.5
PI 170732	S140	1	3	1	3	1	1	3	3	43.00	45.00	46.5	96.5	3.5	65	45
PI 173101	S145	1	3	1	3	1	1	3	3	30.00	32.00	39.5	89	4	42	50

(cont. on next page)

Table 3.1. (cont.)

PI	Number	Growth Type	Stem hairs	Branching	Leaf Hairs	Leaf axil flower number	Carpel number for each capsule	Capsule Hair	Split status of capsules	First flowering time	50 % flowering time	Length of first capsule	Plant length	Branch number	Capsule number	Seed number in a capsule
PI 173100	S146	1	3	1	3	1	1	3	3	44.00	46.00	47	114	3.5	53	47.5
PI 170769	S147	1	3	1	3	1	1	3	3	35.50	38.00	44	102	3.5	45	47.5
Özberk	S148	1	3	1	3	1	1	3	3	38.00	42.00	52.5	104	3.5	59	45
PI 170767	S149	1	3	1	3	1	1	3	3	33.00	36.00	49	118	3	53.5	45
PI 170765	S150	1	3	1	3	1	1	3	3	35.00	39.00	43	117	2.5	44	45
PI 238445	S153	1	3	1	3	1	1	3	3	35.00	38.00	37.5	104	4.5	70.5	47.5
PI 238444	S154	1	3	1	3	1	1	3	3	35.00	39.00	25.5	81	3	46.5	50
PI 238442	S155	1	3	1	3	2	1	3	3	30.50	35.50	52.5	86.5	2.5	42	47.5
PI 238441	S156	1	3	1	3	1	1	3	3	39.50	42.00	50	98.5	2.5	57	45
ORHANGAZİ	S157	1	5	1	5	1	1	7	3	34.50	40.00	40	103	3	49	45
TAN-99	S158	1	3	1	3	1	1	3	3	35.50	40.50	46.5	103	5.5	73	47.5
KEPSUT-99	S159	1	3	1	3	1	1	3	3	33.00	35.00	47.5	111	2	50.5	50
OSMANLI-99	S160	1	3	1	3	1	1	3	3	40.00	42.50	41.5	107	4	90.5	47.5
TR 45524	S161	1	3	1	3	1	1	3	3	35.00	38.00	48	116	5	73	42.5
TR 45572	S162	1	3	1	3	1	1	3	3	36.50	40.00	38	98.5	3	66.5	42.5
TR 39702	S163	1	3	1	3	1	1	3	3	34.50	37.00	30	99	3.5	61.5	42.5
TR 61609	S164	1	3	1	3	1	1	3	3	35.50	39.00	45	109	3	57.5	42.5
TR 38106	S165	1	3	1	3	1	1	3	3	33.50	36.50	25	102	2.5	49	45
TR 76589	S166	1	3	1	3	1	1	3	3	32.00	36.00	36.5	92	3.5	51	45
TR 42870	S167	1	3	1	3	1	1	3	3	31.00	33.00	19	78	3	40.5	42.5
TR 68411	S168	1	3	1	3	1	1	3	3	31.00	36.00	32.5	93	3.5	56.5	45
TR 61927	S169	1	3	1	3	1	1	3	3	35.00	38.00	22.5	81	2.5	48.5	45

(cont. on next page)

Table 3.1. (cont.)

PI	Number	Growth Type	Stem hairs	Branching	Leaf Hairs	Leaf axil flower number	Carpel number for each capsule	Capsule Hair	Split status of capsules	First flowering time	50 % flowering time	Length of first capsule	Plant length	Branch number	Capsule number	Seed number in a capsule
TR 45642	S170	1	3	1	3	1	1	3	3	35.00	39.00	37.5	101	4	61.5	45
TR 38253	S171	1	3	1	3	1	1	3	3	30.00	32.00	31.5	85	2.5	38.5	45
TR 42145	S173	1	3	1	3	1	1	3	3	31.00	32.00	34	106	2.5	46.5	45
TR 39695	S174	1	3	1	3	1	1	3	3	36.00	38.00	29	79	4	62	47.5
TR 52540	S175	1	3	1	3	1	1	3	3	36.00	39.00	34	111	3	55	45
TR 45543	S176	1	3	1	3	1	1	3	3	35.00	38.00	32.5	87	5	63.5	45
TR 52533	S177	1	3	1	3	1	1	3	3	37.00	38.00	42	112	3	45.5	42.5
TR 42635	S178	1	3	1	3	1	1	3	3	36.00	39.00	29	104	4	78	52.5
TR 50128	S179	1	3	1	3	1	1	3	3	32.00	36.00	25	92	3.5	52.5	47.5
TR 45596	S180	1	3	1	3	1	1	3	3	34.00	37.00	36.5	101	4	50.5	45
TR 64094	S181	1	3	1	3	1	1	3	3	37.00	39.00	47.5	109	2.5	41.5	50
TR 45673	S182	1	3	1	3	1	1	3	3	38.00	41.50	37.5	97.5	3.5	44	47.5
TR 39716	S183	1	3	1	3	1	1	3	3	29.00	41.50	32.5	96.5	4	53.5	52.5
TR 37513	S184	1	3	1	3	1	2	3	3	38.00	42.50	44	101	3.5	46	47.5
TR 45707	S185	1	3	1	3	1	1	3	3	39.00	42.50	39	100	4.5	52	47.5
TR 38356	S186	1	3	1	3	1	1	3	3	37.00	39.00	28.5	77.5	3.5	40.5	52.5
TR 68905	S187	1	3	1	3	1	1	3	3	31.00	33.00	24	88.5	3.5	39	50
Cumhuriyet	S188	1	3	1	3	1	1	3	3	35.00	39.00	41	105	3.5	58.5	47.5
	Mug	1	3	1	3	1	1	3	3	33.20	36.20	33.83	95.2	3.53	67.2	46.54
	Özb	1	3	1	3	1	1	3	3	35.26	37.63	34.96	95.1	3.56	55.23	47.1
	Göl	1	3	1	3	1	1	3	3	33.80	37.53	36.6	115	4.53	75.4	46.93
	Tan	1	3	1	3	1	1	3	3	33.90	38.00	36.96	110	3.63	68.53	45.73

Table 3.2. A averages for morphological characters measured in 2009 and 2010.

Trait	Averages ± Standard Error	Minimum	Maximum
Growth Type	1.0±0.0	1	1
Stem hairs	3.22±0.07	3	7
Branching	1.0±0.0	1	1
Leaf Hairs	3.05±0.02	3	5
Axil flower number	1.02±0.01	1	2
Number of carpels/capsule	1.0±0.0	1	2
Capsule Hair	3.17±0.06	3	7
Capsule splitting	3.0±0.0	3	3
Days to 1st flower	35.5±0.26	29	45.5
Days to %50 flower	39.08±0.30	32	49
Height of 1st capsule	36.28±0.58	19	56.5
Plant height	101±1.01	70.5	132
Number of branches	3.66±0.05	2	5.5
Number of capsules	73.22±1.64	38.5	197
Number of seeds/capsule	52.23±0.94	42.5	84

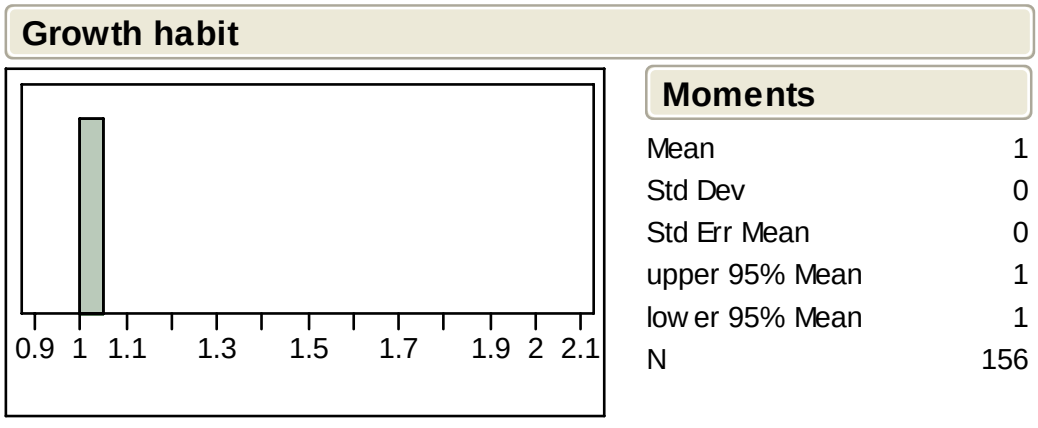

Figure 3.1. Growth habit histogram analysis of 156 Turkish sesame accessions.

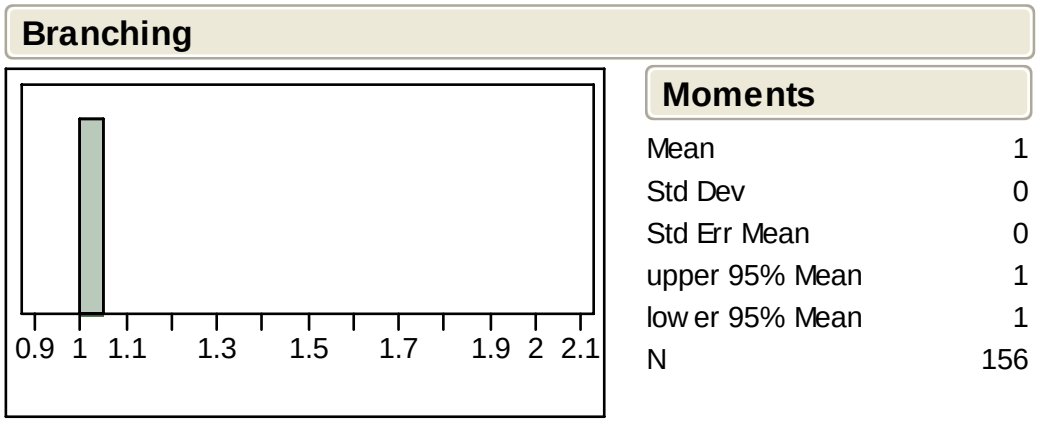

Figure 3.2. Branching histogram analysis of 156 Turkish sesame accessions.

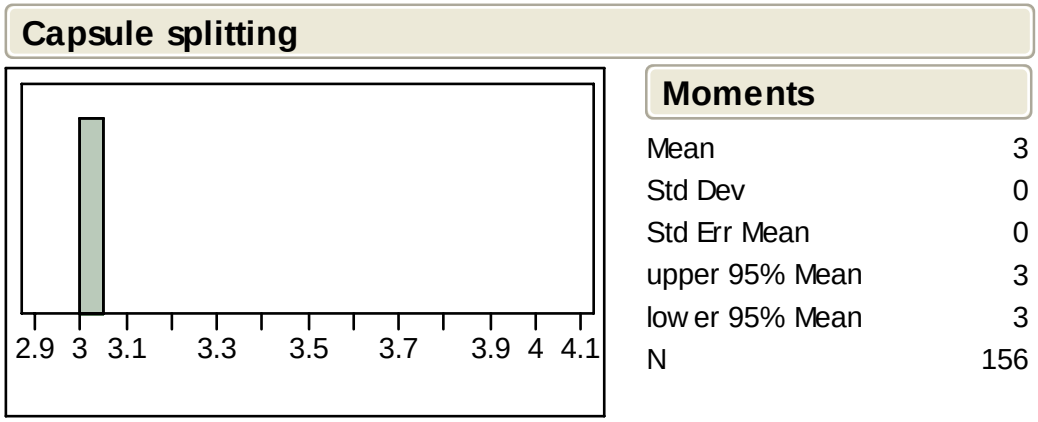

Figure 3.3. Capsule splitting histogram analysis of 156 Turkish sesame accession.

According to the phenotypic observations and analyses, growth type, branching and split status of capsules characters did not have variation among Turkish sesame accessions (Figures 3.1 to 3.3).

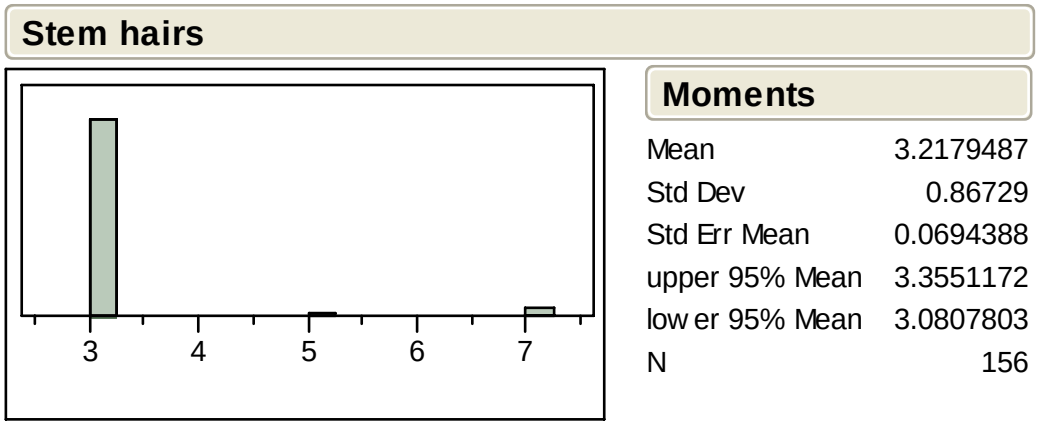

Figure 3.4. Stem hairs histogram analyses and distribution of 156 Turkish sesame accessions.

Stem hairs character showed variation among only a few accessions. Most accessions had few stem hairs with a score of 3. However, three accessions had scores of 5 and 7 accessions had scores of 7. The average for stem hairs character was 3.21.

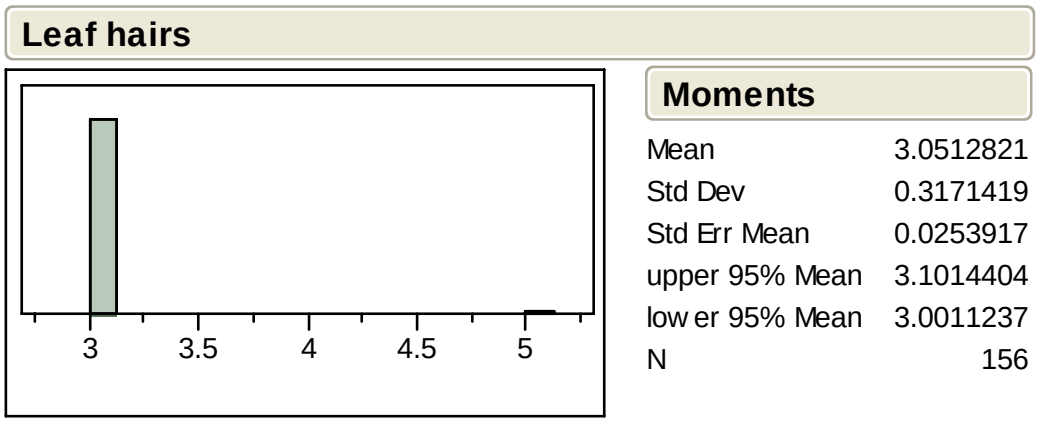

Figure 3.5. Leaf hairs histogram analysis and distribution of 156 Turkish sesame accessions.

Leaf hairs character showed polymorphism among only five accessions. The rest of the accessions had no variation. The average for the leaf hairs character was 3.05.

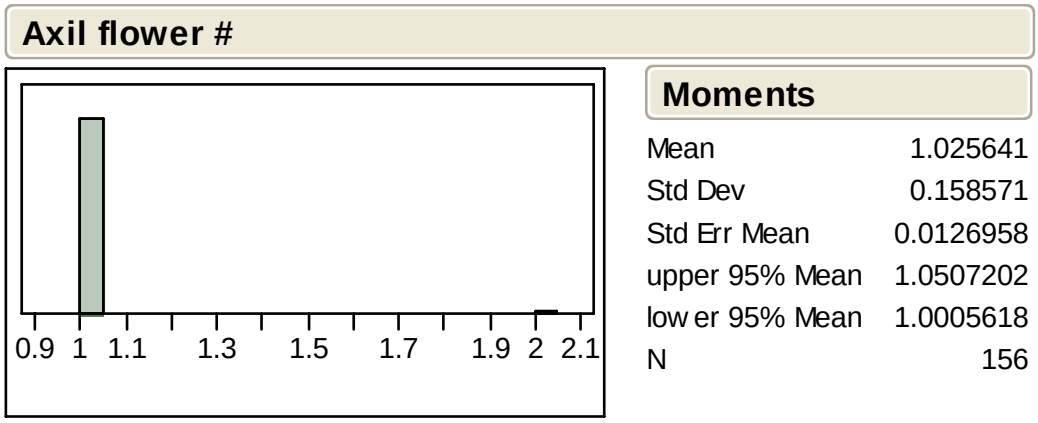

Figure 3.6. Axil flower number histogram analysis and distribution of 156 Turkish sesame accessions.

Axil flower number is like leaf hairs character, and showed polymorphism for only four accessions. The average of this character was 1.02.

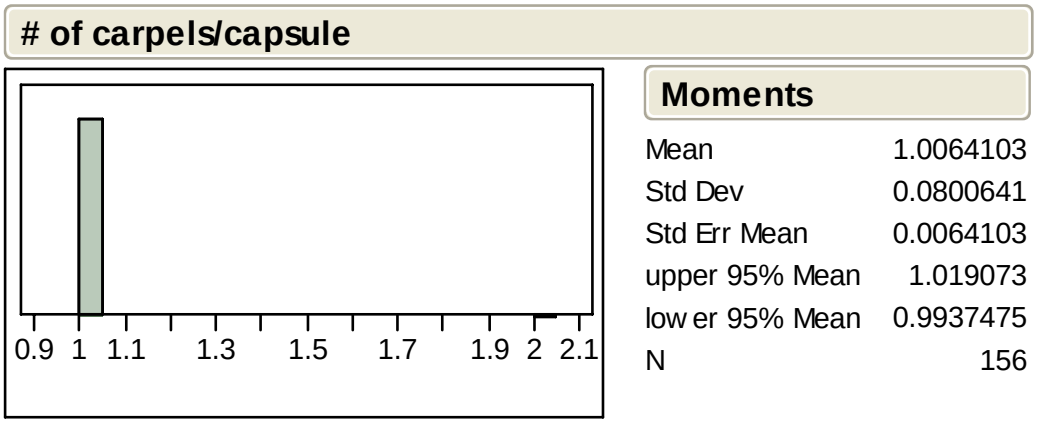

Figure 3.7. Number of carpels/capsule character histogram analysis and distribution of 156 Turkish sesame accessions.

Number of carpels/capsule character did not show variation except for only one accession.

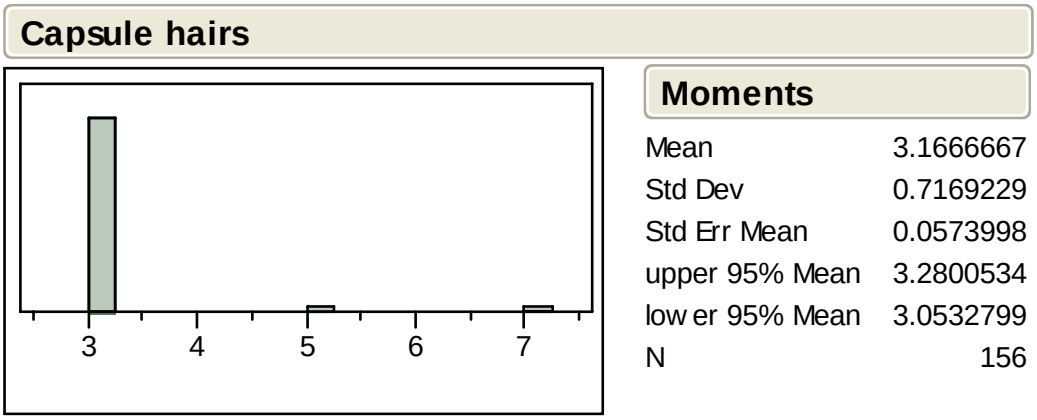

Figure 3.8. Capsule hairs character histogram analysis and distribution of 156 Turkish sesame accessions.

Variation of capsule hairs characters is very similar to stem hairs. Most of the population had score of 3. However, five accessions had a score of 5 and four accessions had scores of 7. The average capsule hairs was 3.16.

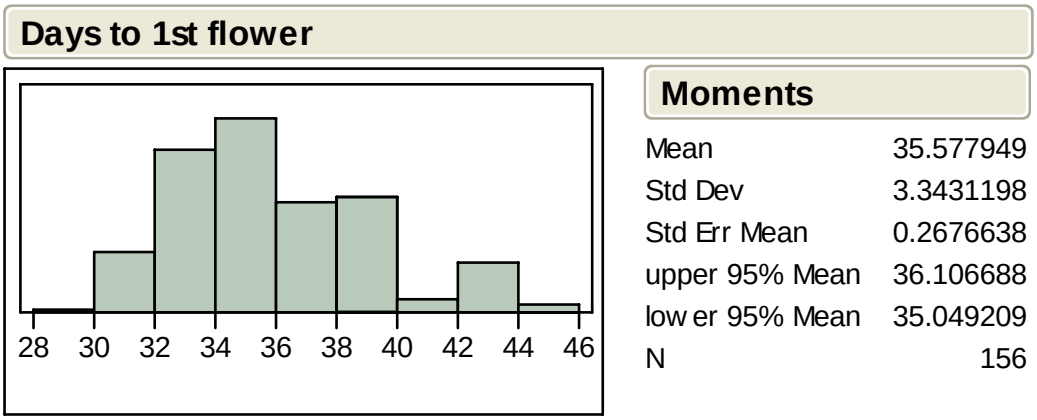

Figure 3.9. Days to 1st flower character histogram analysis and distribution of 156 Turkish sesame accessions.

Days to 1st flower character showed variation among Turkish sesame accessions. Distribution ranged from 29 to 45.5. The most common days to first flower was 35 days, followed by 33 and 39 days, respectively, and the average was 35.57 days.

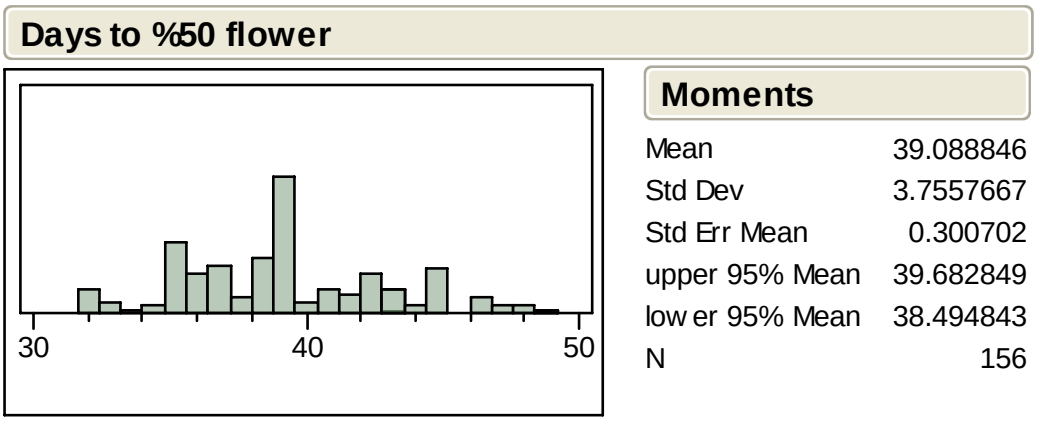

Figure 3.10. Days to %50 flower character histogram analysis and distribution of 156 Turkish sesame accessions.

Days to 50% flowering character had variation among accessions and the distribution showed a wide range. Most plants showed 50% flowering 39 days after planting. The average was 39.08 days.

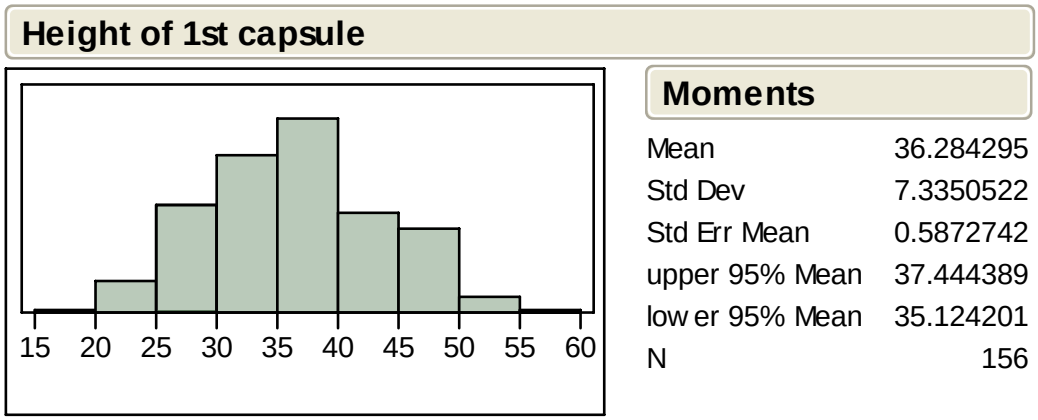

Figure 3.11. Height of 1st capsule character histogram analysis and distribution of 156 Turkish sesame accessions.

Height of 1st capsule character had high variation and showed a wide distribution. The histogram ranged from 19 to 56.5 cm. The most common range value was 36.5 cm.

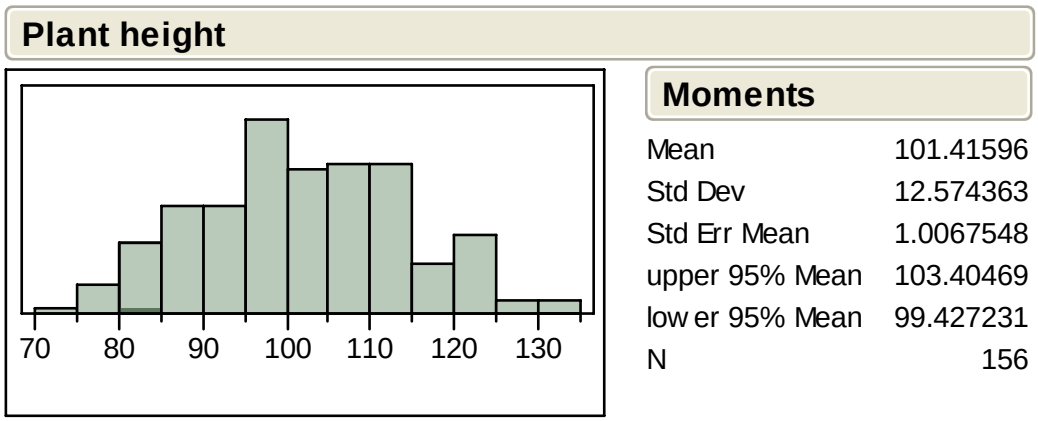

Figure 3.12. Plant height character histogram analysis and distribution of 156 Turkish sesame accessions.

Plant height character showed high variation among Turkish sesame accessions. Distribution ranged from 70.5 to 132 cm. The average height was 101.41 cm.

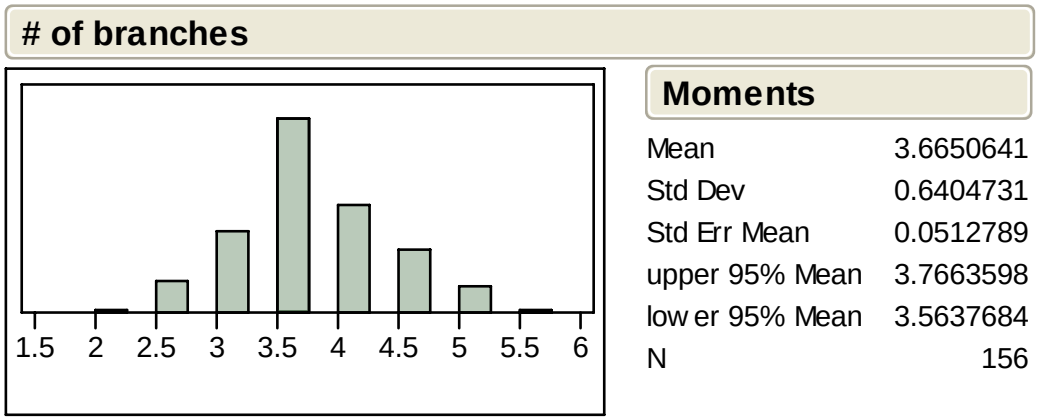

Figure 3.13. Number of branches character histogram analysis and distribution of 156 Turkish sesame accessions.

Number of branches character showed variation among Turkish sesame accessions and the histogram showed that variation had a wide spectrum. Distribution ranged from 2 to 5.5. The average was 3.66.

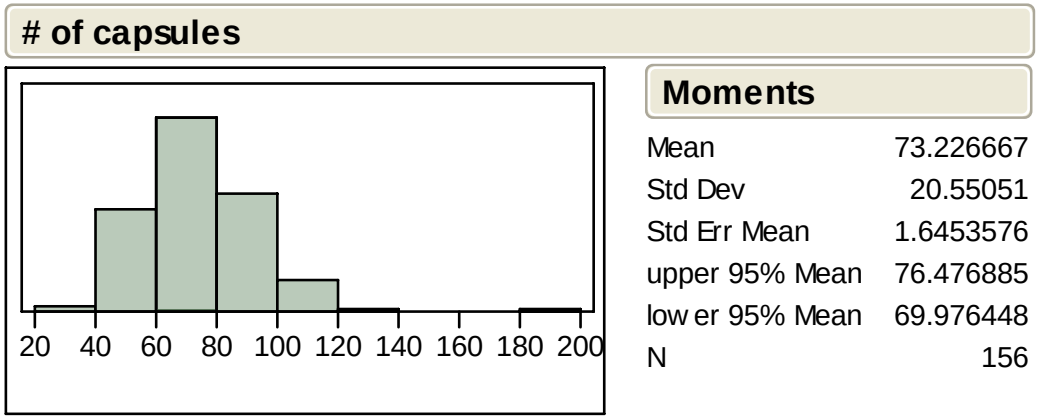

Figure 3.14. Number of capsules character histogram analysis and distribution of 156 Turkish sesame accessions.

Number of capsules character had variation among accessions and the distribution ranged from 38.5 to 197. The average for the number of capsules character was 73.22.

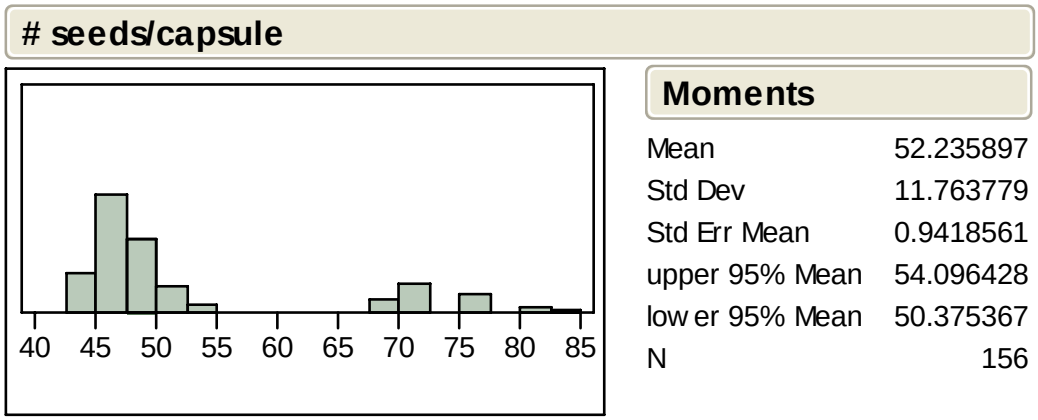

Figure 3.15. Number seeds/capsule character histogram analysis and distribution of 156 Turkish sesame accessions.

Number of seeds/capsule character showed high variation among accessions. The average of number seeds/capsule character was found to be 52.23. The distribution ranged from 42.5 to 84.

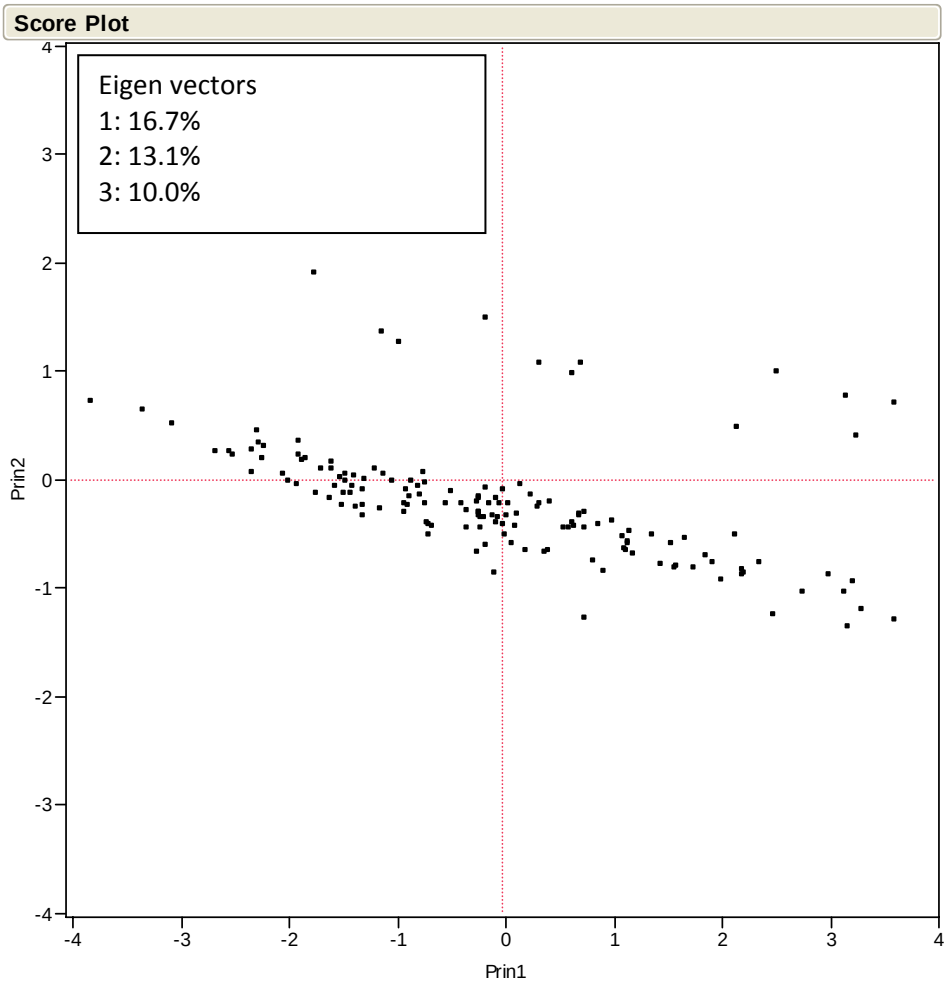

Figure 3.16. PCA for the sesame agronomic character data collected in 2009 & 2010.

Principal component analysis was performed on the combined 2009-2010 data (Figure 3.16.). According to this analysis, the first Eigen vector accounted for 16.73% of the variation among accessions and flowering time traits had the highest correlation with this vector (r = 0.52-0.53). The second Eigen vector accounted for 13.1% of variation and leaf and capsule hairs had the highest correlations with this vector (r = 0.64-0.65). The third vector accounted for 10.0% of variation with numbers of capsules and branches having the highest correlations with this vector (r = 0.55-0.62). The 2-dimensional plot showed that several accessions plotted away from the main group of sesame accessions These results suggest that these accessions are morphologically diverse from the other accessions.

3.2. SSR Marker Design Results

Using the 3662 transcript assemblies from dbEST/GenBank (ftp://ftp.tigr.org/pub/data/plantta/Sesamum_indicum), a total of 179 SSR markers were designed. When they were analysed, it was seen that the most common repeat type was trinucleotide with 111 (62.01%) identified, followed by dinucleotide 46 (25,69%), tetranucleotide 15 (8,37 %) and pentanucleotide 7 (3,91 %).

In dinucleotide microsatellites, the most common repeat type was TC/CT (39,13 %). The second most common repeat types were defined as AT/TA and AG/GA which occurred in the same range (23.91 %). TG/GT and AC/CA were found at 6.52 % and no GC/CG microsatellites were identified.

Also, in this study, 3328 transcript assemblies from dbEST/GenBank (http://www.ncbi.nlm.nih/entrez) were used to design SSR markers. They were integrated to form contigs. They were then used for *Sesamum indicum L.* as the input file on the BatchPrimer3 Primer Design Input Site (http://probes.pw.usda.gov/cgi-bin/batchprimer3/batchprimer3.cgi) and a total of 139 EST-based SSR markers were found. In these SSRs, the most common repeat type was trinucleotides (72 trinucleotides, 51.79 %) like first designed SSRs. The second most common repeat type was dinucleotides (34 dinucleotides, 24.46 %) and they were followed by tetranucleotides (22 tetranucleotides, 15,82 %), hexanucleotides (9 hexanucleotides, 6,47 %) and pentanucleotides (2 pentanucleotides, 1.43 %). Dinucleotide

microsatellites repeat distribution was seen similar to first designed SSRs. First and second designed SSRs are compared in table 3.2.

Table 3.3. Comparison of first and second designed SSRs.

1.SISSR	2.SISR
46 dinucleotide	34 dinucleotide
111 trinucleotide	72 trinucleotide
15 tetranucleotide	22 tetranucleotide
7 pentanucleotide	2 pentanucleotide
	9 hexanucleotide
total:179 microsatellites	total:139 microsatellites

3.3. Parental Testing of EST Based SSRs

A total of 179 and 139 EST-based SSRs were tested separately in Korean-Japan and African sesame accessions which will be used as parents for future mapping studies. A total of 153 SSRs designed in the first experiment and 126 SSRs designed in the second experiment were amplified in these accessions. However, polymorphism was not observed on agarose gels.

The 153 amplified SSR products of Korean-Japan and African sesame accessions were analysed in Qiaxcel capillary electrophoresis system which is able to separate DNA fragments with high resolution (2-5 bp). As a result, five polymorphic SSRs were identified in parental survey.

3.4. SSR Analysis Results

In this study, genetic diversity of 161 Turkish sesame accessions was determined by 30 SSR markers. Characteristics of SSR markers are shown in Table A-1. Five of these SSR markers were polymorphic and totally 19 polymorphic alleles were found. The amplified fragments ranged in size from approximately 145 to 1575 bp.

As a result, a phylogenetic tree of the 161 Turkish sesame accessions was constructed with NTSYS-pc version 2.2 software program using DICE matrix and UPGMA (Unweighted Pair Group Method) arithmetical averages in SHAN module. Mantel test (Mantel 1967) was applied to compare the data matrix and the tree and the correlation was 0.78. This means that the correlation between sample genotypic data and the dendrogram was acceptable. The dendrogram scale varied from 0.68 to 1.0 with a mean similarity of 0.813. This means Turkish sesame accessions have fairly low diversity. According to the dendrogram, sesame accessions clustered into 3 groups: A, B, C (Figure 3.2.) The group A, had 26 accessions. The African sesame accession was identical with two Turkish sesame accessions in A group. Also, there were nine identical sesame accessions in this group. Group B was the largest group and included 135 sesame accessions. B group included 109 ıdentical accessions. These results suggest that the Turkish sesame genebank may have many redundant accessions. However, more markers must be tested to confirm these initial results. In the C group, there was only one Turkish sesame accession and this group included the Korean sesame accession. Clustered from all groups, *Sesamum alatum L.*, which was the outgroup showed 0.626 similarity with national *Sesamum indicum L.* accessions. Also, it was determined that Turkish sesame accessions were more similar to the African sesame accession than the Korean sesame accession. *Sesamum alatum L.* was most similar to the Korean sesame accession.

In previous studies, genetic variability in sesame has also been researched by molecular techniques. Isozymes, RAPD, ISSR and AFLP have been used as molecular markers to date. Isshiki et al. (1997) used isozymes for determination of genetic variation in 68 accessions of cultivated sesame (from Japan, Korea, and Thailand) and only one enzyme showed variation. Bhat et al. (1999) studied genetic diversity of exotic sesame and Indian germplasm via RAPD markers. They found a high level of genetic diversity. Kim et al. (2002) determined genetic diversity among 75 sesame accessions of Korean and exotic sesame using fourteen ISSR markers. The accessions clustered into seven groups and showed that different geographical origins are not completely distinct. In Turkey, Ercan et al. (2004) determined genetic diversity among 38 Turkish sesame accessions using 12 RAPD primer and they found important variation. Laurentin et al. (2006) performed AFLP in 32 sesame accessions from the Venezuelan germplasm and they found 93% polymorphism. However, Ali et al. (2007) used AFLP for determining the genetic diversity of 96 sesame accessions from different parts of the

world and they determined low (35%) genetic diversity. According to these results, Korean and Indian sesame accessions have high genetic diversity. However, there is no exact relationship between genetic diversity and geographical origin. In our study, Turkish sesame accessions showed very low polymorphism. However, only 19 alleles were determined as polymorphic. Therefore, it is suggested that other marker systems can be used to confirm the low apparent diversity of Turkish sesame. For example, genomic SSR marker systems can be used as a potentially more polymorphic marker system.

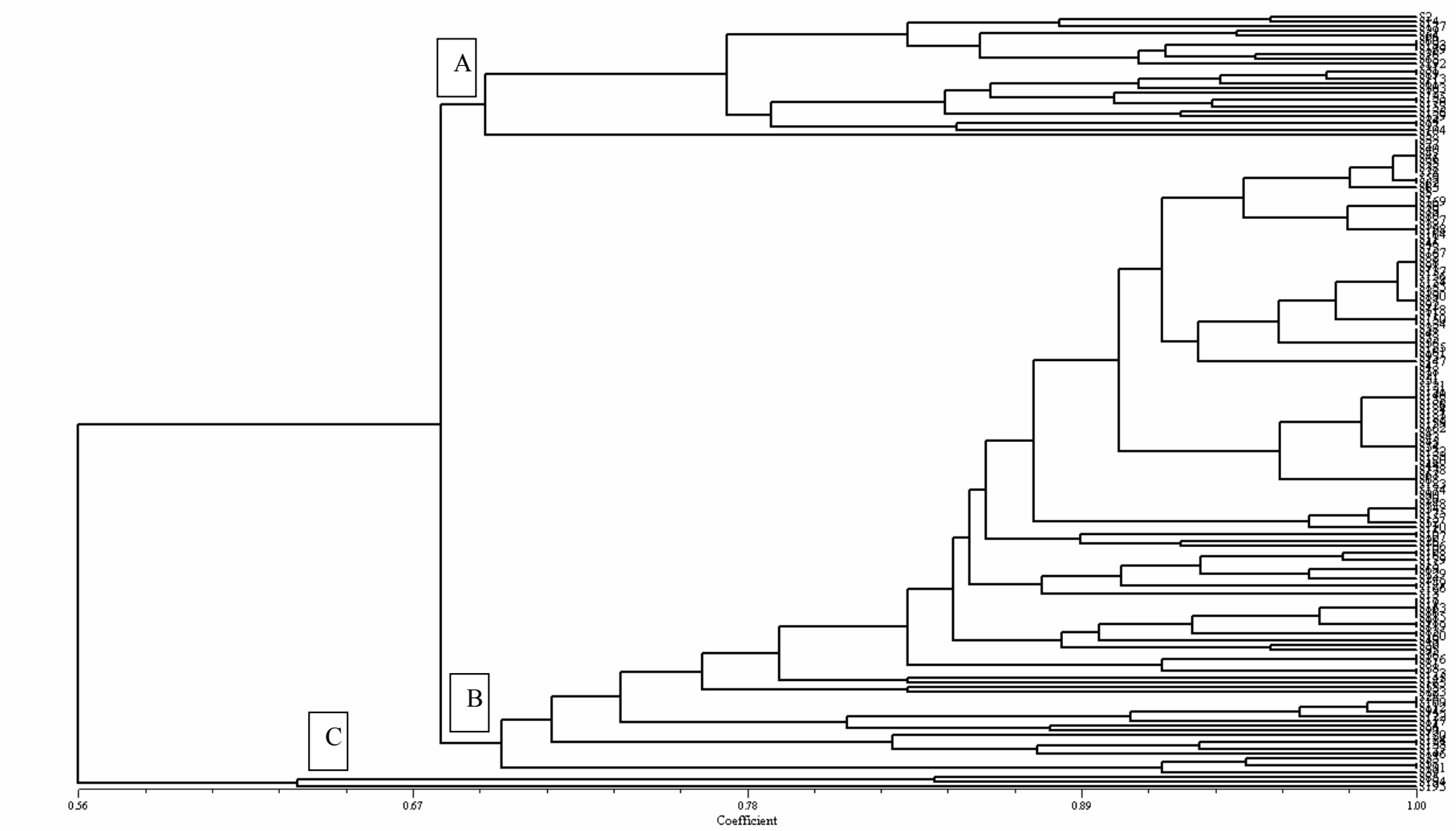

Figure 3.17. The phylogenetic tree with 161 national sesame accessions and 3 outgroups.

CHAPTER 4

CONCLUSION

Sesame (*Sesamum indicum L.*), is one of the most important oilseed crops because of its nutritional composition and health related value. Its production is dispersed all over the world and throughout Turkey. Although, early molecular studies have not confirmed it, sesame has a large genetic variability as shown morphologically. The aim of our study was characterization of the genetic diversity of Turkish sesame accessions. Accessions were characterized morphologically. In addition EST-based SSR markers were developed for molecular genetic characterization of the national sesame collection.

In this study, 156 Turkish sesame accessions were used for morphological characterization and 161 sesame accessions were used for molecular characterization. The EST based SSR markers were also screened in African and Korean-Japan sesame accessions for future map construction studies. According to morphological analysis, Height of first capsule, capsule number per plant, plant height, number of branches, number seeds/capsule, days to 1st flower and days to %50 flower characters showed variation among accessions. However, stem hair, leaf hairs, axil flower number, number of carpels/capsule and capsule hairs showed variation for only a few accessions. Also, growth habit, branching and capsule splitting characters had no variation In parental survey result, only five polymorphic markers were found among 318 EST based SSR markers. These means EST based SSR marker system is not efficient for genetic diversity and mapping studies. In accordance with genetic characterization, 19 EST based SSR fragments were analysed. Dendrogram analysis showed that Turkish sesame accessions clustered into three groups. These results may help in selection of accessions as breeding materials for new cultivars. In addition our study will be useful for managing and developing germplasm collections by eliminating redundant accessions. When the phylogenetic tree was examined, B group had lots of redundant sesame accessions. If we eliminate the redundant accessions and leave only one accession as representative, the collection can be reduced to 72 accessions. This suggests that the Turkish sesame genebank is very redundant. However, to confirm this result, more

markers must be tested and we should switch to a more polymorphic system like genomic SSRs.

REFERENCES

Ali, G.M., S. Yasumaoto and M. Seki-Katsuta. 2007. Assessment of genetic diversity in sesame (*Sesamum indicum L.*) detected by Amplified Fragment Length Polymorphism markers. *Electronic Journal of Biotechnology*. 10: 13-23.

Agarwal, M., N. Shrivastava and H. Padh. 2008. Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Rep*. 27: 617–631.

Ashri, A., G. Robbelen and R.K. Downey. eds. 1989. Oil Crops of the World. New York. McGraw-Hill.

Ashri, A., J. Janick. ed. 1998. Sesame Breeding. Plant Breeding Reviews. NJ,16: 179–228.

Bandopadhyay, R., S. Sharma, S. Rustgi, S. Singh, A. Kumar, H.S. Balyon and P. K. Gupta. 2004. DNA polymorphism among 18 species of Triticum-Aegilops complex using wheat EST-SSRs. *Plant Sci.* 166: 349–356.

Baydar, H, İ. Turgut, K. Turgut. 1999. Variation of Certain Characters and Line Selection for Yield, Oil, Oleic and Linoleic Acids in the Turkish Sesame (*Sesamum indicum L.*) Populations. *Tr. J. of Agriculture and Forestry.* 23: 431-441.

Baydar, H. 2005. Breeding for the improvement of the ideal plant type of sesame. *Plant Breeding.* 124: 263-267.

Bhat, K.V., P.P. Babrekar and S. Lakhanpaul. 1999. Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum L.*) germplasm using random amplified polymorphic DNA (RAPD) markers. *Euphytica.* 110: 21–33.

Brar, G.S. 1982. Variations and correlations in oil content and fatty acid composition of sesame. *Indian J. Agric. Sci*. 52: 434-439.

Brar, G.S. and K.L. Ahuja. 1979. Sesame: its culture, genetics, breeding and biochemistry. *Annu. Rev. of Plant Sci*. 245-313.

Bedigian, D., C.A. Smyth, and J.R. Harlan. 1986. Patterns of morphological variation in sesame. *Economic Botany.* 40: 353-365

Bedigian, D., K.F. Kiple, and C.K. Ornelas-Kiple eds. 2000. 'Sesame' The Cambridge World History of Food. New York. Cambridge University Press. 1: 411–421.

Bedigian, D. 2003. Evolution of sesame revisited: domestication, diversity and prospects. *Genetic Resources and Crop Evolution.* 50: 779 – 787.

Bisht, I.S., R.K. Mahajan, T.R. Loknathan and R.C. Agrawal. 1998. Diversity in Indian sesame collection and stratification of germplasm accessions in different diversity groups. *Genetic Resources and Crop Evolution.* 45: 325-335.

Buckler, E.S. and J.M. Thornsberry. 2002. Plant molecular diversity and applications to genomics. Current *Opinion in Plant Biology.* 5: 107–111.

Chen, P.R., K.L. Chien, T.C. Su, C.J. Chang, T.L. Liu, H. Cheng, C. Tsai. 2005. Dietary sesame reduces serum cholesterol and enhances antioxidant capacity in hypercholesterolemia. *Nutrition Research.* 25: 559–567.

Costa, F.T., S.M. Neto, C. Bloch, O.L. Franco. 2007. Susceptibility of Human Pathogenic Bacteria to Antimicrobial Peptides from Sesame Kernels. *Current Microbiology* 55: 162–166.

Cuvelier, M.E., H. Richard and C. Berset. 1992. Comparison of the antioxidative activity of some acid-phenols: Structure-activity relationship. *Bioscience Biotechnology and Biochemistry.* 56: 324–325.

Ellis, J.R. and J.M. Burke. 2007. EST-SSRs as a resource for population genetic analyses *Heredity*. 99: 125–132.

Ercan, A.G., M. Taskin and K. Turgut. 2004. Analysis of genetic diversity in Turkish sesame (*Sesamum indicum L.*) populations using RAPD markers. *Genetic Resources and Crop Evolution.* 51: 599–607.

Frank, J. 2005. Beyond vitamin E supplementation: An alternative, strategy to improve vitamin E status. *Journal of Plant Physiology.* 162: 834-843.

Fraser, L.G., C.F. Harvey, R.N. Crowhurst, and H.N. Silva. 2004. EST-derived microsatellites from Actinidia species and their potential for mapping. *Theor. Appl. Genet.* 108: 1010–1016

Fulton, T.M., J. Chungwonse, S.D. Tanksley. 1995. Microprep Protocol for Extraction of DNA from Tomato and other Herbaceous Plants. *Plant Molecular Biology Reporter.* 13: 207-209.

Hamrick, J.L., M.J. Godt., Brown H.D., M.T. Clegg, A.L. Kahler and B.S. Weir eds. 1990. Plant population genetics, breeding and genetic resources: Allozyme diversity in plant species. 43–63.

Hess, D.E., Dodo H. 2004. Potential for sesame to contribute to integrated control of Striga hermonthica in the West African Sahel. *Crop Protection.* 23: 515–522.

Isshiki, S. and T. Umezaki. 1997. Genetic variations of isozymes in cultivated sesame (*Sesamum indicum L.*). *Euphytica.* 93: 375–377.

Janick, J. ed. 2005. *Plant Breeding Reviews*, 25. New Jersey. John Wiley.

Johnson, L.A., T.M. Suleiman and E.W. Lusas. 1979. Sesame protein: A review and prospectus. J. *Amer. Oil Chem. Soc*. 56: 463-468.

Kantety, R.V., M.L. Rota., D.E. Matthews and M.E. Sorrells. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.* 48: 501–510.

Kaur, I.P., A. Saini. 2000. Sesamol exhibits antimutagenic activity against oxygen species mediated mutagenicity. *Mutation Research.* 470: 71–76.

Kapadia, G.J., M.A. Azuine, H. Tokuda, M. Takasaki, T. Mukainaka , T. Konoshima and H. Nishino. 2002. Chemopreventive effect of resveratrol, sesamol, sesame oil and sunflower oil in the epstein–barr virus early antigen activation assay and the mouse skin two-stage carcinogenesis. *Pharmacological Research.* 45: 500-505.

Kim, D.H., G. Zur, Y. Danin-Poleg, S.W. Lee, K.B. Shim, C.W. Kang and Y. Kashi. 2002. Genetic relationships of sesame germplasm collection as revealed by inter-simple sequence repeats. *Plant Breeding.* 121: 259-262.

Laurentin, H. 2007. Genetic diversity in sesame (*Sesamum indicum L.*): molecular markers, metabolic profiles and effect of plant extracts on soil-borne pathogenic fungi. Dissertation: http://webdoc.sub.gwdg.de.

Laurentin, H., P. Karlovsky. 2006. Genetic relationship and diversity in a sesame (*Sesamum indicum L.*) germplasm collection using amplified fragment length polymorphism (AFLP). *BMC Genetics.* 7/10: 1-13.

Li, G. and C.F. Quiros. 2001. Sequence- Related Amplified Polymorphism (SRAP), A New Marker System based on A Simple PCR Reaction. *Theor Appl Genet*. 103: 455-461.

Mabberley, D.J. 1997. The Plant-Book. A Portable Dictionary of the Higher Plants. 2nd edn. Cambridge University Pres. Cambridge.

Meudt, H.M. and Andrew C.C. 2007. Almost Forgotten or Latest Practice AFLP applications, analyses and Advances. *TRENDS in Plant Science.* 12: 107-117

Morgante, M. and Olivieri A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal.* 3: 175-182.

Pashley, C.H., J.R. Ellis, D.E. McCauley, and J.M. Burke. 2006. EST databases as a source for molecular markers: lessons from Helianthus. *J Hered.* 97: 381–388.

Pathirana, R., 1994. Natural cross-pollination in sesame (*Sesamum indicum L.*). *Plant Breeding.* 112: 167-170.

Peter, K.V., ed. 2004. Woodhead Publishing in Food Science and Technology: *Handbook of spices*. 2. Cambridge.

Shahidi, F., C.M. Liyana-Pathirana., D.S. Wall. 2006. Antioxidant activity of white and black sesame seeds and their hull fractions. *Food Chemistry.* 99: 478–483.

Sharmila, V., S.K. Ganesh, M. Gunasekaran. 2007. Generation mean analysis for quantitative traits in sesame (*Sesamum indicum L.*) crosses. *Genetics and Molecular Biology*. 30: 80-84.

Tanksley, S.D., and S.R. McCouch. 1997. Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. *Science.* 277

Uzun, B., D. Lee, P. Doniniand, M. Çağırgan. 2003. Identification of a molecular marker linked to the closed capsule mutant trait in sesame using AFLP. *Plant Breeding.* 122: 95-97.

Uzun, B., M.İ. Çağıran. 2006. Comparison of determinate and indeterminate lines of sesame for agronomic traits. *Field Crops Research.* 96: 13-18.

Uzun, B., S. Ülger and M.İ. Çağıran. 2002. Comparison of Determinate and Indeterminate Types of Sesame for Oil Content and Fatty Acid Composition. *Turk J Agric For*. 26: 269-274.

Xiurong, Z., Z. Yingzhong1, C. Yong, F. Xiangyun, G. Qingyuan1, Z. Mingdeand T. Hodgkin. 2000. Establishment of sesame germplasm core collection in China. *Genetic Resources and Crop Evolution.* 47: 273–279.

Varshney, R.K., A. Graner, and M.E. Sorrells. 2005. Genic microsatellite markers in plants: features and applications. *TRENDS in Biotechnology*. 23: 49-55.

Yermanos, D.M., S. Hemstreet, W. Saeeb, and C.K. Huszar. 1972. Oil content and composition of the seed in the world collection of sesame introductions. *J. Amer. Oil Chem. Soc.* 49: 20-23.

APPENDIX A

ALL SSR PRIMERS FROM SESAME DBEST/GENBANK

Table. A-1. EST-based SSR primers from TIGR dbEST/GenBank with their repeat type, repeat ,Tm and product size.

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>AB194714	trinucleotide	GTGGTGGTGGTG	54.614	57.971	267
>BU667375	tetranucleotide	TCCCTCCCTCCCT	63.037	54.949	180
>BU667382	dinucleotide	CTCTCTCTCTCTCTCT	58.670	54.362	242
>BU667391	dinucleotide	ATATATATATATATATATAT	58.701	55.191	169
>BU667423	trinucleotide	GAGGAGGAGGAG	54.923	54.878	283
>BU667455	trinucleotide	CAGCAGCAGCAGCAGC	54.974	55.017	236
>BU667464	dinucleotide	TCTCTCTCTCTC	56.089	55.189	370
>BU667477	trinucleotide	CCACCACCACCAC	54.937	55.906	101
>BU667552	dinucleotide	TCTCTCTCTCTC	57.194	55.276	147
>BU667564	trinucleotide	CTGCTGCTGCTGCT	55.514	55.059	220
>BU667567	trinucleotide	AAGAAGAAGAAG	63.248	54.320	167
>BU667791	trinucleotide	TCTTCTTCTTCTTC	56.178	54.923	107

(cont. on next page)

Table. A-1. (cont.)

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>BU667806	dinucleotide	GAGAGAGAGAGAGACAGAGAGAGAGAGA	54.601	55.252	268
>BU667825	trinucleotide	TTCTTCTTCTTCT	58.727	59.304	200
>BU667875	trinucleotide	GGAGGAGGAGGAGG	54.736	55.446	328
>BU667918	trinucleotide	GATGATGATGATG	56.460	54.810	326
>BU668080	dinucleotide	GAGAGAGAGAGAGTGAGTGAGAGAGAGAG	59.135	55.316	170
>BU668121	dinucleotide	ATATATATATATATATATAT	58.544	55.301	185
>BU668131	trinucleotide	TTCTTCTTCTTCTTC	57.004	55.240	392
>BU668208	dinucleotide	AGAGAGAGAGAGAGAGAGAG	55.072	56.379	132
>BU668263	trinucleotide	GCTGCTGCTGCT	55.630	54.689	187
>BU668301	trinucleotide	TGGTGGTGGTGGTGG	55.024	55.773	172
>BU668365	dinucleotide	AGAGAGAGAGAGAGAGAGAGAG	53.744	55.327	210
>BU668378	dinucleotide	AGAGAGAGAGAGA	56.022	55.408	174
>BU668385	trinucleotide	GATGATGATGATGATGATGATG	55.055	54.888	222
>BU668438	dinucleotide	TCTCTCTCTCTCTCTCT	54.580	54.937	226
>BU668444	dinucleotide	TGTGTGTGTGTGTGT	59.673	55.319	306
>BU668467	trinucleotide	CCACCACCACCACCACCATCACCACC	56.460	54.783	152
>BU668583	dinucleotide	AGAGAGAGAGA	56.225	55.038	336
>BU668658	trinucleotide	ATTATTATTATT	57.331	58.715	368
>BU668708	trinucleotide	ACCACCACCACC	57.345	54.935	251
>BU668767	trinucleotide	CCGCCGCCGCCG	55.007	55.174	188
>BU668842	trinucleotide	GTAGTAGTAGTAGT	62.560	55.744	320
>BU668875	trinucleotide	CGGCGGCGGCGGC	57.421	54.789	375

(cont. on next page)

Table. A-1. (cont.)

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>BU669001	dinucleotide	TCTCTCTCTCTCTCTCT	56.211	59.488	357
>BU669013	trinucleotide	CCTCCTCCTCTTCCTCCTC	57.854	55.571	294
>BU669034	tetranucleotide	TTCTTTCTTTCTTTCTT	58.900	56.278	250
>BU669056	trinucleotide	GAGGAGGAGGAG	54.386	57.012	140
>BU669095	trinucleotide	GACGACGACGAC	54.024	55.897	133
>BU669103	trinucleotide	CGGCGGCGGCGGCGGCGGCGG	62.842	56.536	107
>BU669113	trinucleotide	GAAGAAGAAGAAG	58.877	62.363	122
>BU669124	dinucleotide	GAGAGAGAGAGA	55.107	59.139	396
>BU669143	trinucleotide	GCGGCGGCGGCGGC	61.787	55.024	339
>BU669158	trinucleotide	GGCGGCGGCGGC	55.073	58.597	368
>BU669217	trinucleotide	GAAGAAGAAGAAG	56.759	54.949	164
>BU669409	tetranucleotide	GTATGTATGTATGTATGTATGTAT	54.751	63.249	361
>BU669482	dinucleotide	TCTCTCTCTCT	56.241	57.264	259
>BU669623	pentanucleotide	ATTAGATTAGATTCGATTAGAT	55.029	55.672	179
>BU669641	dinucleotide	AGAGAGAGGGAGAGAGAGA	55.204	55.648	225
>BU669689	trinucleotide	TGGTGGTGGTGGCGGTGGTGGTG	54.660	54.797	157
>BU669703	trinucleotide	TCCTCCTCCTCCTCCTCCT	56.160	55.784	267
>BU669739	trinucleotide	CTGCTGCTGCTG	59.830	55.103	226
>BU669782	trinucleotide	CCACCACCACCACCAC	54.784	55.052	343
>BU669912	trinucleotide	ACAACAACAACAA	54.002	55.017	267
>BU669941	trinucleotide	TCTTCTTCTTCTTC	56.178	54.923	107

(cont. on next page)

Table. A-1. (cont.)

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>BU669957	dinucleotide	AGAGAGAGAGAGAGAGA	56.884	56.118	400
>BU670021	trinucleotide	CCGCCGCCGCCG	55.156	55.192	341
>BU670139	tetranucleotide	CAGGCAGGCAGGCAGGCAGGC	57.222	57.328	349
>BU670140	trinucleotide	GCAGCAGCAGCA	54.735	54.840	170
>BU670197	trinucleotide	TGATGATGATGA	54.789	54.441	211
>BU670208	trinucleotide	AGAAGAAGAAGAAG	56.280	54.572	148
>BU670211	trinucleotide	TCCTCCTCCTCTTCCTCCTC	55.140	56.705	295
>BU670257	tetranucleotide	CAAGCAAGCAAGC	55.041	55.715	347
>BU670263	trinucleotide	GATGATGATGATGAT	55.711	55.526	176
>BU670264	dinucleotide	TCTCTCTCTCCTCTCTCTCCTCTCTCTC	60.397	55.003	383
>BU670270	trinucleotide	GAGGAGGAGGAG	54.923	55.032	275
>BU670310	dinucleotide	ACACACACACACACACACACACAC	56.834	55.006	159
>BU670320	trinucleotide	AAGAAGAAGAAG	55.024	54.586	205
>BU670327	trinucleotide	AGAAGAAGAAGAAGAAGGAGAAGAAGAAG	58.071	55.489	192
>BU670450	dinucleotide	TATATATATATAT	55.950	58.076	293
>BU670499	trinucleotide	TATTATTATTAT	55.067	54.937	121
>BU670541	trinucleotide	AGAAGAAGAAGAAGAAG	58.071	54.943	146
>BU670593	trinucleotide	GTTGTTGTTGTT	55.539	55.539	336
>BU670618	trinucleotide	GGAGGAGGAGGA	55.240	54.751	346
>BU670656	dinucleotide	TCTCTCTCTCTCT	56.225	57.360	351
>BU670662	trinucleotide	AGCAGCAGCAGCAGCA	55.626	55.227	104
>BU670685	dinucleotide	TCTCTCTCTCTCTCTCTC	59.711	60.034	195

(cont. on next page)

Table. A-1. (cont.)

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>SIU25817	trinucleotide	GAAGAAGAAGAAGAA	55.174	54.818	336
>TA1012_4182	trinucleotide	CCACCACCACTACCACCACCACC	57.478	56.318	150
>TA1021_4182	dinucleotide	ACACACACACA	58.646	55.175	159
>TA1031_4182	tetranucleotide	CCTGCCTGCCTGCCTG	56.777	55.127	186
>TA1044_4182	dinucleotide	TCTCTCTCTCTCTCTCTCTC	57.739	54.923	331
>TA1045_4182	trinucleotide	GCAGCAGCAGCAGC	56.976	55.423	322
>TA1051_4182	tetranucleotide	AAAGAAAGAAAGA	60.901	56.152	138
>TA106_4182	trinucleotide	TGGTGGTGGTGG	54.751	54.817	172
>TA1075_4182	tetranucleotide	CTTTCTTTCTTTCT	58.850	55.013	257
>TA1078_4182	trinucleotide	TCTTCTTCTTCT	57.131	55.013	176
>TA110_4182	dinucleotide	CTCTCTCTCTCT	54.166	54.308	249
>TA1123_4182	trinucleotide	AAGAAGAAGAAG	54.783	55.289	286
>TA113_4182	tetranucleotide	TCTGTCTGTCTGTCT	55.024	55.327	247
>TA1136_4182	tetranucleotide	GATGGATGGATGG	55.655	55.443	391
>TA1137_4182	trinucleotide	GCAGCAGCAGCAGCAG	54.922	58.389	115
>TA1138_4182	trinucleotide	CCACCACCACCACC	55.194	59.776	261
>TA1151_4182	trinucleotide	CTGCTGCTGCTGC	55.789	53.785	293
>TA1155_4182	dinucleotide	ATATATATATATA	54.630	60.000	222
>TA1158_4182	dinucleotide	TCTCTCTCTCTCCTCTCTCTCTCTCT	54.779	57.212	154
>TA13_4182	trinucleotide	GGTGGTGGTGGTG	55.214	54.495	146
>TA142_4182	trinucleotide	TGGTGGTGGTGG	55.003	54.668	157
>TA156_4182	trinucleotide	GCGGCGGCGGCGG	56.225	55.582	201

(cont. on next page)

Table. A-1. (cont.)

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>TA16_4182	trinucleotide	CACCACCACCACTACCACCA	57.883	55.145	240
>TA17_4182	trinucleotide	CACCACCACCACTACCACCA	57.883	55.145	240
>TA181_4182	trinucleotide	CTACTACTACTA	54.280	54.126	158
>TA19_4182	trinucleotide	TATTATTATTAT	54.386	54.182	177
>TA191_4182	pentanucleotide	ACACCACACCACACCACACCAC	56.004	56.930	303
>TA201_4182	trinucleotide	AGCAGCAGCAGCAGCA	55.626	55.227	104
>TA221_4182	dinucleotide	AGAGAGAGAGAGAGAGAGAG	55.072	56.379	132
>TA227_4182	trinucleotide	CCGCCGCCGCCGC	58.228	54.937	257
>TA239_4182	dinucleotide	TCTCTCTCTCTCTCT	55.103	55.240	257
>TA247_4182	dinucleotide	TATATATATATATAT	55.401	56.013	293
>TA259_4182	trinucleotide	GATGATGATGATGAT	54.538	58.528	341
>TA268_4182	trinucleotide	GGTGGTGGTGGTGG	55.103	57.131	163
>TA271_4182	trinucleotide	CCACCACCACCACCA	54.994	55.024	370
>TA28_4182	trinucleotide	AGAAGAAGAAGAAGAAGGAGAAGAAGAAG	58.071	54.656	192
>TA286_4182	trinucleotide	AAGAAGAAGAAG	54.783	55.289	286
>TA29_4182	trinucleotide	GCAGCAGCAGCAGC	54.982	55.024	129
>TA327_4182	dinucleotide	CTCTCTCTCTCTCTCT	56.453	55.353	204
>TA342_4182	pentanucleotide	AGAGAAGAGAAGAG	61.200	54.937	323
>TA349_4182	trinucleotide	GCAGCAGCAGCAGC	56.575	58.350	285
>TA35_4182	dinucleotide	TGTGTGTGTGTGTG	55.189	55.205	349
>TA350_4182	dinucleotide	ATATATATATATA	54.630	60.000	222
>TA353_4182	dinucleotide	TCTCTCTCTCTCCTCTCTCTCTCTCT	54.779	57.212	154

(cont. on next page)

Table. A-1. (cont.)

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>TA381_4182	trinucleotide	GATGATGATGAT	55.007	53.977	378
>TA388_4182	dinucleotide	TATATATATATATATATATATATATA	55.369	55.140	346
>TA420_4182	trinucleotide	GCAGCAGCAGCAGCAG	54.922	58.389	115
>TA432_4182	trinucleotide	CCACCACCACTACCACCACCACC	57.478	56.318	150
>TA437_4182	tetranucleotide	CCTGCCTGCCTGCCTG	56.777	55.127	186
>TA439_4182	tetranucleotide	AAAGAAAGAAAGA	60.901	56.152	138
>TA46_4182	pentanucleotide	TTTCTTTTCTTTTCTT	55.327	54.024	189
>TA495_4182	trinucleotide	TCTTCTTCTTCTTCTTCTTC	58.147	55.073	192
>TA496_4182	tetranucleotide	CTTTCTTTCTTTCT	58.850	55.013	257
>TA514_4182	trinucleotide	CTGCTGCTGCTGC	55.789	53.785	293
>TA517_4182	trinucleotide	CCACCACCACCACC	55.194	59.776	261
>TA544_4182	trinucleotide	GATGATGATGATGATGAAGATGATGA	56.799	54.840	209
>TA546_4182	dinucleotide	ACACACACACA	58.646	55.175	159
>TA547_4182	trinucleotide	TCTTCTTCTTCT	57.131	55.013	176
>TA567_4182	tetranucleotide	GATGGATGGATGG	55.655	55.443	391
>TA580_4182	trinucleotide	GCAGCAGCAGCAGC	56.976	55.423	322
>TA582_4182	trinucleotide	CCACCACCACCA	54.765	54.923	291
>TA593_4182	trinucleotide	GCAGCAGCAGCAG	55.707	56.460	326
>TA597_4182	trinucleotide	GGTGGTGGTGGTG	55.214	54.495	146
>TA599_4182	trinucleotide	GTGGTGGTGGTG	55.711	52.859	325
>TA600_4182	trinucleotide	GTGGTGGTGGTG	55.145	57.883	240
>TA603_4182	trinucleotide	TATTATTATTAT	54.386	54.182	177

(cont. on next page)

Table. A-1. (cont.)

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>TA613_4182	trinucleotide	AGAAGAAGAAGAAGAAGGAGAAGAAGAAG	58.071	54.656	192
>TA614_4182	trinucleotide	GCAGCAGCAGCAGC	54.982	55.024	129
>TA617_4182	dinucleotide	TGTGTGTGTGTGTG	55.189	55.205	349
>TA634_4182	pentanucleotide	TTTCTTTTCTTTTCTT	55.327	54.024	189
>TA657_4182	trinucleotide	GCGGCGGCGGCGGC	58.019	57.424	261
>TA66_4182	trinucleotide	GCGGCGGCGGCGGC	58.019	57.424	261
>TA660_4182	trinucleotide	CAGCAGCAGCAGC	57.452	54.479	154
>TA666_4182	trinucleotide	CCGCCGCCGCCGCC	54.045	54.937	342
>TA669_4182	dinucleotide	TATATATATATA	54.974	58.010	255
>TA674_4182	dinucleotide	CTCTCTCTCTCT	54.166	54.308	249
>TA676_4182	trinucleotide	TGGTGGTGGTGG	54.751	54.817	172
>TA694_4182	tetranucleotide	TCTGTCTGTCTGTCT	55.024	55.327	247
>TA7_4182	trinucleotide	GCAGCAGCAGCAG	55.707	56.460	326
>TA714_4182	trinucleotide	TGGTGGTGGTGG	55.003	54.668	157
>TA767_4182	trinucleotide	TAGTAGTAGTAG	54.126	54.280	158
>TA777_4182	trinucleotide	AGCAGCAGCAGCAGCA	55.626	55.227	104
>TA791_4182	trinucleotide	CCGCCGCCGCCGC	58.228	54.937	257
>TA80_4182	dinucleotide	TATATATATATA	54.974	58.010	255
>TA807_4182	trinucleotide	TGGTGGTGGTGGTGG	55.024	54.994	370
>TA810_4182	dinucleotide	AGAGAGAGAGAGAGAGAGAG	55.072	56.379	132
>TA816_4182	trinucleotide	GCGGCGGCGGCGG	56.225	55.582	201

(cont. on next page)

Table. A-1. (cont.)

Sequence ID	Repeat Type	Repeat	Left Tm	Right Tm	Product size
>TA823_4182	pentanucleotide	ACACCACACCACACCACACCAC	56.004	56.930	303
>TA825_4182	dinucleotide	TCTCTCTCTCTCTCT	55.103	55.240	257
>TA827_4182	dinucleotide	TATATATATATATAT	55.401	56.013	293
>TA834_4182	trinucleotide	GGTGGTGGTGGTGG	55.103	57.131	163
>TA835_4182	trinucleotide	GATGATGATGATGAT	54.538	58.528	341
>TA87_4182	trinucleotide	CAGCAGCAGCAGC	57.452	54.479	154
>TA876_4182	trinucleotide	GATGATGATGAT	55.007	53.977	378
>TA88_4182	trinucleotide	CCGCCGCCGCCGCC	54.045	54.937	342
>TA890_4182	dinucleotide	TATATATATATATATATATATATATA	55.369	55.140	346
>TA907_4182	trinucleotide	TCTTCTTCTTCTTCTTCTTC	58.147	55.073	192
>TA934_4182	trinucleotide	GCAGCAGCAGCAGC	56.575	58.350	285
>TA973_4182	pentanucleotide	AGAGAAGAGAAGAG	61.200	54.937	323
>TA995_4182	dinucleotide	CTCTCTCTCTCTCTCT	56.453	55.353	204
>TA996_4182	trinucleotide	GATGATGATGATGATGAAGATGATGA	56.799	54.840	209
>U25817	trinucleotide	GAAGAAGAAGAAGAA	55.174	54.818	336

Table A-2. EST-based SSR primers from NCBI dbEST/GenBank with their repeat type, repeat number, length, Tm and product size

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
Contig17	Pentanucleotide	(TTTCT)3	15	60.01	305
Contig28	Hexanucleotide	(TCCTCT)3	18	60.63	313
Contig30	Trinucleotide	(GAA)5	15	59.92	322
Contig36	Trinucleotide	(CAG)4	12	59.4	302
Contig42	Trinucleotide	(CGT)4	12	60.5	261
Contig46	Trinucleotide	(TGG)4	12	59.87	305
Contig57	Trinucleotide	(CTG)4	12	59.01	315
Contig65	Hexanucleotide	(ATGTAT)2	18	59.54	257
Contig75	Trinucleotide	(TGG)5	15	60.21	300
Contig81	Dinucleotide	(CT)8	16	60.0	303
Contig117	Pentanucleotide	(ACACC)4	20	59.59	305
Contig128	Tetranucleotide	(GAAA)3	12	60.94	298
Contig160	Dinucleotide	(TA)6	12	60.25	314
Contig163	Tetranucleotide	(TTTC)3	12	60.32	315

(cont. on next page)

Table A-2. (cont.)

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
Contig207	Trinucleotide	(AGC)5	15	60.11	300
Contig210	Trinucleotide	(GAT)5	15	60.07	316
Contig213	Dinucleotide	(AT)7	14	60.0	339
Contig218	Dinucleotide	(TA)13	26	60.04	311
Contig237	Tetranucleotide	(GATG)3	12	60.02	304
Contig258	Dinucleotide	(TC)6	12	60.11	289
Contig264	Trinucleotide	(CAA)4	12	60.23	292
Contig270	Dinucleotide	(CT)8	16	60.05	301
Contig289	Dinucleotide	(GA)6	12	60.0	209
Contig298	Tetranucleotide	(TTCT)4	16	60.5	259
Contig331	Hexanucleotide	(GCACCT)4	24	60.57	303
Contig339	Hexanucleotide	(AATGCT)2	18	60.02	272
Contig345	Trinucleotide	(GGT)4	12	59.65	295
Contig381	Dinucleotide	(AG)10	20	60.28	303
Contig406	Dinucleotide	(AT)6	12	59.5	257
Contig410	Trinucleotide	(TAG)4	12	60.00	298

(cont. on next page)

Table A-2. (cont.)

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
Contig424	Trinucleotide	(CCT)4	12	59.98	296
Contig426	Trinucleotide	(GAT)5	15	60.13	287
Contig452	Dinucleotide	(AG)6	12	59.80	332
Contig462	Trinucleotide	(GTG)4	12	59.63	304
Contig477	Tetranucleotide	(TCCC)3	12	60.3	298
Contig485	Hexanucleotide	(GATTTT)2	18	60.00	293
Contig494	Trinucleotide	(GCA)5	15	60.0	298
Contig524	Trinucleotide	(CGG)7	21	60.00	263
Contig532	Dinucleotide	(TG)7	14	59.3	282
Contig558	Trinucleotide	(GAT)4	12	60.0	342
Contig576	Hexanucleotide	(GCAAGC)2	18	59.5	181
Contig585	Dinucleotide	(TA)6	12	59.00	298
Contig586	Dinucleotide	(TC)6	12	60.00	292
Contig606	Trinucleotide	(CCA)5	15	60.00	297
Contig617	Trinucleotide	(CCA)4	12	60.34	294
Contig622	Trinucleotide	(CCA)4	12	59.81	319

(cont. on next page)

Table A-2. (cont.)

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
Contig633	Trinucleotide	(GCA)4	12	59.52	299
Contig642	Tetranucleotide	(TATG)3	12	60.05	287
Contig659	Trinucleotide	(GAA)4	12	59.65	304
Contig681	Trinucleotide	(TGA)4	12	60.50	291
Contig685	Tetranucleotide	(TGGT)3	12	60.20	302
Contig692	Tetranucleotide	(CCTG)4	16	60.54	341
Contig693	Hexanucleotide	(AAGAAC)3	18	60.02	203
Contig695	Dinucleotide	(TC)6	12	60.12	303
Contig717	Trinucleotide	(GCG)4	12	60.05	321
Contig732	Trinucleotide	(AGA)4	12	59.60	235
Contig741	Dinucleotide	(GA)8	16	60.35	297
Contig760	Trinucleotide	(GGA)4	12	60.82	304
Contig764	Dinucleotide	(AT)10	20	60.25	317
Contig776	Trinucleotide	(GTA)4	12	60.15	313
Contig786	Trinucleotide	(GCA)4	12	60.01	300
Contig792	Trinucleotide	(CGG)4	12	59.02	294

(cont. on next page)

Table A-2. (cont.)

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
Contig853	Dinucleotide	(TA)7	14	59.76	298
Contig866	Trinucleotide	(GAT)4	12	59.60	308
Contig869	Dinucleotide	(AC)12	24	58.62	300
Contig881	Trinucleotide	(CCG)4	12	58.61	281
Contig886	Trinucleotide	(GAT)7	21	59,89	294
Contig907	Dinucleotide	(TC)6	12	60.04	292
Contig921	Tetranucleotide	(TCAA)3	12	60.20	367
Contig925	Tetranucleotide	(AAAG)3	12	60.53	288
Contig937	Trinucleotide	(CCG)4	12	59.20	312
Contig949	Dinucleotide	(TC)7	14	60.70	299
Contig971	Trinucleotide	(GCG)4	12	59.55	299
Contig972	Trinucleotide	(CCG)4	12	60.72	298
Contig975	Trinucleotide	(AAG)4	12	60.15	298
Contig988	Trinucleotide	(TCT)6	18	59.62	343
Contig996	Trinucleotide	(AGA)5	15	60.54	296
Contig1011	Dinucleotide	(TC)10	20	60.16	323

(cont. on next page)

Table A-2. (cont.)

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
Contig1025	Trinucleotide	(TGG)4	12	60.12	301
Contig1028	Dinucleotide	(TC)6	12	60.01	330
Contig1029	Trinucleotide	(ATT)4	12	59.43	364
Contig1033	Dinucleotide	(TG)7	14	60,17	276
Contig1044	Trinucleotide	(AAG)4	12	59,40	173
Contig1050	Dinucleotide	(AG)6	12	60.12	250
Contig1059	Trinucleotide	(GGT)4	12	60,04	306
Contig1080	Trinucleotide	(GCA)4	12	60,48	305
Contig1083	Dinucleotide	(CT)6	12	59.65	330
Contig1087	Dinucleotide	(TC)9	18	61,08	299
Contig1093	Trinucleotide	(CTG)4	12	59.03	266
Contig1100	Trinucleotide	(CGC)4	12	60.2	258
Contig1112	Hexanucleotide	(GCCACC)3	18	59,71	279
Contig1116	Tetranucleotide	(TCAA)3	12	60.34	285
Contig1154	Trinucleotide	(GCA)4	12	60.17	285
Contig1166	Trinucleotide	(GAG)4	12	60.23	286

(cont. on next page)

Table A-2. (cont.)

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
Contig1185	Trinucleotide	(GAC)4	12	59.93	315
Contig1192	Trinucleotide	(TCC)6	18	60.41	289
Contig1216	Tetranucleotide	(TTGC)3	12	60,28	293
Contig1228	Hexanucleotide	(CATTCA)3	18	59.91	107
Contig1231	Trinucleotide	(CTG)4	12	59,41	306
Contig1232	Trinucleotide	(AAG)4	12	61.02	274
Contig1253	Trinucleotide	(GCG)4	12	60,61	297
Contig1258	Trinucleotide	(CCG)4	12	60.27	298
Contig1281	Trinucleotide	(GAT)5	15	59.94	312
Contig1285	Trinucleotide	(GGC)4	12	59.98	314
Contig1289	Tetranucleoitde	(CTTT)3	12	60,7	296
Contig1292	Tetranucleotide	(CAAG)3	12	60,17	301
Contig1308	Trinucleotide	(TTC)5	15	59,75	309
Contig1309	Dinucleotide	(TG)7	14	59,83	307
Contig1322	Tetranucleotide	(TTGA)3	12	59,65	283
Contig1326	Tetranucleotide	(ACCC)3	12	59.58	268

(cont. on next page)

Table A-2. (cont.)

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
Contig1331	Trinucleotide	(TGG)4	12	60.61	299
Contig1335	Tetranucleotide	(GAAA)3	12	60.05	285
Contig1342	Trinucleotide	(GAG)4	12	60.26	306
Contig1349	Tetranucleotide	(AAGG)3	12	59.54	291
Contig1351	Dinucleotide	(TC)6	12	59.67	313
Contig1353	Trinucleotide	(CAG)5	15	59.80	316
Contig1359	Trinucleotide	(CCA)6	18	59,96	252
Contig1377	Trinucleotide	(AAC)4	12	59,27	300
Contig1383	Trinucleotide	(TCT)4	12	60,11	298
Contig1385	Trinucleotide	(TAT)4	12	59,47	301
Contig1394	Trinucleotide	(ACA)4	12	60.35	284
Contig1418	Trinucleotide	(TCT)4	12	59.92	295
Contig1436	Trinucleotide	(AGC)4	12	59.81	213
BU670139	Tetranucleotide	(CAGG)5	20	60,05	305
BU670140	Trinucleotide	(GCA)4	12	59.86	252
BU670685	Dinucleotide	(TC)6	18	60,62	267

(cont. on next page)

Table A-2. (cont.)

Seq ID	Repeat Type	Repeat Number	SSR Length	Tm	Product size
AB194714	Trinucleotide	(GTG)4	12	59.43	299
BU668708	Trinucleotide	(ACC)4	12	60.16	291
BU669027	Tetranucleotide	(GCTA)3	12	60,18	304
BU669798	Tetranucleotide	(CTCA)3	12	59.76	183
BU669124	Dinucleotide	(GA)6	12	60.74	280
BU669113	Trinucleotide	(GAA)4	12	59.97	129
BU669957	Dinucleotide	(AG)8	16	60.80	272
BU668438	Dinucleotide	(TC)8	16	59,28	313
BU667825	Trinucleotide	(TTC)4	12	58.73	120
BU669908	Tetranucleotide	(TTGT)4	16	59.81	314
BU670499	Trinucleotide	(TAT)4	12	59.5	109
BU668365	Dinucleotide	(AG)11	22	59,47	193
BU669924	Dinucleotide	(GA)8	16	59,26	264