# MORPHOMETRIC, MtDNA AND MICROSATELLITE ANALYSIS IN HONEYBEE POPULATIONS (*Apis mellifera* L.) OF NORTH AND NORTHWEST IRAN

A THESIS SUBMITTED T0
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

HOUMAN JABBARI FARHOUD

IN PARTIAL FULFILLMENT OF THE REQUIRMENTS FOR
MASTER DEGREE OF SCIENCE
IN
BIOTECHNOLOGY

SEPTEMBER 2004

Approval of the Graduate School of Natural and applied Sciences

_____

Prof. Dr. Canan ÖZGEN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

_____

Assoc. Prof. Dr. F. Dilek SANİN
Chairman of the Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

| | |
|---|---|
| _____ | _____ |
| Prof. Dr. Mahinur S. AKKAYA | Assoc. Prof. Dr. Meral KENCE |
| Co-supervisor | Supervisor |

Examining Committee Members

| | |
|---|---|
| Prof. Dr. Neşe ÇAĞATAY | _____ |
| Prof. Dr. Işık BÖKESOY | _____ |
| Prof. Dr. Musa DOĞAN | _____ |
| Assoc. Prof. Dr. Meral KENCE | _____ |
| Prof. Dr. Aykut KENCE | _____ |

# PLAGIARISM

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

Name, Last name : Houman, Jabbari Farhoud

Signature :

# ABSTRACT

MORPHOMETRIC, MtDNA AND MICROSATELLITE ANALYSIS IN HONEYBEE POPULATIONS (*Apis mellifera* L.) OF NORTH AND NORTHWEST IRAN

Jabbari Farhoud , Houman
M.Sc. Department of Biotechnology
Supervisor: Assoc. Prof. Dr. Meral Kence
Co-supervisor: Prof. Dr. Mahinur S. Akkaya
September 2004, 71 pages

Morphometric measurements, mitochondrial DNA analyses and 5 microsatellite loci were used to investigate variation in the honeybee populations of Iran and comparing it with the Turkish populations. Five honeybee populations were sampled from North and west north of Iran.

In morphometric aspect of the study 23 characters were measured from left forewings and hindlegs of honey bee samples. The data were analysed by multivariate statistical analyses.

By using mtDNA analyses length polymorphism of the intergenic region COI-COII of mitochondrial DNA was studied. After amplification of this region by the polymerase chain reaction, *Dra*I enzyme was used for restriction of amplified region. Results of mtDNA studies show no diversity between four populations and all of them exhibit the same C1 pattern.

Five microsatellite loci (A7, A24, A28, A43 and A113) were used in this studies.A high level of average heterozygosity changing between 0.611 and 0.709 was detected in Iranian honey bee populations, and a significant degree of polymorphism was observed. Although Urmia, Sarein and Viladereg populations are similar, Amol population which has located in northern Iran shows a significant difference from other populations. Result obtained form morphometric studies are supporting microsatellite analyses. By comparing data obtained form Iranian honey bee populations with Turkish population (Hakkari), western populations (Urmia, Sarein and Viladereg) are more similar to Hakkari population. It is found Amol is significantly different form other populations and better represents Iranian honeybee.

Keywords: Microsatellite, mtDNA, morphometry, honey bee, *Apis mellifera*, Iran

# ÖZ


İRAN'IN KUZEY VE KUZEYBATISINDAKİ BALARISI POPULASYONLARININ (*Apis mellifera* L.) MORFOMETRİK, MtDNA VE MİKROSATELİT ANALİZİ

Houman Jabbari Farhoud

Tez yöneticisi: Doç.Dr. Meral Kence

Ortak tez yöneticisi: Prof. Dr. Mahinur S. Akkaya

Eylül 2004, 71 sayfa

Bu çalışmada, morfometrik ölçümler, mitokondrial DNA (mtDNA) ve mikrosatelit analizleri İran ve Doğu Anadolu bölgesindeki balarısı populasyonlarının genetik çeşitliliğinin araştırılmasında kullanılmıştır. İran'ı temsil eden bal arısı populasyonu İran'ın kuzey ve kuzey batısından toplanan beş farklı balarısı populasyonları ile oluşturulmuştur.

Morfometrik ölçümler, balarası örneklerinin sol arka kanatlarına ve arka bacaklarına ait 23 karakterinde gerçekleştirilmiştir. Çalışmaya ait veriler multivariyete istatistik analizi ile değerlendirilmiştir.

Mitokondrial DNA analizi kulanılarak, mtDNA'nin COI-COII intergenik bölgesi araştırılmıştır. Bu bölgenin Polimeraz Zincir Reaksiyonu (PZR) ile

çoğaltılmasının ardından, *Dra*I enzimi çoğaltılmış bölgelerin kesimi için kullanılmıştır. Analiz sonucunda İran'dan alınan 4 populasyon (Urmia, Sarein, Viladareg, Amol) arasında genetik çeşitliliğin olmadığı ve aynı C1 patternini verdiği gözlenmiştir.

Çalışmada beş mikrosatellit lokusu (A7, A24, A28, A43, A113) kullanılmıştır. Populasyonlar arasındaki ortalama heterozigotluk değeri 0.611 ve 0.709 arasında değişmektedir ve polimorfizmin önemli olduğu gözlenmiştir. Urmia, Sarein, Viladarag populasyonları arasında genetic benzerliğin yüksek olmasına rağmen, İran'ın küzeyinde bulunan Amol populasyonunun önemli derecede farklı oldugu gözlenmiştir. Morfometrik çalışmalarından elde edilen sonuçlar mikrosatelit analizlerini desteklemektedir. İran'ın batısındaki populasyonlar ile Türkiyeden elde edilen Hakkari populasyonunun genetic olarak benzer olduğu gözlenmiştir. Diğer taraftan, Amol balarısı populasyonu diğer bölgeler ile önemli derecede farklılık göstermiştir ve İran balarısnı daha iyi temsil etmektedir.

Anahtar kelimeler: Mikrosatelit, mtDNA, morfometri, Balarısı, *Apis mellifera,* İran.

AKNOWLEDGMENTS

I wish to thank my advisor Assoc. Prof. Dr Meral KENCE for her guidance, advices and patience throughout this research.

I would like to thank to the members of my dissertation committee to Prof. Dr. Aykut KENCE, Prof. Dr. Neşe ÇAĞATAY, Prof. Dr. Işık BÖKESOY , and Prof. Dr. Musa DOĞAN for their suggestions, support and availability for consultation.

I wish to express my deep gratitude to my laboratory mates Çağrı BODUR, Deniz ÖZÜT, Rahşan İVGİN TUNCA, Damla BETON for their help in all steps of laboratory work. I also thank to Mehmet DEĞİRMENCİ for his technical assistance.

Many thanks for Ali Jafari and bee keepers who helped me in collection of samples.

I would like to thank my parents for encouraging me to pursue my own interest as I grew up, and for their support, both financial and emotional,

which allowed me to complete my bachelor's degree and to go on to complete this achievement.

This study was supported by Research Fund Project of Middle East Technical University (BAP-2002-07-02-00-51).

# TABLE OF CONTENTS

# CHAPTER 1

# INTRODUCTION

Honeybees first time were classified by Carlous Linnaeus (Dietz, 1986). All of bees species belong to the order of insecta known as Hymenoptera, (literally "membrane wings"). This order, comprise some 100,000 species, also includes wasps, ants, ichneumons and sawflies. The sub-family Apinae or honeybees, comprises a single genus, Apis (Table 1).

Fossil evidence is sparse but Apidae family probably appeared on earth about the same time as flowering plants in the Cretaceous period, 146 to 74 million years ago, and at about 40 million years ago the genus Apis originated (Ruttner 1988). The oldest known fossil bee, a stingless bee named *Trigona prisca*, was found in the Upper Cretaceous of New Jersey, U.S.A., and dates from 96 to 74 million years ago. It is indistinguishable from modern Trigona. The ancestor of the honeybees may have been living about this time, but fossils of the true Apis type were first discovered in the Lower Miocene (22 to 25 million years ago) of Western Germany (http://www.angus.co.uk/bibba/bibborig.html). African and European subspecies of the honey bee have been geographically separated for around 10,000 years Ruttner 1986).

Table1. Classification of honeybee

| | |
|---|---|
| Kingdom | Animalia |
| Phylum | Arthropoda |
| Class | Insecta |
| Order | Hymenoptera |
| Suborder | Aculeata |
| Superfamily | Apidea |
| Family | Apidae |
| Subfamily | Apinae |
| Tribe | Apini |
| Genus | *Apis* |
| Species | *mellifera* |

Two hypotheses have been stated about evolution of honeybees. According to the one hypothesis, honeybees evolved in near Ethiopia in Africa and extend over Africa and then Middle East and Europe (Wilson 1971). However the second hypothesis explains that, although representatives of most of bee species were original to all the continents, bees belonging to the genus Apis were originally to be found only in the Old World, namely southern part of Caspian Sea and then to Africa and Europe (Ruttner 1988).

*Apis mellifera* L. (the western honeybee) was originally found in Africa, near and Middle East and Europe. Then it was brought to new world continents by the European settlers (Ruttner 1988).

The low migration ability and small population size of *A. mellifera* led to a high level of genetic differentiation and large number of recognized subspecies (Garnery 1992). These subspecies have been studied mainly on morphometrical ground (Ruttner 1988). One of the main studies on *Apis mellifera*, which was carried out by Ruttner (1988) reports that there are 4 main evolutionary lineages and 24 subspecies of *A. mellifera*. This report is based on multivariate analysis of morphometrical characters of *A. mellifera* biogeographical and taxonomical evaluations. The 4 evolutionary branches of *A. mellifera* proposed by Ruttner 1988 were lineage A lineage M, lineage C. Lineage O stands for oriental branch (near east), races lineage A represents tropical African races, lineage M represents western Mediterranean (northern African, western and northern European). And lineage C represents central Mediterranean and south Eastern European (Ruttner 1988).

Turkey and especially Turkish border of Iran are referred as the genetic center of the honeybees on the basis of morphometrical studies (Ruttner 1988).

Although it has long been known that there are many subspecies of honeybee, these have been the subject of scientific study for more than two centuries. Only in recent years a comprehensive classification has been attempted which takes into account, not only the differences in physical characters between subspecies and their present geographical distribution, but also the geological evidence pointing to their origins, and to the course of their subsequent evolution and distribution.

The different races of *A. mellifera* can generally be differentiated in physiological terms. Bees from warmer climates tend to be smaller in size and lighter in color than those adapted to the colder regions, although this rule is not invariable. The effect of altitude seems to be similar to that of increasing latitude. Accurate differentiation between races of similar appearance requires precise morphometric examination of representative samples of bees. (http://www.angus.co.uk/bibba/bibborig.html)

According to the previous studies 7 species of honeybee has been reported on the world: *Apis floraea, Apis dorasata, Apis koschevinkovi, Apis cerana, Apis mellifera*, and *Apis laborisa*. From these species at least 2 have been reported in Iran. *Apis mellifera* exists in all of Iran except desert regions and *Apis floraea* lives in central and southern parts of Iran. Early studies show at least 2 groups of honeybee, can be discriminated in north of Iran, and in west and center of Iran. Between center and west populations some differences can be observed, although in most of the forms, these 2 groups integrate in each other (G. Tahmasbi 1996).

## 1.1. Morphologic Analysis

Morphologic analysis has been used to discriminate honeybee groups and for many years it was the only way to describe the genetic diversity of honeybee. It measures both genetically based and environmentally induced variation using one or several variables (Daly, 1988). Morphometry is so powerful, since it has not only able to the discriminate subspecies and branches, but also in the sense of Louveaux (1966), that is, sets of populations from the same geographic area are characterized by a specific biological cycle of colonies.

Ruttner (1988) by re-examining his morphological data introduced a fourth branch (O) composed of the subspecies *meda-anatolica- caucasica*. Currently 26 subspecies of *Apis mellifera* (Table2) are formally recognized based primarily on morphometric characters (Ruttner 1988, 1992; Sheppard *et al.* 1997).

Multivariate analysis use the measurement of different morphological characters to obtain a probability value. Morphometrics cannot be used effectively to identify hybrids and, therefore, is not effective in studying gene introgression and hybridization. Nevertheless, morphometrics has been used to follow changes in Africanization of some populations in New world (Bohreham and Roubik, 1987).

Table 2: Honeybee subspecies according to their geographical distribution (Ruttner 1992 ; Sheppard *et al.* 1997)

| |
|---|
| Near East Subspecies |
| *Apis mellifera anatoliaca* Maa (1953) |
| *A. m. adami* Ruttner ( 1975) |
| *A.m. cypria* Pollman (1879) |
| *A. m. syriaca* Buttel-Reepen (1906) |
| *A. m. meda* Skorikov ( 1929) |
| *A. m. caucasica* Gorbachev (1916) |
| *A. m. armeniaca* Skorikov (1929) |
| Tropical Africa Subspecies |
| *Apis mellifera lamarckii* Cockerell (1906) |
| *A. m. yemenitica* Ruttner (1975) |
| *A. m. litorea* Smith ( 1961) |
| *A m. scutellata* Lepeletier (1836) |
| *A.m. adansonii* Latreille (1804) |
| *A. m. monticolo* Smith (1961) |
| *A. m. capensis* Escholtz (1821) |
| *A. m. unicolor* Latreille (1804) |
| West Mediterranean (West and North Europe and North Africa) Subspecies |
| *Apis mellifera sahariensis* Baldensperger (1924) |
| *A. m. intermissa* Buttel-Reepen (1906) |
| *A. m. iberica* Goetze (1964) |
| *A. m. mellifera* Linnaeus (1758) |
| *A. m. major* Ruttner (1978) |
| Central Mediterranean and Southeast Europe Subspecies |
| *Apis mellifera sicula* Montagano (1911) |
| *A. m. ligustica* Spinola (1806) |
| *A. m. cecropia* Kiesenwetter (1860) |
| *A. m. macedonica* Ruttner (1987) |
| *A. m. carnica* Pollman (1879) |
| *A. m. rutnerii* Sheppard *et al.* (1997) |

Studies on Morphologic variation for honeybees of Iran reveal that Iranian honeybees are mostly same with Iranian subspecies, *Apis mellifera meda,* and has a large distance with European subspecies that has imported to Iran (G. Tahmasbi 1996).

## 1.2. Allozyme Studies

Allozyme characters of individuals are independent of environmental influences, being the direct product of structural genes (Daly, 1988). However the existence of protein variants is subject to selection and their detection depends on their expression (Hall, 1988). Some of the polymorphic proteins available for honeybees also include the enzymes esterase, hexokinase, alcohol dehydrogenase, malate dehydrogenase, and a non enzymatic protein P-3 (Contel *et al.*, 1977; Mestriner, 1969; Mestriner and Contel, 1972; Sheppard and Berlocher, 1985), (Kandemir *et al.* 2000).

In recent years a new emphasis has been given to the use of easily processed DNA markers in both applied and evolutionary entomology. New methods have become available that could bring together traditional entomology and molecular biology. The interest of the method has been approved because it is fast, technically easy and requires no prior knowledge of the genome analysis.

## 1.3. Mitochondrial DNA:

Mitochondrial DNA is especially useful for inferring phylogenetic relationship among the components of a species. It has become a useful approach to population studies addressing the problem of Africanized bees.

Since all the individuals of colony are the progeny of the queen, their mtDNA is identical, so the colony's individuals are relevant in analysis. Changes in mtDNA sequences reflect the accumulation of mutations. Therefore in colonization process of honeybees, the important role is played by swarms (maternal division of colony), i.e., with colonies, and not by males (Garnery *et al* 1992). So, inheritance of mtDNA makes it even more suited to studies on honeybees. Drawback of mtDNA is its relatively small size. This restricts the number of polymorphisms that can be detected.

The mitochondrial genome of honeybee contains a non coding region located between a leucine tRNA gene and the Cytochrome oxidase II gene CoII; (Cornuet *et al.* 1991). This non coding sequence has a complex structure which gives rise to both sequence and length variation (Hall and Smith 1991). The three lineages differ in length. In these studies a particular intergenic non-coding region which is located between Cytochrom Oxidase I (Co I) $T_{RNA}LUE^{UUR}$ gene and Cytochrome Oxidase II (Co II) gene has been used frequently. The non-coding region of western honeybee, *Apis mellifera*, has some unique characters that make it different from other species of honeybee. Western honeybee has a non-coding region which can be divided in three distinct regions on the basis of sequence and secondary structure. These are called, P unit, and Q1, Q2, and Q3 subunits. P unit is observed in all populations except, eastern Mediterranean region. The polymorphism in P region is represented as $P_0$ (67bp) in African population and P (54bp) in northern and Western Europe. Size variation is observed in Q unit and the entire sequence has composed of approximately 190-200 base pairs (Garnery *et al.* 1993).

The number of Q units also shows variation between different populations. In eastern Mediterranean populations there is only single Q unit while in the

African populations the $P_0$ unit is followed with one to four copies of the Q unit which are $P_0Q$, $P_0QQ$, $P_0QQQ$ or $P_0QQQQ$. In northern European populations, also a similar model is observed. The only difference is P unit which is not $P_0$ in this population (Cornuet *et al.*1991, Garnery *et al.* 1993). Preliminary works on mtDNA restriction site polymorphisms and sequence polymorphisms have revealed three lineages of honeybee mtDNA. Western European, Mediterranean and African (Smith, 1991a, b; Garnery *et al.* 1992). Smith *et al.* (1997) and Palmer *et al.* (2000) have studied mtDNA for identification of lineages in honeybees of Turkey. On the basis of mtDNA haplotypes, they concluded that Turkish honeybees generally belong to eastern Mediterranean (C) lineage.

## 1.4. Random Amplified Polymorphic DNA (RAPD):

Random amplified polymorphic DNA or RAPD is another nuclear DNA marker that is commonly used to discriminate subspecies (Williams *et al.*, 1990). The simplicity of the RAPD analysis makes them popular for population genetics studies. These markers are generated by a PCR (Polymerase Chain Reaction) using a single ten-oligonucleotide primer of random sequence. Nevertheless, RAPD markers are generally dominant, that is, homozygous and heterozygous individuals cannot be distinguished. The other reasons that limit using of these markers more, are related with reliability and reproducibility of them (Black, 1993; McPheron *et al.*, 1993).

## 1.5. Microsatellite Markers:

Microsatellite markers are areas of DNA having short sequences of nucleotides which are repeated several times. These markers are flanked by identified primer sequences which make the markers unique within the genome. Microsatellite markers are commonly found in the majority of a species; however different individuals have different numbers of sequence

repeats. Replication slippage and unequal crossing over is the main mechanism for mutations in microsatellites. It occurs because of the mistakes in dissociation and reannealing during replication. Nascent DNA strand anneals in wrong position because of the repetitive sequences.

Another mechanism in occurring mutations in microsatellite is unequal crossingover. It is usually observed between DNA molecules having repeated unit regions. Misaligment and crossingovers, in chromosomes results in deletion and insertion.

Most microsatellite loci investigated so far have been taken from published sequences. In contrast to mammalian, there is little information about microsatellites in other vertebrates and insects. Few studies in this subject was committed on Drosophila species. According to these studies presence of $(CT)_n$ and $(GT)_n$ can be observed, which are similarly distributed in distantly separated species (Pardue, et al 1987). Information obtained from these studies were used to distinguish $(CT)_n$ and $(GT)_n$ repetitive sequences in honeybee. A set of 52 $(CT)_n$ and 23 $(GT)_n$ microsatellites have been isolated and sequenced from partial genomic libraries of honeybee. Generally, $(CT)_n$ and $(GT)_n$ microsatellites are observed in every 15 kb and 34 kb in honeybee genome (Estoup *et al,* 1993).

Such genetic differences can be used in population genetics. The primer sequences are used with PCR to determine the marker size for individual samples. Results obtained from microsatellite studies confirm Ruttner *et al* (1978) hypothesis based on morphological data. Microsatellite analysis performed by Estoup *et al.* (1995) clearly confirms the existence of three evolutionary branches, A, M and C, presented by Ruttner. These studies also

show microsatellites are useful in discrimination of subspecies and populations within subspecies.

The structure and the core sequence of 552 microsatellites with polymorphic alleles were reported by Solignac *et al.* (2003). High number of microsatellite markers are being used in doing genomic research on honeybees such as genetic mapping, library screening, artificial cloning.

Eight microsatellites loci were identified which provide genetic information in a population of honeybees in Iberia (P.De La Rua *et al.* 2002). Six of microsatellite loci showed intermediate polymorphism and was found that microsatellite alleles in *A. m. iberica* observed in the populations of south-eastern Spain are similar to that found in populations from Western Europe.

Few studies have been done in Turkish honeybee populations. In a study by Kandemir (1999), Kırklareli were discriminated from Van, Kastamonu and Balıkesir populations. According to results obtained form studying 3 microsatellite loci in Turkey honeybee population, Kırklareli samples are considerably different from the rest of the locations and concluded the hypothesis that Anatolian honeybees belong to C lineage (Palmer 2000) is not supported by the microsatellite data by (Bodur 2001).

## 1.6. The Goal of the Study:

Although bee keeping is common in different parts of Iran, less importance is given to conserving genetic diversity of honeybees. Migratory bee keeping and importing of queens are two main factors in increasing of homozigosity and loosing of diversity between local honeybee populations. This may result in disappearing of local populations. Genetic diversity between populations

can be used in studies for improving of honeybee races which are economically valuable and more resistant to diseases.

Our study focuses on lineage C type of *A. mellifera* species of origin north-western Iran. According to studies done on the *A. mellifera*, subspecies of Turkey revealed great amount of variability in their electrophoretic loci. (Kandemir *et al.,* 2000). Reasoning from these experiences we expect many variations in these regions, linage C races of *A. mellifera* subspecies.

The aim of this study is to determine amount of genetic diversity in 5 Iranian honeybee populations and comparing it with Turkish honeybee population. Samples were collected from North and North West of Iran. Therefore microsatellite and mitochondrial DNA analysis, and morphometric studies used, to find out the extent and level of variation in honeybee populations in Iran. These data also will be used to study the hypothesis which is about importance of Near East and Iran in evolution of honeybees (Garnery 1992, Ruttner 1988).

# CHAPTER 2

# MATERIALS AND METHODS

## 2.1. Biological Material

Accurate sampling is an important part of biological studies. In honeybees since most of the individuals of a colony are descendents of a single queen. It is suggested to collect few workers from a high number of colonies, instead of sampling from few colonies in a location.

With considering sampling conditions stated above, honeybee workers collected from 5 different locations belonging to 5 provinces Urmia, Sarein, Viladereg, Amol and Alborz region which are located in north and North West of Iran. Names of the provinces and their locations are given in figure 1. Samples have been kept in absolute ethanol until DNA isolation step.

After samples were taken out from alcohol, head, wings, thorax and legs are removed. Wings and fore legs are used for preparing of slides which are used in morphometric measurements. DNA extraction is done on head and thorax. For mtDNA experiments, DNA extracted from thorax samples showed better results.


Figure 1: Map of Iran and collection sites.

## 2.2. Morphometric Measurements:

Seven individuals of each colony were used in morphometric measurements. The left forewings and hindlegs of each individual were fixed on a slide by using a transparent tape. An Olympus dissection microscope was used to magnify the samples by 1.2X (1mm=3.08cm). The image was transmitted to a monitor by a JVS camera control unit.

Forewing width (FW), cubital index of 'a' vein (a), cubital index of 'b' vein (b), c and b values, $L_1$ and $L_2$ parameters were measured from forewings. (Figures 2 ). Figure 3 shows the 11 angles measured from forewings.

For each sample 23 characters were measured.11 angular and 8 metric parameters of forewing and 4 metric parameters of hindlegs (Figure 4).
4 metric parameters of hindleg are:
Femur length (FeL)
Tibia length (TiL)
Metatarsus (Mew)
Metatarsus length (ML)

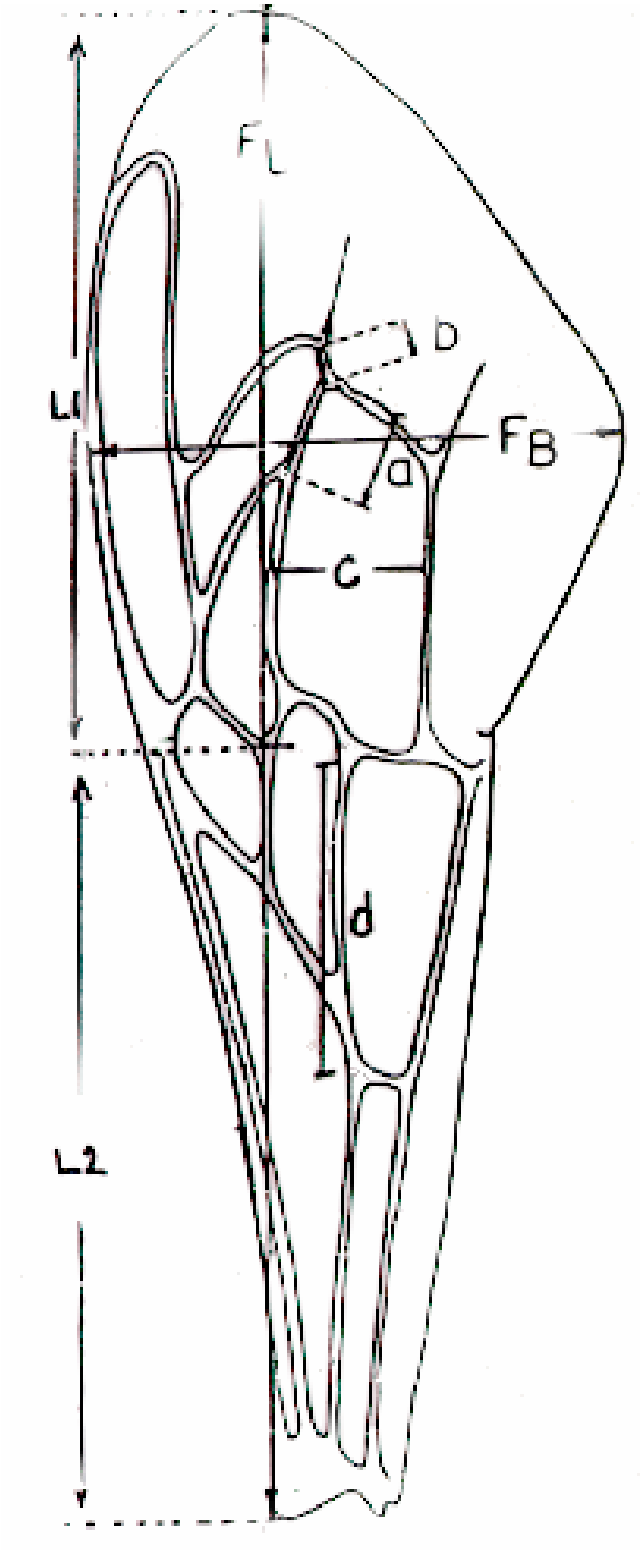


Figure 2: parameters measured in forewing.

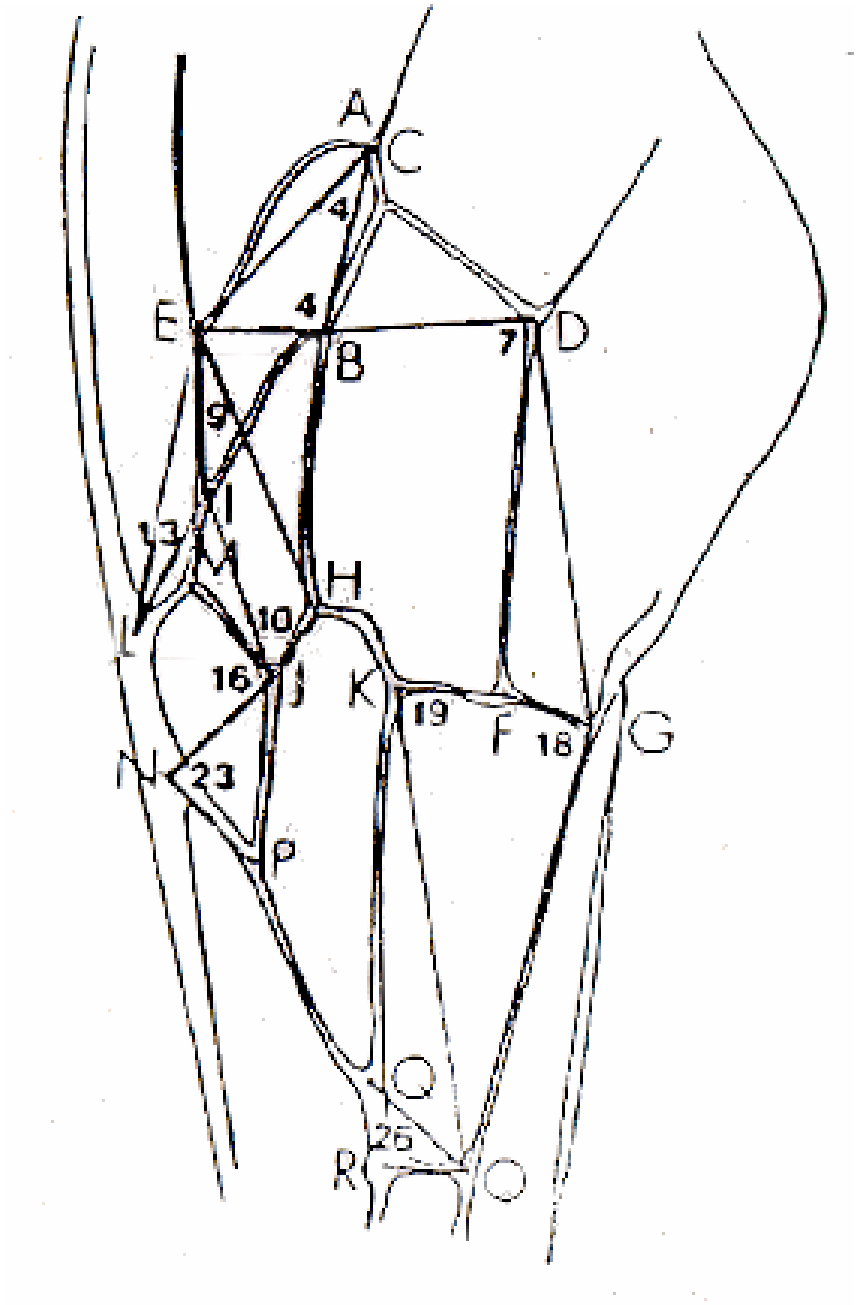


Figure 3: The 11 angles measured from forewings.

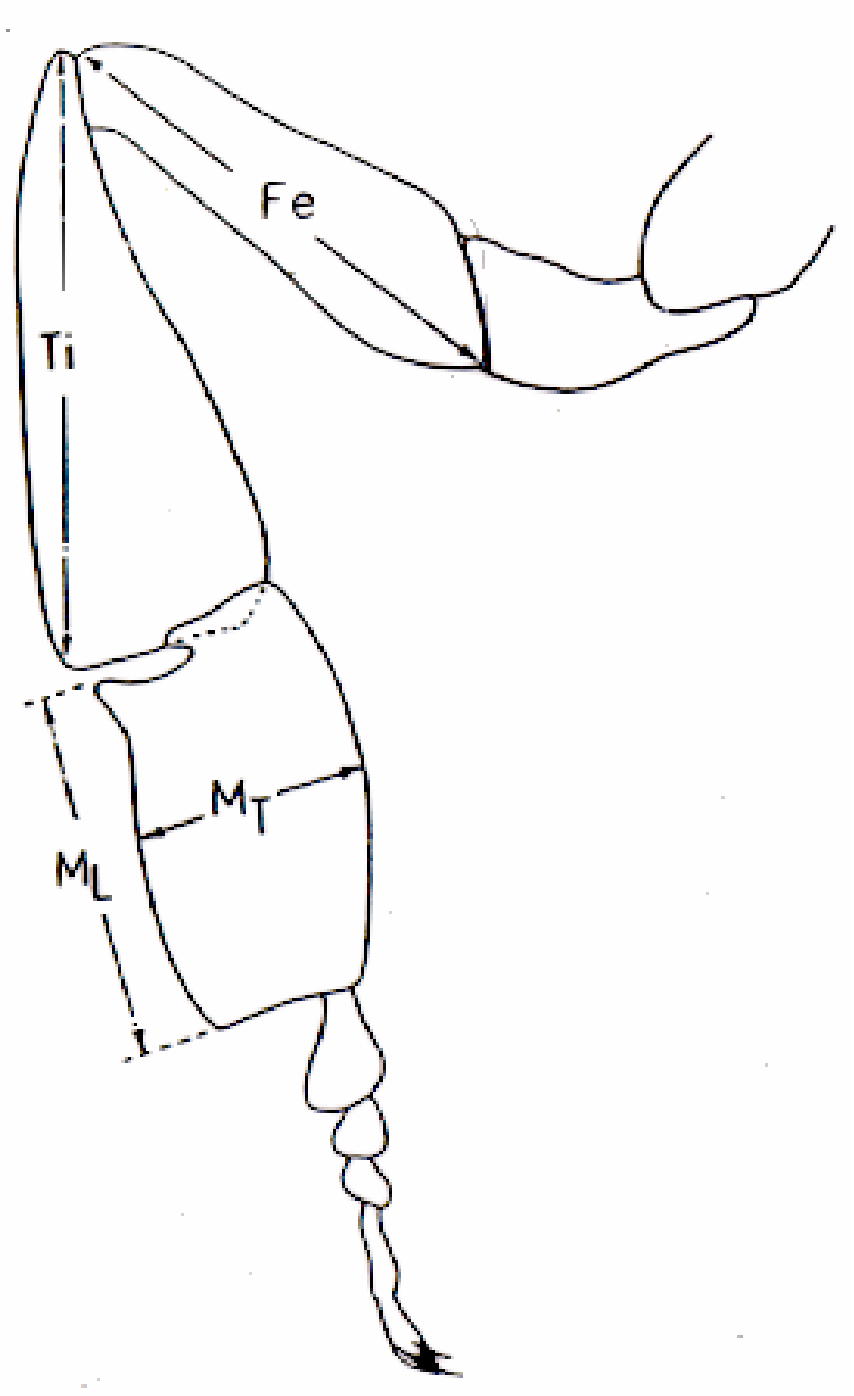


Figure 4: The 4 metric parameters of hindleg.

## 2.3. Statistical Methods for Morphometric Measurements:

In this study NTSYS (Rohlf, 1992) computer program was used for analyzing morphometric data. Principle Component Analysis (PCA) was performed by SYNTAX software. PCA is usually used in analyzing the relationship between variables and is one of the most important methods in ordination analysis.

## 2.4. DNA Isolation:

Each head was then grinded in a 1.5 ml tube with a sterile pestle immediately after immersing the tube containing head into liquid nitrogen and 750 µl of Wilson buffer (Appendix B) was added into the tube. Twenty five µl of 10 mg/ml Proteinase K was added into each tube. Tubes are mixed and then incubated for two hours in a water bath at 50°C. After incubation tubes were centrifuged at 10000 rpm for 10 minutes, the upper phase solution was poured into a new tube. Seven hundred and fifty µl of phenol:chloroform:isoamylalcohol (25:24:1 vol.) was added and tubes were centrifuged at 10000 rpm for 20 minutes after gentle inversions of five minutes. Then 600 µl of aqueous phase was removed into a new tube and then the previous extraction step was repeated twice, first by adding 600 µl of phenol:chloroform:isoamylalcohol (25:24:1 vol.) and then 450 µl of chloroform: isoamylalcohol (24:1 vol.) to the removed 450 µl of aqueous phase. Recovered 300 µl of aqueous phase was transferred into a new tube and added with 30 µl of 3 M sodium acetate, 600 µl of absolute alcohol and stored at -20 ºC overnight after mixing for a few minutes.

The tubes were centrifuged at 13,000 rpm for 30 minutes and the supernatant was discarded. 900 µl of 70% ethanol was added and the tubes were centrifuged at 13000 rpm for 20 minutes. After pouring alcohol off, the

pellet was dried in a desiccators for 20 minutes. The pellets in the tubes were added 50 µl of sterile water and kept at room temperature for one hour. DNA solutions were examined under UV illumination at 230, 260 and 280 nm for detection of absorptions of RNA, DNA and protein parts respectively, if available in solution and run on 1 % agarose gel electrophoretically to confirm the presence of DNA.

## 2.4. Mitochondrial DNA Amplification by PCR

The $t_{RNA}$leu-COII intergenic region of the mtDNA was amplified with the corresponding 1x buffer, 2 mM $Mgcl_2$, 0.3mM of each dNTP, 0.25µM of primers E2 (5'-GGCAAGAATAAGTGCATTG-3') and H2 (5'-CAATATCATTGATGACC-3') (Garnery *et al.*1992) and 3 units of Taq DNA polymerase in a total volume of 100 µl. Polymerase chain reaction were submitted to an initial denaturation of 5 minutes at 97 $^{o}$C, followed by 30 cycles of 95$^{o}$C for 30 seconds, 50$^{o}$C for 90 seconds and 72$^{o}$C for 90 seconds. 5µl of PCR product was electrophorased in a 2% agarose gel in order to check the result of the reaction.

### 2.4.1. Digestion of PCR Product:

Restriction enzyme digest is usually performed in a volume of 2µl on 0.2-1.5µg of substrate DNA using 2 to 10 fold excess of enzyme over DNA. Using an unusual excess amount of enzyme may produce abnormal results. In a sterile tube 13.3µl sterile deionised water, 2µl restriction enzyme buffer, 0.2µl acetylated BSA (10µg/µl) and 3µl DNA, were mixed by pipetting. Then 0.5µl of *Dra*I restriction enzyme (TTTAAA) was added to the mixture. After mixing by pippet, tubes were centrifuged for a few seconds in a microcentrifuge. These steps should be done on ice to prevent losing activity

of the *Dra*I enzyme. After adding 1 drop of mineral oil on each tube, the mixture was incubated at 37$^{o}$C for 3-4 hours.

### 2.4.2. Preparation of the Gel

After digestion step samples were electrophorased two times .10% polyacrylamide gel for separating long fragments and 3.6% nondenaturating polyacrylamide gel for separating shorter restricted fragments .

### 2.4.3. Pouring the Gel

Before preparing of gel solution two glass plates1 mm. plastic spacers are squeezed in a vertical gel caster. Gel mix was poured between the two glass plates immediately after the addition of N, N, N', N'-tetramethylethylenediamine (TEMED), by using a syringe. After pouring the gel, comb is inserted between the glasses.

### 2.4.4 Loading and Running the Gel:

After digestion step, 5µl of bromophenol solution was added to each tube. In the first electrophorase, 8-10µl of mixture was loaded in 10% acrylamide gel and run at 45 Ampere, 250 volt for 2.5 hours. In the second electrophorase 1.5µl of mixture was loaded in the 3.6% acrylamide gel and run for over night at 35 volt with 10-13 Ampere.

### 2.4.5. Staining the Gel:

After running the gel, glasses are separated carefully in a way that gel remains on one of them without tearing. Gel was soaked in ethidium bromide solution for 20-30 minutes and shaked gentelly for better staining. Restriction fragments are visualized under ultraviolet.

## 2.5. Microsatellite Amplification by PCR

In this study five microsatellites loci (A7, A24, A28, A43, and A133), were selected from available *A. mellifera* microsatellites (Estoup *et al.* 1993). The core regions and primer sequences of loci are given in (Table 3 and 4).

Table 3. Core sequences of microsatellites

| **Locus** | **Core sequence** |
|---|---|
| **A7** | $(CT)_3(T)_7CCTTCG(CT)_{24}$ |
| **A24** | $(CT)_{11}$ |
| **A28** | $(CCT)_3GCT(CCT)_6(CT)_5$ |
| **A43** | $(CT)_{13}$ |
| **A113** | $(TC)_2C(TC)_2TT(TC)_5TT(TC)_8TT(TC)_5$ |

Table 4. Forward and reverse primers used to amplify microsatellites

| Locus | **Forward primer** | **Reverse primer** |
|---|---|---|
| A7 | 5'GTTAGTGCCCTCCTCTTGC3' | 5'CCCTTCCTCTTTCATCTTCC3' |
| A24 | 5'CACAAGTTCCAACAATGC3' | 5'CACATTGAGGATGAGCG3' |
| A28 | 5'GAAGAGCGTTGTTGCAGG3' | 5'GCCGTTCATGGTTACCACG3' |
| A43 | 5'CACCGAAACAAGATGCAAG3' | 5'CCGCTCATTAAGATATCCG3' |
| A113 | 5'CTCGAATCGTGGCGTCC3' | 5'CCTGTATTTTGCAACCTCGC3' |

Radioactive PCR amplifications were carried out according to Estoup *et al.* 1995. Twenty five microliter of amplification reactions were performed with 50 ng of template DNA, 400 nM of each primer, 75 µM of each 2'-deoxythymidine 5'-triphosphate (dTTP), 2'-deoxyguanidine 5'-triphosphate (dGTP) and 2'-deoxycytidine 5'-triphosphate (dCTP), 7.5 µM of 2'-deoxyadenosine 5'triphosphate (dATP), 0.25 µCi of $\alpha^{33}$P-dATP, 20 µg/ml bovine serum albumin (BSA), 1x reaction buffer containing $(NH_4)_2SO_4$, 0.4 unit of Taq polymerase, and 1-1.2 mM $MgCl_2$. PCR started with a denaturation step of 3 minutes at 94 ºC and continued with 30 cycles, containing; a 30 second denaturation segment at 94 ºC, a 30 second annealing segment at the optimum temperature, and a 30 second elongation segment at 72 ºC. The final elongation step was extended to 10 minutes in order to allow all the products to be fully extended. The annealing temperatures and MgCl2 concentrations that were used for each microsatellite loci, are given in table below.

| Locus | Annnealing temperature ($^oC$) | $MgCl_2$ concentration (M) |
|---|---|---|
| A7 | 60 | 1,2 |
| A24 | 56 | 1,2 |
| A28 | 54 | 1,7 |
| A43 | 55 | 1,5 |
| A113 | 60 | 1,2 |

### 2.5.1 Sequencing Polyacrylamide Gel Electrophoresis

A sequencing polyacrylamide gel electrophoresis apparatus was used for discriminating of alleles which are different in one or more nucleotides.

### 2.5.2. Cleaning the Glass Plates

Glass plates with edges of twenty and forty five centimeters were used in electrophoresis. One side of each plate were cleaned carefully first by distilled water and then by absolute ethanol in order to prevent any debris on the surface to interfere with the progress of DNA fragments during electrophoresis. Then a silanizing solution was applied to one clean surface of a glass plate to make it easier to remove one of the plates after electrophoresis, the intact gel remaining on the other plate.

### 2.5.3. Preparation of the Gel

A 6% denaturing polyacrylamide gel was used in electrophoresis. A 6% acrylamide-urea mix (Appendix B) containing 8 molar of urea was prepared and put in a light-tight bottle and kept at 4°C. Six hundred and fifty µl of 10% (v/v) ammoniumpersulfate (APS) and 30 µl of N, N, N', N'-tetramethylethylenediamine (TEMED) was added to 50 ml of acrylamide/urea mix just before pouring the gel.

### 2.5.4. Pouring the Gel

A gel caster, a comb and 0.4 mm. plastic spacers were used. Gel mix was poured on one of the glass plates which is fixed horizontally in the gel caster,

immediately after the addition of TEMED by using a syringe. The upper glass plate is slided slowly on the other plate as the gel is poured. Being spacers adhered (by water drops) on the lower plate, gel solution fills the area between plates. After pouring the gel, comb is inserted and metal clamps were used at the edges of the plates to squeeze them.

### 2.5.5. Loading and Running the Gel

PCR reactions containing 25 µl of DNA solution were added 10 µl of loading dye solution (Appendix B) each and 2.5 µl of these mixes were loaded to the gel placed in the vertical gel apparatus by using an ordinary micropipettor. A sequencing reaction done by USB Sequenase Version 2.0 DNA Sequencing Kit using $\alpha^{33}$P-dATP, was exploited as size marker to determine the exact sizes of DNA fragments. Upper and lower reservoirs of sequencing gel electrophoresis apparatus were filled with 1x tris-boric acid-edta (TBE) buffer, and it was run at 40 Watts and for 2.5-3 hours depending on microsatellite loci.

### 2.5.6. Autoradiography

After the run, siliconized plate was removed and the gel which remained on the other plate, was taken onto a chromatography paper (Whatman 3MM). Gel was covered with an ordinary stretch film and dried on a vacuum dryer at 78 °C for 30 minutes. Special autoradiography films (Kodak Biomax MR) handled in a dark room, were exposed to the dried gels in light-tight metal cassettes for 2-5 days depending on the time passed after the radioactive

material purchased. The exposed films were developed in the medical center of Middle East Technical University.

## 2.6. Statistical Methods for Microsatellite Data

Genetic variation in populations is determined by using different statistical analysis. Population genetic softwares were used for performing Hardy Weinberg equilibrium, allele frequencies, heterozygosities, gene diversity, pairwise $F_{ST}$ measures; population differentiation, linkage disequilibrium tests and genetic distance calculations and phenogram constructions.

### 2.6.1. Testing Hardy-Weinberg Equilibrium

Option one sub-option three of Genepop on the web, which is the internet version of original Genepop program (Raymond and Rousset) was used to calculate deviation from Hardy – Weinberg equilibrium. The null hypothesis assumes random association of gametes, according to exact HW tests of Haldane (1954), Weir (1990), Guo and Thompson (1992) and others. The probability of observed sample is used to define the rejection zone, and the *p*-value of the test corresponds to the sum of the probabilities of all tables, having same allelic counts, with the equal or lower probability. Estimation of $F_{IS}$ value was also calculated by Weir and Cockerham's (1984). Positive values of $F_{IS}$ are indications of a deficiency of heterozygote individuals and negative values indicate an excess of heterozygotes.

### 2.6.2. Linkage Disequilibrium:

If there is linkage equilibrium between two loci, it is not possible to predict the genotype of one chromosome at one locus by knowing the genotype of the other. Linkage disequilibrium exist when there is a nonrandom association between a chromosome's genotype at one locus and its genotype at the other locus (Freeman and Herron, 2001). Genepop 3.4 software for the test of genotypic linkage disequilibrium has the null hypothesis: genotypes at one locus are independent from genotypes at the other locus (Raymond and Rousset, 1995). Results of the test are given with the tables of significant linkage disequilibrium for each population.

### 2.6.3. Population Structure Analysis:

Population differentiation tests and F coefficients of the total population analysis were done to identify the population structures of honeybee populations. Suboptions 1, 2, 3 and 4 from option 3 of Genepop on the web software (Raymond and Rousset) were used for these tests.

#### 2.6.3.1. Population Differentiation:

Genic differentiation and genotypic differentiation analysis were done for four populations of honeybee. In genic differentiation the allelic distribution of alleles in different samples were tested. The null hypothesis is: the allelic distribution is identical through populations. The test performed in a contingency table. For each locus an unbiased estimate of the *P*-value of the probability test was done (Raymond and Rousset, 1995).

### 2.6.3.2. Genotypic Differentiation

Genotypic differentiation test considers the distribution of genotypes in a variety of populations. The null hypothesis for this test is: the genotypic distribution is identical across populations. An unbiased estimate of the P-value of a log-likelihood (G) based exact test is performed (Goudet *et al.* 1996). The test was done for each locus for all populations in the first part and the same test was done for all pairs of populations for all loci and a contingency table was constructed for each locus.

## 2.7. Phenogram Construction

Phenograms are constructed from distance matrices calculated from the microsatellite data.

### 2.7.1 Genetic Distance Calculation:

Nei's (1972) standard genetic distance Ds, standard errors of standard genetic distances (Nei 1978) were calculated among populations in order to create an input distance matrix by using DISPAN software (Ota 1993).

Nei's standard distance is calculated as;

$$Ds = -\ln\left[\frac{J_{XY}}{J_X J_Y}\right]$$

where .

$$J_X = \frac{\sum_j^r \sum_i^{m_j} x_{ij}^2}{r}$$

and

$$J_Y = \frac{\sum_j^r \sum_i^{m_j} y_{ij}^2}{r}$$

are the average homozygosities over loci in populations **X** and **Y**, respectively, and

$$J_{XY} = \frac{\sum_j^r \sum_i^{m_j} x_{ij} y_{ij}}{r}$$

$\mathbf{X_{ij}}$ and $\mathbf{Y_{ij}}$ are the frequencies of the ***i***th allele at the ***j***th locus in populations **X** and **Y**, respectively, $\mathbf{m_j}$ is the number of alleles at the ***j***th locus, and **r** is the number of loci examined.

### 2.8.2. Tree Construction:

Phenograms were constructed using the neighbour-joining (NJ) method of Saitou and Nei (1987) from the matrix of Ds distance previously obtained, by using DISPAN software (Ota 1993).

# CHAPTER 3

# RESULTS

## Morphometric Analysis

In morphometric studies 23 parameters were measured in 250 honeybee individuals.135 individual from 5 Iranian populations and 115 individuals from Turkey (Hakkari population). The first three components explain 94.82% of the total variation (Table 5). Loading of the 23 variables are given in Table 6. In the first component although FL, BDG and d variables have the highest loading, MW and FB show negative loading. FGD and ML gave the highest positive loading and BDG gave negative loading in second component. The third component is positively loaded by 'a' and ML and negatively loaded by EAB variables.

Table 5. Eigenvalues and cumulative variation explained by the first three principal components.

| | | Eigenvalue | % | Cumulative % |
|---|---|---|---|---|
| **Principle Components** | 1 | 2.67 | 50.54 | 50.54 |
| | 2 | 1,35 | 25.51 | 76.05 |
| | 3 | 0.99 | 18.77 | 94.82 |

Table 6: Loading of morphometric parameters on the first 3 components.

| Variable | 1.Component | 2.Component | 3.Component |
|---|---|---|---|
| FL | .82100 | -.17945 | -.20955 |
| L1 | -.11005 | -.16685 | -.65859 |
| L2 | -.12208 | -.36541 | -.12064 |
| FB | -.45758 | .61173 | -.67506 |
| a | .11620 | .21649 | .94568 |
| b | .65550 | -.24382 | -.43221 |
| c | -.30598 | -.24717 | -.23567 |
| d | .10413 | .25081 | -.19169 |
| FeL | .32857 | .29223 | .12286 |
| TiL | .13085 | -.12342 | .34514 |
| ML | -.43417 | .68189 | .80763 |
| MW | -.47704 | .28232 | .28871 |
| EAB | .25872 | .19388 | -.60909 |
| EBA | .28750 | .91379 | -.22254 |
| BDG | .84926 | -.52150 | .35040 |
| FGD | -.42912 | .91082 | .45669 |
| OKF | .11913 | .66544 | -.33532 |
| ROQ | -.17169 | .58272 | -.30795 |
| HEI | .12451 | .40187 | -.45970 |
| PNJ | .18921 | .33386 | -.47963 |
| NJM | -.14076 | .23317 | -.18777 |
| IJH | .14210 | .33914 | -.25083 |
| ILE | .11371 | -.17378 | .25207 |

Analysing the correlation matrix of the data indicates that d and FL has the highest correlation value. FL and L2, c and FB, TiL and FeL, d and TiL also are variables that have high positive correlation. The highest negative correlation is observed between HEI and OKF variables (Table 7).

Table 7. Correlation matrix of the morphometric data (obtained by PCA).

| | FL | L1 | L2 | FB | a | b | c | d | FeL | TiL | ML | MW | EAB | EBA | BGD | FGD | OKF | ROQ | HEI | PNJ | NJM | IJH | ILE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **FL** | 1.000 | .581 | .643 | .479 | -.025 | .371 | .365 | .658 | .538 | .514 | .481 | .534 | -.153 | .241 | .244 | -.083 | -.105 | -.114 | .153 | .126 | .136 | .158 | .126 |
| **L1** | .581 | 1.000 | -.133 | .402 | -.089 | .473 | .265 | .504 | .452 | .428 | .373 | .319 | -.180 | .226 | .226 | -.031 | -.089 | .009 | .106 | -.069 | -.012 | .009 | .115 |
| **L2** | .643 | -.133 | 1.000 | .296 | .040 | .077 | .287 | .347 | .219 | .238 | .234 | .381 | -.020 | .057 | .164 | -.169 | -.085 | -.194 | .102 | .164 | .150 | .141 | .109 |
| **FB** | .479 | .402 | .296 | 1.000 | .091 | .145 | .601 | .449 | .365 | .293 | .374 | .440 | -.112 | .173 | .138 | -.028 | .024 | -.120 | .058 | .003 | .137 | -.059 | -.042 |
| **a** | -.025 | -.089 | .040 | .091 | 1.000 | -.105 | .152 | -.025 | .038 | -.031 | -.032 | -.040 | -.149 | -.216 | -.457 | .051 | .089 | .080 | -.084 | -.129 | .081 | -.088 | -.083 |
| **b** | .371 | .473 | .077 | .145 | -.105 | 1.000 | .048 | .443 | .428 | .501 | .334 | .257 | -.204 | .218 | .314 | -.139 | .011 | -.031 | .090 | .070 | .086 | .040 | .239 |
| **c** | .365 | .265 | .287 | .601 | .152 | .048 | 1.000 | .274 | .154 | .090 | .307 | .334 | -.104 | .088 | .051 | -.227 | .108 | -.115 | -.012 | .084 | .254 | .055 | .003 |
| **d** | .658 | .504 | .347 | .449 | -.025 | .443 | .274 | 1.000 | .564 | .556 | .441 | .436 | -.209 | .229 | .234 | -.024 | -.074 | .021 | .171 | .080 | .047 | -.020 | .128 |
| **FeL** | .538 | .452 | .219 | .365 | .038 | .428 | .154 | .564 | 1.000 | .561 | .448 | .418 | -.204 | .270 | .085 | .027 | .022 | .053 | .049 | .086 | .276 | -.036 | .068 |
| **TiL** | .514 | .428 | .238 | .293 | -.031 | .501 | .090 | .556 | .561 | 1.000 | .415 | .397 | .001 | .077 | .181 | .159 | -.082 | -.035 | .112 | -.039 | -.037 | .002 | .105 |
| **ML** | .481 | .373 | .234 | .374 | -.032 | .334 | .307 | .441 | .448 | .415 | 1.000 | .529 | -.206 | .272 | .236 | -.107 | .013 | -.019 | .009 | .040 | .204 | -.051 | .104 |
| **MW** | .534 | .319 | .381 | .440 | -.040 | .257 | .334 | .436 | .418 | .397 | .529 | 1.000 | -.107 | .189 | .297 | -.114 | -.018 | .002 | .068 | -0,05 | .167 | .034 | .103 |
| **EAB** | -.153 | -.180 | -.020 | -.112 | -.149 | -.204 | -.104 | -.209 | 0,2 | .001 | -.206 | -.107 | 1.000 | -.732 | -.086 | -.061 | -.075 | -.037 | -.136 | .031 | -.029 | .205 | -.122 |
| **EBA** | .241 | .226 | .057 | .173 | -.216 | .218 | .088 | .229 | .270 | .077 | .272 | .189 | -.732 | 1.000 | .368 | -.004 | .071 | .074 | .111 | .020 | .052 | -.104 | .068 |
| **BDG** | .244 | .226 | .164 | .138 | -.457 | .314 | .051 | .234 | .085 | .181 | .236 | .297 | -.086 | .368 | 1.000 | -.341 | .034 | -.105 | .073 | .050 | .009 | .007 | .248 |
| **FGD** | -.083 | -.031 | -.169 | -.028 | .051 | -.139 | -.227 | -.024 | .027 | -.057 | -.107 | -.114 | -.061 | -0,04 | -.341 | 1.000 | -.085 | -.122 | .114 | .205 | -.037 | -.027 | .096 |
| **OKF** | -.105 | -.089 | -.085 | .024 | .089 | .011 | .108 | -.074 | .022 | -.082 | .013 | -.018 | -.075 | .071 | .034 | -0.08 | 1.000 | .156 | -.757 | .141 | .075 | .002 | .040 |
| **ROQ** | -.114 | .009 | -.194 | -.120 | .080 | -.031 | -.115 | .021 | .053 | -.035 | -.019 | .002 | -.037 | .074 | -.105 | -.122 | .156 | 1.000 | -.211 | .444 | .038 | -.053 | 0.012 |
| **HEI** | .153 | .106 | .102 | .058 | -.084 | .090 | -.012 | .171 | .049 | .112 | .009 | .068 | -.136 | .111 | .073 | .114 | -.757 | -.211 | 1.000 | -.106 | -.048 | -.075 | .012 |
| **PNJ** | .126 | -.069 | .164 | .003 | -.129 | .070 | .084 | .080 | .086 | -.039 | .040 | -.048 | .031 | .020 | .050 | .205 | .141 | -.444 | -.106 | 1.000 | .248 | .272 | -.045 |
| **NJM** | .136 | -.012 | .150 | .137 | .081 | .086 | .254 | .047 | 276 | -.037 | .204 | .167 | -.029 | .052 | .009 | -.037 | .075 | .038 | -.048 | .248 | 1.000 | .098 | -.017 |
| **IJH** | .158 | .009 | .141 | -.059 | -.088 | .040 | .055 | -.020 | -.036 | .002 | -.051 | .034 | .205 | -.104 | .007 | -.027 | .002 | -.053 | -.075 | .272 | .098 | 1.000 | .161 |
| **ILE** | .126 | .115 | .109 | .042 | .083 | .239 | .003 | .128 | .068 | .159 | .104 | .103 | -.12 | .068 | .248 | -.096 | .040 | -002 | .012 | -.045 | -.017 | -.161 | 1.000 |

The clustering patterns of five Iranian and one Turkish populations were plotted as axis 1 vs. 2 (Figures 5-7).Figure 5 shows clustering of 5 honeybee populations of Iran. Populations are clustered into 4 groups and there is no important morphometric variation between Sarein and Viladereg. By discriminant function analysis (DFA), five Iranian populations and Hakkari population are clustered into two separate groups (Figures 6 and 7).

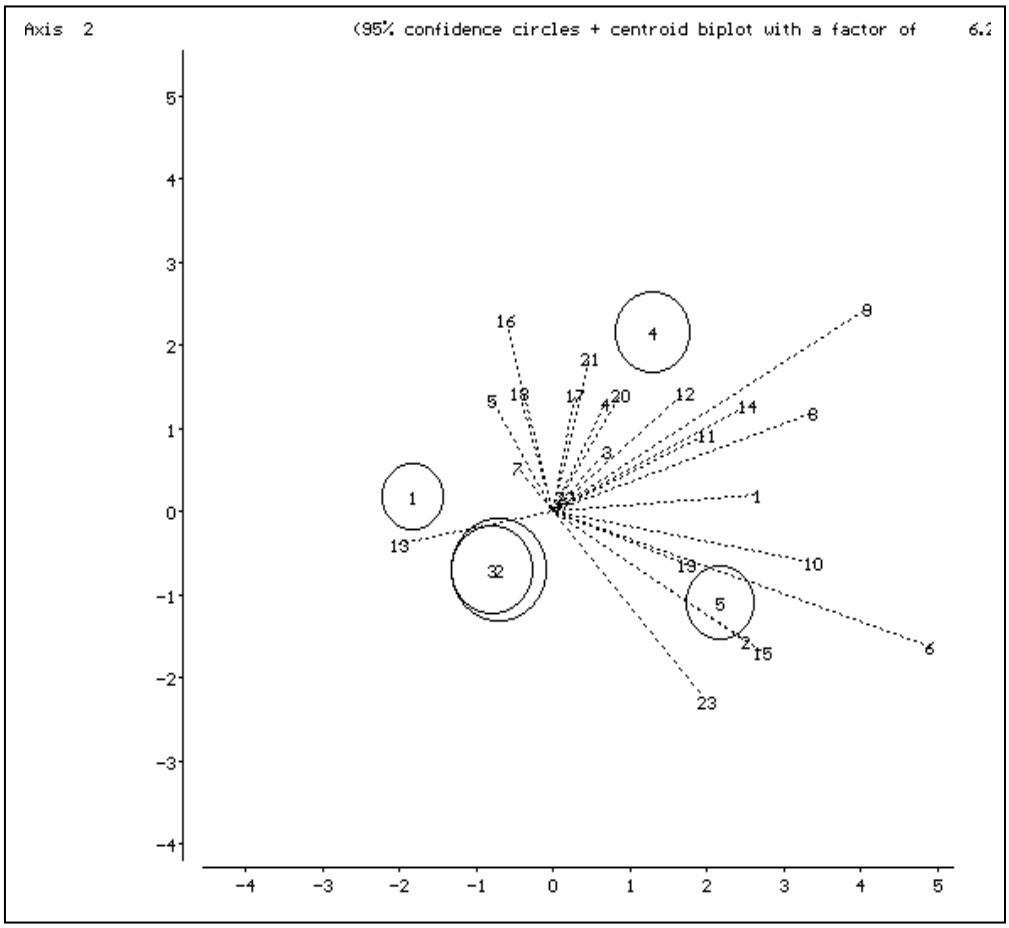


Figure 5: Clustering of five Iranian honeybee populations on the bases of morphometric analysis, axis 1 vs. 2.(1:Urmia, 2: Sarein, 3: Viladereg, 4:Amol, 5: Alborz).

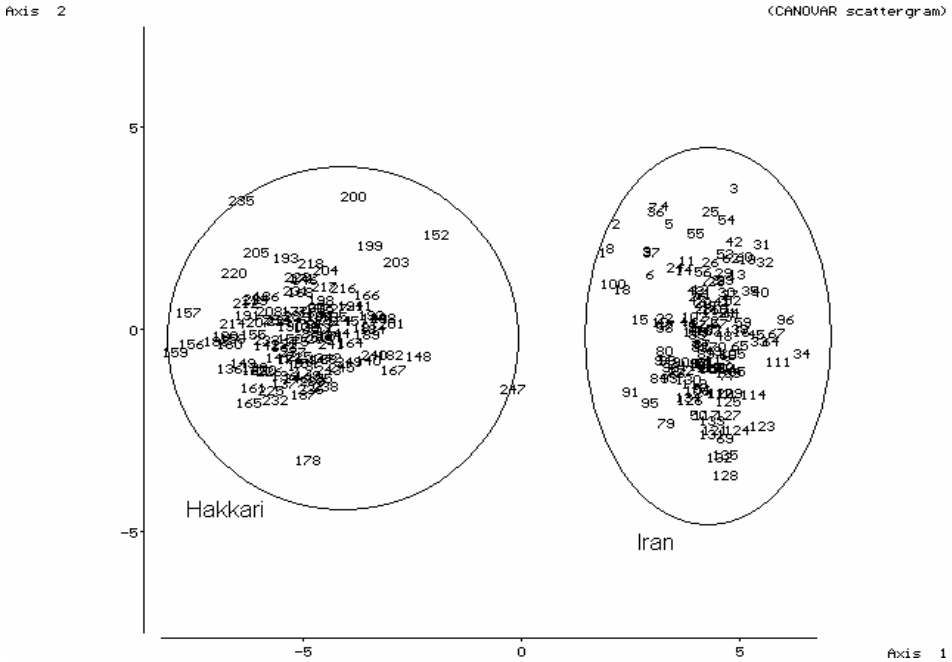


Figure 6: Clustering of honey bees on the basis of morphometric measurements analyzed by DFA, axis 1 vs. 2.

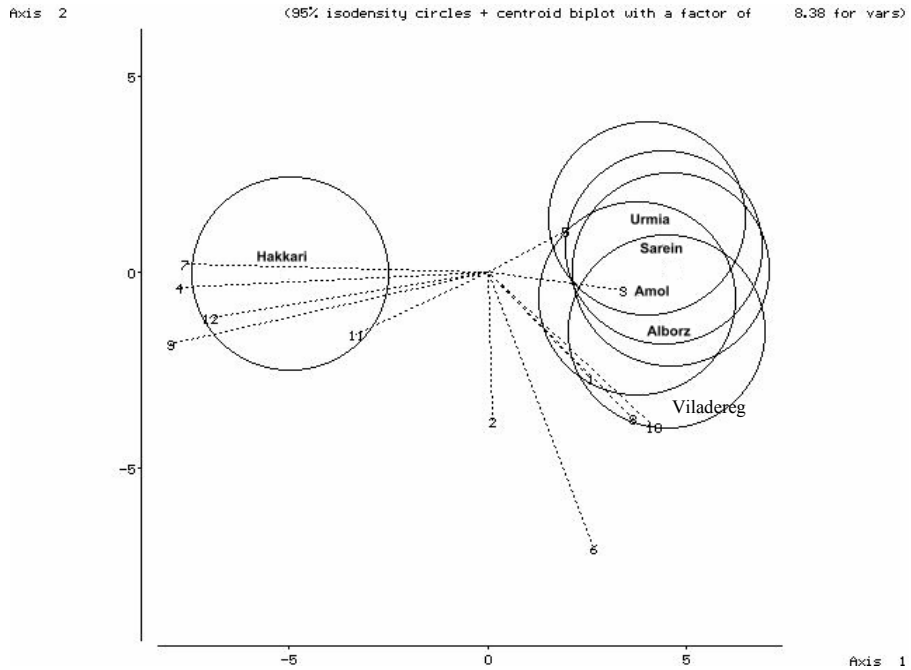


Figure 7: Clustering of Iranian and Turkish honey bee populations on the bases of morphometric analysis , axis 1 vs. 2.

For each population sum of 8 metric characters which are important in determining the body size of honeybee individuals were calculated. The largest value was obtained for Albroz population and Amol has the smallest value (Table 8). The correlation between average size of honeybee individuals and altitude of locations calculated as 0.82.

Table 8. Sum of mean values of 8 metric parameters measured in forewings and hindlegs

| | average of characters | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Location** | FL | L1 | L2 | FB | Fel | TiL | ML | MT | **SUM** |
| Urmia | 5,18 | 0,65 | 2,58 | 5,51 | 7,3 | 9,1 | 5,7 | 3,6 | **39,63** |
| Sarein | 2,06 | 0,65 | 2,58 | 5,51 | 7,2 | 9,3 | 5,7 | 3,5 | **36,57** |
| Viladereg | 2,06 | 0,65 | 2,58 | 5,51 | 7,3 | 9,2 | 5,7 | 3,5 | **36,48** |
| Amol | 2,06 | 0,65 | 2,58 | 5,51 | 7,2 | 9,1 | 5,7 | 3,5 | **36,37** |
| Alborz | 2,06 | 0,65 | 2,58 | 5,51 | 7,3 | 10,5 | 12,0 | 9,7 | **50,20** |
| Hakkari | 8,00 | 1,46 | 0,92 | 3,01 | 7,4 | 9,7 | 8,0 | 5,5 | **43,95** |

## Mitochondrial DNA Analysis:

Only one mitochondrial DNA haplotype was observed from *Dra*I RFLP analysis of four Iranian populations. *Dra*I digestion of PCR products of amplified $t_{RNA}$leu-COII intergenic region of the mtDNA showes only one pattern in four Iranian honeybee populations. 422 base pair bands were observed after running 10% acrylamide gel (Figure 8). Four shorter bands pattern obtained from electrophoresis and staining of 3,6% acrylamide gel.

64, 47, and 41 base pair bands which correspond to pattern C1 observed (Figure 9).

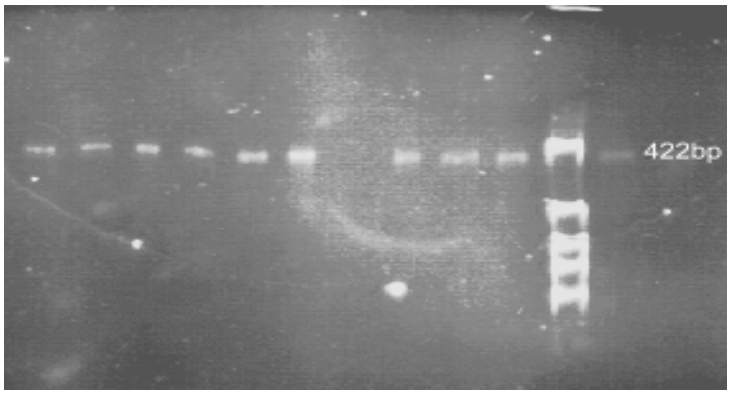


Figure 8: long fragments of mtDNA in 10% acrylamide gel

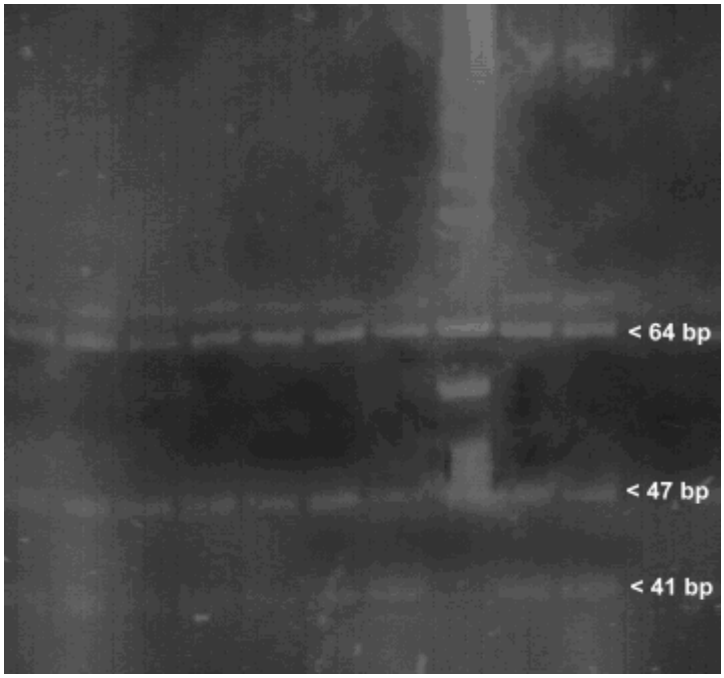


Figure 9: short fragments of mtDNA in 3,6% acrylamide gel

## 3.3. Microsatellite Analysis

### 3.3.1. DNA Extraction Results

The DNA extracts obtained from tissue samples were electrophoresed 1% agarose gel and stained by ethidium bromide to check for the existence of DNA in solution (Figure 10). Samples with DNA bands in agarose gel were used for later studies.

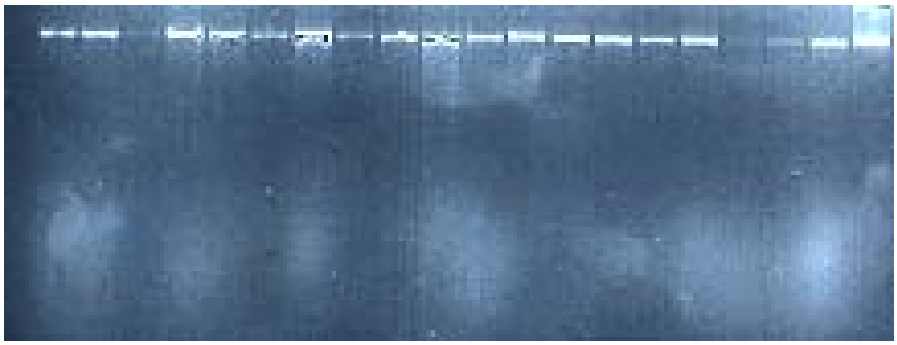

Figure 10. Agarose gel electrophoresis results of DNA extraction from samples. DNA bands and smears of RNA can be observed in the figure.

### 3.3.2. Allele Frequencies, Heterozygosities and Gene Diversities

Totally 95 honeybee workers from 4 populations of Iran were examined for microsatellite analysis. 5 loci were studied in these experiments and 24 alleles for locus A7, 9 alleles for locus 24, 5 alleles for locus A28, 10 alleles for locus A43 and 13 alleles for A113 were found (Figures 10, 11, 12, 13, and 14). Population specific alleles are given in table 9, and all loci are

polymorphic for 4 populations. Between populations, Amol with 16 alleles has the highest number of specific alleles. Allele 220 from locus A133 with 0.27 has the highest frequency. After Amol, Viladereg with 6 allele has the highest population specific allele number. Urmia with 4 and Sarein with 3 specific allele numbers are the next populations. Calculated allele frequencies, observed and expected heterozygosities for loci A7, A24, A28, A43 and A113 are given in tables 10, 11, 12, 13 and 14.

Samples collected from Amol show the highest polymorphism in locus A7 with 24 alleles, and in locus A7 Sarein population with 8 alleles has the lowest polymorphism . In locus A24 studies, Amol population with 8 alleles shows the greatest polymorphism and Urmia has the lowest polymorphism (3 alleles). Locus A28 has the greatest polymorphism in Urmia population with 4 alleles has the highest polymorphism. Amole and Viladereg populations show the lowest polymorphism (2 alleles) of locus A28 . Urmia population was found to have the biggest number of alleles (9) for A43 locus and Sarein population showed the smallest number of A43 alleles (3). The greatest polymorphism for A113 locus (10 alleles) was detected in Amol samples, while the lowest polymorphism (3 alleles) was found in Viladereg samples.

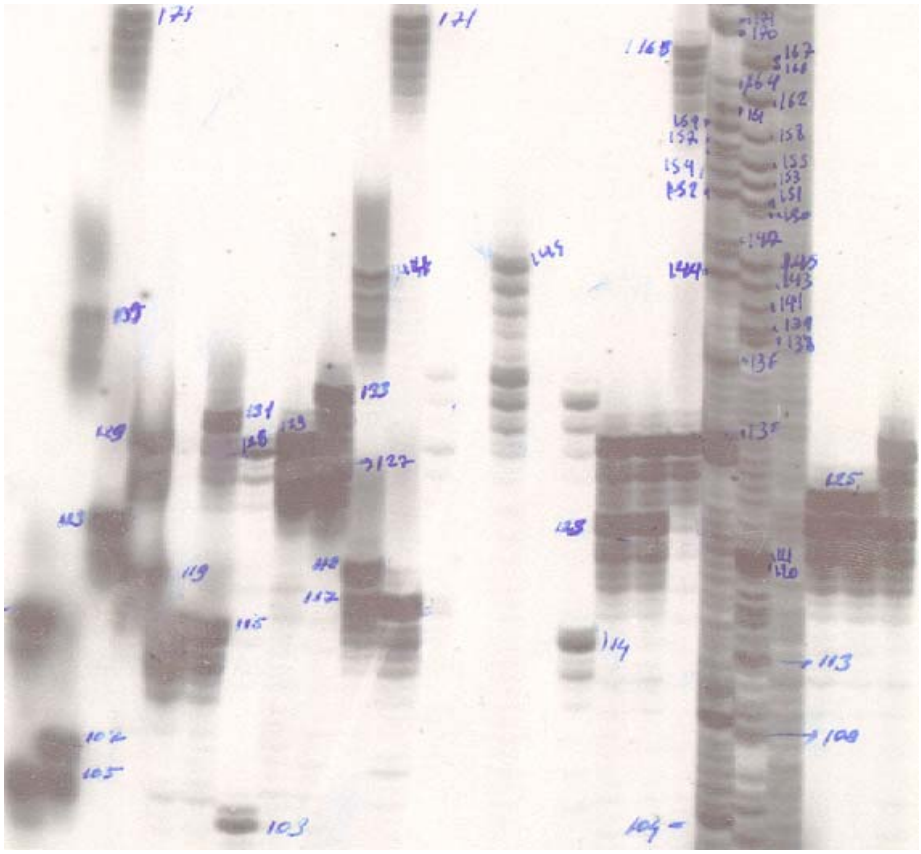

Figure 10. An autoradiogram representing A7 alleles

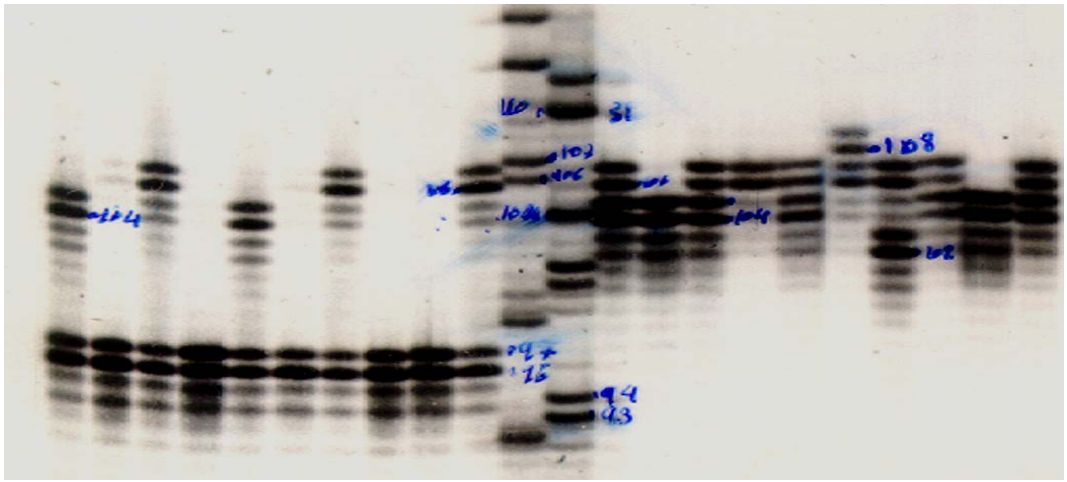

Figure 11. An autoradiogram representing A24 alleles

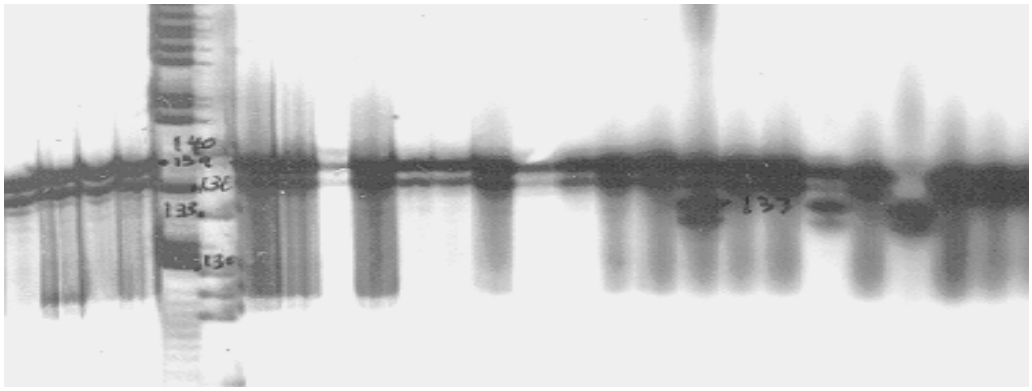

Figure 12. An autoradiogram representing A28 alleles

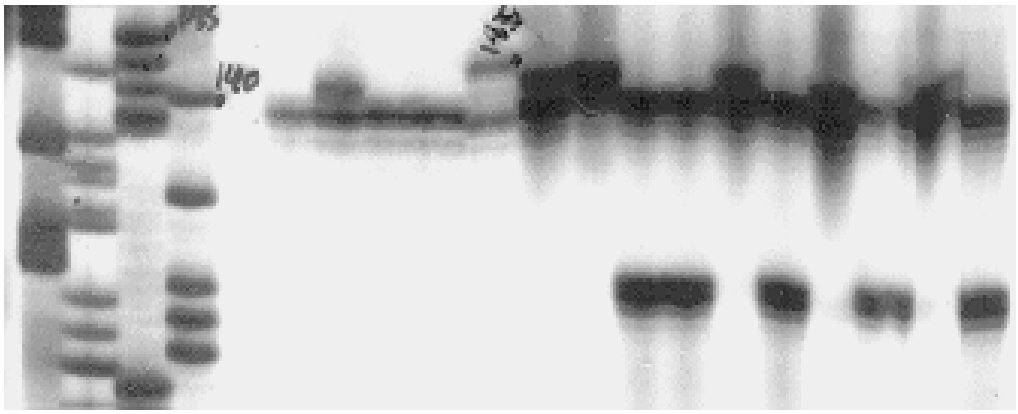

Figure 13. An autoradiogram representing A43 allele

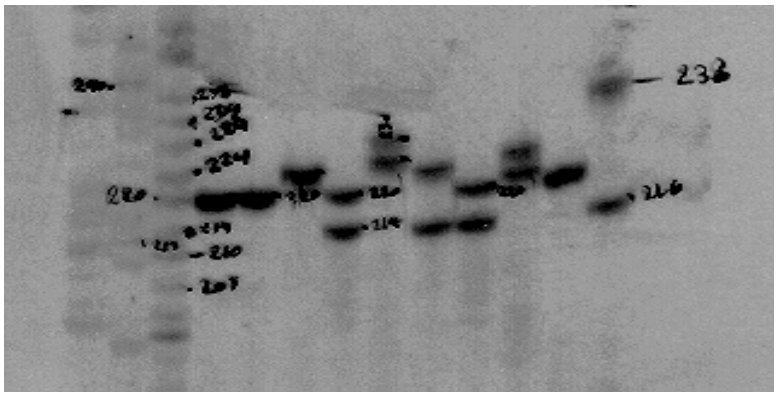

Figure 14. An autoradiogram representing A113 alleles

Table 9. Specific alleles observed in each population

| Locus | **Allele Size** | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|---|
| A28 | **123** | 0.04 | 0 | 0 | 0 |
| | **131** | 0 | 0 | 0.07 | 0 |
| A43 | **119** | 0 | 0 | 0.038 | 0 |
| | **137** | 0 | 0 | 0 | 0.083 |
| | **146** | 0.032 | 0 | 0 | 0 |
| A7 | **103** | 0 | 0 | 0.034 | 0 |
| | **105** | 0 | 0 | 0.033 | 0 |
| | **107** | 0 | 0 | 0.10 | 0 |
| | **109** | 0.083 | 0 | 0 | 0 |
| | **113** | 0 | 0.038 | 0 | 0 |
| | **115** | 0 | 0 | 0.20 | 0 |
| | **121** | 0 | 0 | 0 | 0.029 |
| | **144** | 0 | 0 | 0.033 | 0 |
| | **159** | 0 | 0.038 | 0 | 0 |
| | **171** | 0 | 0 | 0.067 | 0 |
| A24 | **92** | 0 | 0 | 0.10 | 0 |
| | **94** | 0 | 0 | 0.07 | 0 |
| | **98** | 0 | 0 | 0.03 | 0 |
| | **100** | 0 | 0.02 | 0 | 0 |
| | **102** | 0 | 0 | 0 | 0.042 |
| | **108** | 0 | 0 | 0 | 0.04 |
| A113 | **206** | 0 | 0 | 0.03 | 0 |
| | **212** | 0 | 0 | 0 | 0.04 |
| | **214** | 0 | 0 | 0.10 | 0 |
| | **218** | 0 | 0 | 0 | 0.08 |
| | **220** | 0 | 0 | 0.27 | 0 |
| | **232** | 0 | 0 | 0.03 | 0 |
| | **234** | 0 | 0 | 0.03 | 0 |

Table 10. Allele frequencies and heterozygosities for A7 locus (N: Number of individuals, $H_o$: Observed heterozygosity, $H_e$: Expected heterozygosity).

| Allele Size(bp) | Urmia | Sarein | Amol | Viladereg |
|---|---|---|---|---|
| **103** | 0 | 0 | 0.034 | 0 |
| **105** | 0 | 0 | 0.033 | 0 |
| **107** | 0 | 0 | 0.10 | 0 |
| **109** | 0.083 | 0 | 0 | 0 |
| **111** | 0.028 | 0 | 0.033 | 0.088 |
| **113** | 0 | 0.038 | 0 | 0 |
| **115** | 0 | 0 | 0.20 | 0 |
| **117** | 0 | 0 | 0.10 | 0.12 |
| **119** | 0 | 0 | 0.034 | 0.029 |
| **121** | 0 | 0 | 0 | 0.029 |
| **123** | 0.028 | 0.15 | 0.033 | 0.059 |
| **125** | 0 | 0.15 | 0 | 0.088 |
| **127** | 0.22 | 0.23 | 0.033 | 0.029 |
| **128** | 0.028 | 0 | 0.033 | 0 |
| **129** | 0.028 | 0.19 | 0.10 | 0.12 |
| **131** | 0.083 | 0.15 | 0.033 | 0.029 |
| **133** | 0.056 | 0 | 0.033 | 0.059 |
| **135** | 0.11 | 0 | 0 | 0.24 |
| **137** | 0.083 | 0.038 | 0 | 0.029 |
| **139** | 0.23 | 0 | 0.033 | 0.029 |
| **141** | 0.056 | 0 | 0.068 | 0.059 |
| **144** | 0 | 0 | 0.033 | 0 |
| **159** | 0 | 0.038 | 0 | 0 |
| **171** | 0 | 0 | 0.067 | 0 |
| | N=18 | N=13 | N=15 | N=17 |
| $H_o$ | 0.944 | 0 | 0.600 | 0.941 |
| $H_e$ | 0.0556 | 1.000 | 0.400 | 0.588 |

Table 11. Allele frequencies and heterozygosities for A24 locus (N: Number of individuals, $H_o$: Observed heterozygosity, $H_e$: Expected heterozygosity).

| Allele Size(bp) | Urmia | Sarein | Amol | Viladereg |
|---|---|---|---|---|
| **92** | 0 | 0 | 0.10 | 0 |
| **94** | 0 | 0 | 0.07 | 0 |
| **96** | 0.10 | 0.48 | 0.20 | 0.13 |
| **98** | 0 | 0 | 0.03 | 0 |
| **100** | 0 | 0.02 | 0 | 0 |
| **102** | 0 | 0 | 0 | 0.04 |
| **104** | 0.42 | 0.14 | 0.40 | 0.37 |
| **106** | 0.48 | 0.36 | 0.20 | 0.42 |
| **108** | 0 | 0 | 0 | 0.04 |
| | N=26 | N=29 | N=15 | N=12 |
| $H_o$ | 0.538 | 0.552 | 0.600 | 0.667 |
| He | 0.592 | 0.627 | 0.770 | 0.696 |

Table 12. Allele frequencies and heterozygosities for A28 locus (N: Number of individuals $H_o$: Observed heterozygosity, $H_e$: Expected heterozygosity).

| Allele Size(bp) | Urmia | Sarein | Amol | Viladereg |
|---|---|---|---|---|
| **123** | 0.04 | 0 | 0 | 0 |
| **131** | 0 | 0 | 0.07 | 0 |
| **133** | 0.02 | 0.06 | 0 | 0 |
| **137** | 0.24 | 0.10 | 0 | 0.11 |
| **138** | 0.70 | 0.84 | 0.93 | 0.89 |
| | N=25 | N=34 | N=15 | N=9 |
| $H_o$ | 0.60 | 0.265 | 0.133 | 0 |
| $H_e$ | 0.459 | 0.287 | 0.129 | 0.209 |

Table 13. Allele frequencies and heterozygosities for A43 locus ($H_o$: Observed heterozygosity, $H_e$: Expected heterozygosity, N: Number of individuals).

| Allele Size(bp) | Urmia | Sarein | Amol | Viladereg |
|---|---|---|---|---|
| **119** | 0 | 0 | 0,038 | 0 |
| **127** | 0,107 | 0 | 0,077 | 0,042 |
| **137** | 0 | 0 | 0 | 0,083 |
| **139** | 0,071 | 0,567 | 0,424 | 0,042 |
| **140** | 0,214 | 0 | 0 | 0,333 |
| **141** | 0,107 | 0,367 | 0,269 | 0,042 |
| **142** | 0,357 | 0 | 0 | 0,25 |
| **143** | 0,036 | 0,066 | 0,192 | 0 |
| **144** | 0,072 | 0 | 0 | 0,208 |
| **146** | 0,036 | 0 | 0 | 0 |
| | N=28 | N=15 | N=13 | N=12 |
| $H_o$ | 0,4026 | 0,2793 | 0,366 | 0,402 |
| $H_e$ | 0,4285 | 0,367 | 0,346 | 0,416 |

Table 14. Allele frequencies and heterozygosities for A113 locus (N: Number of individuals, $H_o$: Observed heterozygosity, $H_e$: Expected heterozygosity).

| Allele Size(bp) | Urmia | Sarein | Amol | Viladereg |
|---|---|---|---|---|
| **206** | 0 | 0 | 0.03 | 0 |
| **212** | 0 | 0 | 0 | 0.04 |
| **214** | 0 | 0 | 0.10 | 0 |
| **216** | 0.020 | 0 | 0.03 | 0.08 |
| **218** | 0 | 0 | 0 | 0.08 |
| **220** | 0 | 0 | 0.27 | 0 |
| **222** | 0.10 | 0 | 0.10 | 0 |
| **224** | 0.25 | 0.15 | 0.17 | 0.26 |
| **226** | 0.33 | 0.43 | 0.17 | 0.31 |
| **228** | 0.20 | 0.29 | 0 | 0.23 |
| **230** | 0.10 | 0.13 | 0.07 | 0 |
| **232** | 0 | 0 | 0.03 | 0 |
| **234** | 0 | 0 | 0.03 | 0 |
| | N=20 | N=24 | N=15 | N=13 |
| $H_o$ | 0.60 | 0.417 | 0.667 | 0.462 |
| $H_e$ | 0.791 | 0.70 | 0.873 | 0.797 |

Average heterozygosities for each population (Table 15), gene diversities ($H_T$) for each locus and all the loci are given in table 16. Sarein samples have the lowest average heterozygosity, 0.6114 and the lowest total number of alleles for five loci; 22 alleles out of 61 detected. The greatest value for gene diversity ($H_T$) measure is found on A7 locus with 0.933. Allele number for each locus and percentages of polymorphic loci in populations are given in table 17.

### 3.3.3. Hardy-Weinberg Tests

Table 18 shows the P values and standard errors in each population for each locus. Out of 20 cases 5 (A43 locus in Urmia population, A113 locus in Sarein population, A7 and A133 loci in Amol population and A113 locus in Viladereg) significant deviations are detected at 0.05 level.

### 3.3.4. Linkage Disequilibrium Tests

The exact P values calculated for each pair of loci in each population and significant linkage disequilibrium detected at 0.05 significance level are given in table 19. All five loci showed significant pairwise linkage disequilibrium in one of four populations.

Table 15. Average heterozygosities for all five loci

| population | Average Heterozygosity | Standard error |
|---|---|---|
| **Urmia** | 0.708930 | 0.077226 |
| **Sarein** | 0.611495 | 0.096213 |
| **Amol** | 0.694913 | 0.144365 |
| **Viladereg** | 0.693247 | 0.120971 |

Table 16. Gene diversities ($H_T$) for each loci and for all loci.

| Locus | Gene Diversity($H_T$) |
|---|---|
| **A28** | 0.284562 |
| **A43** | 0.829485 |
| **A7** | 0.931529 |
| **A24** | 0.703250 |
| **A113** | 0.813262 |
| **All** | 0.688558 |

Table 17. Allele number for each locus and percents of polymorphic loci in populations.

| **locus** | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **A7** | 12 | 8 | 17 | 14 |
| **A 24** | 3 | 4 | 6 | 5 |
| **A28** | 4 | 3 | 2 | 2 |
| **A43** | 8 | 3 | 5 | 7 |
| **A113** | 6 | 4 | 10 | 6 |
| Total No. Of alleles | 33 | 22 | 40 | 34 |
| No. of allels per locus | 6,6± 1.9 | 4,4± 0,8 | 8,0± 1.2 | 6,8± 1.2 |
| Polymorphic loci (0.99 criterion) (%) | 100 | 100 | 100 | 100 |

Table 18. Results of Hardy-Weinberg exact tests, the P values and standard errors in each population for each locus.

| Population | Locus | P value | St.error | W&C ($F_{IS}$) |
|---|---|---|---|---|
| **Urmia** | A24 | 0.5287 | 0.0000 | +0.092 |
| **Urmia** | A28 | 0.3626 | 0.0000 | -0.314 |
| **Urmia** | A113 | 0.0716 | 0.0054 | +0.246 |
| **Urmia** | A43 | 0.0000 | 0.0000 | -0.049 |
| **Urmia** | A7 | 0.7457 | 0.0218 | -0.057 |
| **Sarein** | A24 | 0.5958 | 0.0000 | +0.122 |
| **Sarein** | A28 | 0.1608 | 0.0000 | +0.080 |
| **Sarein** | A113 | 0.0344 | 0.0000 | +0.411 |
| **Sarein** | A43 | 0.1364 | 0.0000 | -0.340 |
| **Sarein** | A7 | 0.3446 | 0.0174 | -0.160 |
| **Amol** | A24 | 0.1238 | 0.0089 | +0.227 |
| **Amol** | A28 | 1.000 | 0.0000 | +0.080 |
| **Amol** | A113 | 0.0258 | 0.0074 | +0.243 |
| **Amol** | A43 | 0.2770 | 0.0000 | -0.113 |
| **Amol** | A7 | 0.0000 | 0.000 | +0.370 |
| **Viladereg** | A24 | 0.3307 | 0.0098 | +0.043 |
| **Viladereg** | A28 | 0.0588 | 0.0000 | 1.000 |
| **Viladereg** | A113 | 0.0080 | 0.0017 | +0.431 |
| **Viladereg** | A43 | 0.0576 | 0.0084 | -0.038 |
| **Viladereg** | A7 | 0.6200 | 0.0322 | -0.032 |

Table 19. Genotypic disequilibrium for each pair of loci in each population

| Population | Locus | Locus | P-Value | S.E |
|---|---|---|---|---|
| **Urmia** | A24 | A28 | 0.87375 | 0.00525 |
| **Urmia** | A24 | A133 | 0.10763 | 0.00624 |
| **Urmia** | A28 | A133 | 0.85113 | 0.00803 |
| **Urmia** | A24 | A43 | 0.22172 | 0.01241 |
| **Urmia** | A28 | A43 | 0.28296 | 0.01464 |
| **Urmia** | A133 | A43 | 0.17689 | 0.02116 |
| **Urmia** | A24 | A7 | 1,00000 | 0.00000 |
| **Urmia** | A28 | A7 | 0.39049 | 0.02494 |
| **Urmia** | A133 | A7 | 1,00000 | 0.00000 |
| **Urmia** | A43 | A7 | 1,00000 | 0.00000 |
| **Sarein** | A24 | A28 | 0.50178 | 0.00977 |
| **Sarein** | A24 | A133 | 1,00000 | 0.00000 |
| **Sarein** | A28 | A133 | 0.82335 | 0.00423 |
| **Sarein** | A24 | A43 | 0.84254 | 0.00583 |
| **Sarein** | A28 | A43 | 0.00111 | 0.00003 |
| **Sarein** | A133 | A43 | 1,00000 | 0.00000 |
| **Sarein** | A133 | A7 | 0.47450 | 0.01896 |
| **Amol** | A24 | A28 | 0.91409 | 0.00421 |
| **Amol** | A24 | A133 | 1,00000 | 0.00000 |
| **Amol** | A28 | A133 | 0.44411 | 0.00998 |
| **Amol** | A24 | A43 | 0.80047 | 0.01463 |
| **Amol** | A28 | A43 | 0.76601 | 0.00432 |
| **Amol** | A133 | A43 | 0.70241 | 0.01928 |
| **Amol** | A24 | A7 | 1,00000 | 0.00000 |
| **Amol** | A28 | A7 | 0.76320 | 0.00885 |
| **Amol** | A133 | A7 | 1,00000 | 0.00000 |
| **Amol** | A43 | A7 | 1,00000 | 0.00000 |
| **Viladereg** | A24 | A28 | 0.55623 | 0.00756 |
| **Viladereg** | A133 | A43 | 1,00000 | 0.00000 |
| **Viladereg** | A133 | A7 | 1,00000 | 0.00000 |
| **Viladereg** | A43 | A7 | 0.41146 | 0.02340 |

### 3.3.5. Population Differentiation

Results for the tests of population differentiation are divided into two; one being related with genic differentiation the other with the genotypic differentiation.

#### 3.3.5.1. Genic Differentiation

Tables 20.a -20.e show the P values and standard errors of genic differentiation for each pair of populations for each locus. With a P value of zero, differentiation for each locus is found to be highly significant.

Table 20.a. Population pairwise results for genic differentiation for locus A7

| Locus | Populations | Probability | S.E. |
|---|---|---|---|
| A7 | Urmia & Sarein | 0.00022 | 0.00012 |
| A7 | Urmia & Amol | 0.00004 | 0.00004 |
| A7 | Urmia & Viladereg | 0.00128 | 0.00030 |
| A7 | Sarein & Amol | 0.00000 | 0.00000 |
| A7 | Sarein & Viladereg | 0.00211 | 0.00068 |
| A7 | Amol & Viladereg | 0.00621 | 0.00126 |

Table 20.b. Population pairwise results for genic differentiation for locus A24

| Locus | Populations | Probability | S.E. |
|---|---|---|---|
| A24 | Urmia & Sarein | 0.00000 | 0.00000 |
| A24 | Urmia & Amol | 0.00259 | 0.00067 |
| A24 | Urmia & Viladereg | 0.33767 | 0.00681 |
| A24 | Sarein & Amol | 0.00003 | 0.00003 |
| A24 | Sarein & Viladereg | 0.00264 | 0.00076 |
| A24 | Amol & Viladereg | 0.16157 | 0.00659 |

Table 20.c. Population pairwise results for genic differentiation for locus A28

| Locus | Populations | Probability | S.E. |
|---|---|---|---|
| A28 | Urmia & Sarein | 0.04334 | 0.00257 |
| A28 | Urmia & Amol | 0.00127 | 0.00044 |
| A28 | Urmia & Viladereg | 0.57877 | 0.00495 |
| A28 | Sarein & Amol | 0.01698 | 0.00133 |
| A28 | Sarein & Viladereg | 0.86329 | 0.00169 |
| A28 | Amol & Viladereg | 0.13115 | 0.00198 |

Table 20 d. Population pairwise results for genic differentiation for locus A43

| Locus | Populations | Probability | S.E. |
|---|---|---|---|
| A43 | Urmia & Sarein | 0.00000 | 0.00000 |
| A43 | Urmia & Amol | 0.00000 | 0.00000 |
| A43 | Urmia & Viladereg | 0.17891 | 0.00635 |
| A43 | Sarein & Amol | 0.16905 | 0.00500 |
| A43 | Sarein & Viladereg | 0.00000 | 0.00000 |
| A43 | Amol & Viladereg | 0.00000 | 0.00000 |

Table 20.e. Population pairwise results for genic differentiation for locus A113

| Locus | Populations | Probability | S.E. |
|---|---|---|---|
| A113 | Urmia & Sarein | 0.09862 | 0.00459 |
| A113 | Urmia & Amol | 0.00007 | 0.00005 |
| A113 | Urmia & Viladereg | 0.14496 | 0.00548 |
| A113 | Sarein & Amol | 0.00000 | 0.00000 |
| A113 | Sarein & Viladereg | 0.01238 | 0.00145 |
| A113 | Amol & Viladereg | 0.00000 | 0.00000 |

### 3.3.6. Population comparison and structure analysis by F coefficients

The $F_{ST}$ indicates the power of the locus in differentiating the populations. $F_{ST}$ values are significantly different from zero in pair wise comparison of all four populations. $F_{ST}$ values obtained for all populations and are changing between 0.0341 and 0.0870.Table 21 representing significant $F_{ST}$ P values at 0.05 significance level is also provided. The $F_{IS}$ represents the reduction of the number of heterozygotes as a result of non-random mating. The negative value shows excess of heterozygotes and the positive values indicates excess of homozygotes. The $F_{IS}$ values are between -0.0642 and 0.3476 and the mean value is 0.1112. The $F_{IT}$ value which represents reduction of heterozygosity varied from -0.0044 to 0.3699 with an average 0.159 being significantly different from zero. For locus A43 there is deviation in favor of heterozygotes however for other four loci there is a significant deviation in favor of homozygote individuals. $F_{ST}$ measures for each pair of population are given in the form of triangular matrices in tables 22 to 26.

Nei's standard genetic distances (1972) for four Iranian populations using five microsatellite loci and with the addition of microsatellite analysis of one Turkish population were calculated and matrices are given for distances and their standard errors (Tables 27 – 31).

Table21: F coefficeints for the total populations for each loci and all loci.

| **Locus** | **$F_{ST}$** | **$F_{IT}$** | **$F_{IS}$** |
|---|---|---|---|
| **A24** | 0.0870 | 0.1996 | 0.1233 |
| **A28** | 0.0439 | 0.0515 | 0.0079 |
| **A133** | 0.0341 | 0.3699 | 0.3476 |
| **A43** | 0.0562 | -0.0044 | - 0.0642 |
| **A7** | 0.0444 | 0.0990 | 0.0574 |
| **All:** | 0.0538 | 0.1590 | 0.1112 |

Table 22: $F_{ST}$ distance measure of Iranian honeybee populations for locus A7

| Population | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | 0.0480 | 0 | | |
| **Amol** | 0.0520 | 0.0559 | 0 | |
| **Viladereg** | 0.0425 | 0.0578 | 0.0318 | 0 |

Table23 : $F_{ST}$ distance measure of Iranian honeybee populations for locus A24

| Population | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | 0.1526 | 0 | | |
| **Amol** | 0.0476 | 0,1001 | 0 | |
| **Viladereg** | -0,025 | 0.1030 | 0,0055 | 0 |

Table 24: $F_{ST}$ distance measure of Iranian honeybee populations for locus A28

| Population | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | 0.0403 | 0 | | |
| **Amol** | 0.1363 | 0.0309 | 0 | |
| **Viladereg** | 0.0345 | -0.0333 | -0.0081 | 0 |

Table 25: $F_{ST}$ distance measure of Iranian honeybee populations for locus A43.

| Population | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | 0.1157 | 0 | | |
| **Amol** | 0.0656 | -0.0049 | 0 | |
| **Viladereg** | 0.0111 | 0.0736 | 0.0426 | 0 |

Table 26: $F_{ST}$ distance measure of Iranian honeybee populations for locus A113.

| Population | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | -0.0017 | 0 | | |
| **Amol** | 0.0552 | 0.1153 | 0 | |
| **Viladereg** | -0.0226 | 0.0015 | 0.0614 | 0 |

Table 27. Nei's standard genetic distance matrix of four Iran's honeybee populations for five microsatellite loci; A7, A24, A28, A43 and A113.

| Population | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | 0.2566 | 0 | | |
| **Amol** | 0.2570 | 0.1342 | 0 | |
| **Viladereg** | 0.0152 | 0.2458 | 0.2037 | 0 |

Table 28. Standard errors of distances given in table 25.

| Population | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | 0.1821 | 0 | | |
| **Amol** | 0.1613 | 0.0911 | 0 | |
| **Viladereg** | 0.0474 | 0.2182 | 0.2053 | 0 |

Table 29. Nei's standard genetic distance matrix of three Iran's honeybee populations and one Turkey's population (Hakkari) for five microsatellite loci; A7, A24, A28, A43 and A113.

| Population | **Urmia** | **Sarein** | **Amol** | **Hakkari** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | 0.1122 | 0 | | |
| **Amol** | 0.2663 | 0.1313 | 0 | |
| **Hakkari** | 0.1036 | 0.1223 | 0.2302 | 0 |

Table 30. Standard errors of distances given in table 27.

| Population | **Urmia** | **Sarein** | **Amol** | **Hakkari** |
|---|---|---|---|---|
| **Urmia** | 0 | | | |
| **Sarein** | 0.0668 | 0 | | |
| **Amol** | 0.1592 | 0.0927 | 0 | |
| **Hakkari** | 0.0468 | 0.1060 | 0.2480 | 0 |

.

Table 31: Pairwise $F_{ST}$ for population pairs.

| | **Urmia** | **Sarein** | **Amol** | **Viladereg** |
|---|---|---|---|---|
| **Urmia** | 0.0000 | | | |
| **Sarein** | 0.1088 | 0.0000 | | |
| **Amol** | 0.0883 | 0.0702 | 0.0000 | |
| **Viladereg** | 0.0086 | 0.1122 | 0.0771 | 0.0000 |

### 3.3.7. Tree Construction

By using genetic distances given in tables 20 and 21 two tree was constracted according to Neighbour joining method of Saitou and Nei (1987). In the first phenogram a tree was drawn with four Iran populations for five microsatellite loci (Figure 15). In the second phenogram a tree was constructed by adding microsatellite data of one Turkish population to data obtained from three Iranian populations (Figure 16).

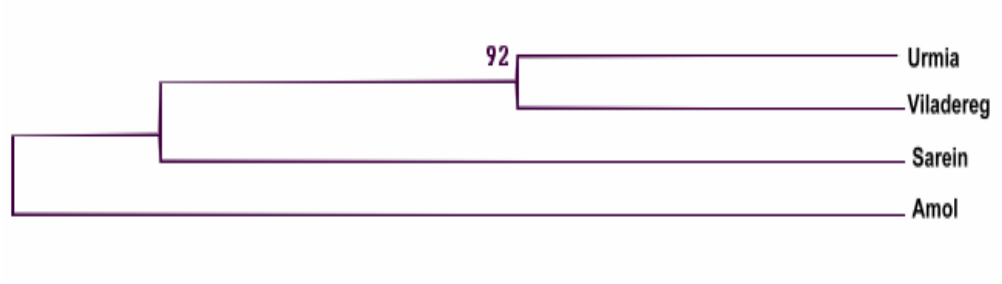


Figure15. UPGMA tree of four honeybee populations of Iran.

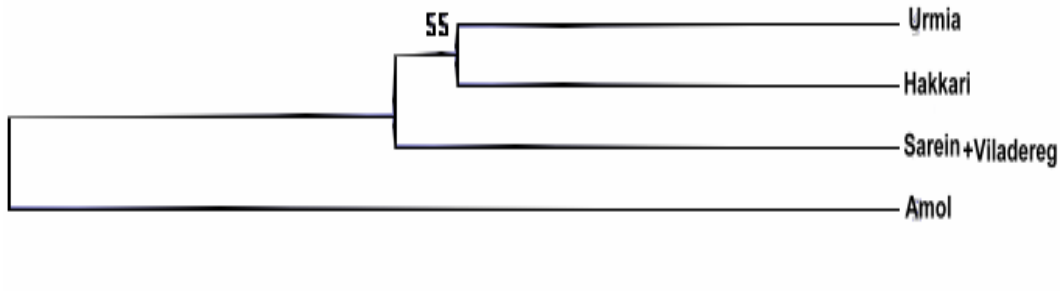


Figure16. UPGMA tree of three honeybee populations of Iran and one population of Turkey.

# CHAPTER 4

# DISCUSSION

In this study morphometric measurements, mitochondrial DNA analysis and 5 microsatellite loci were used to investigate the genetic variation in honeybee populations from north and North West of Iran and comparing it with eastern Anatolian population (Hakkari).

Analysis of 23 morphometric parameters of the honeybees of five Iranian and one Turkish population was done by PCA and DFA methods. In considering five Iranian populations according to PCA four distinct groups were obtained. Although Urmia, Amol and Alborz were clustered in 4 separate groups, Sarein and Viladereg were grouped closer and distinctly from other Iranian populations. After adding morphometric data from Turkish population (Hakkari) to five Iranian populations, two distinct clusters obtained. Hakkari samples were clustered separately from Urmia, Amol, Sarein, Viladereg and Alborz. Iranian populations again further divided in five closer groups.

According to analysis of correlation matrix, ‘d’ and FL has the highest correlation value. FL and L2, c and FB, TiL and FeL , d and TiL also are variables that have high positive correlation.The highest negative correlation is between HEI and OKF variables. Multigroup discrimenat analysis showed that FL, BDG and d variables have the highest discriminat weight among other characters.

Statistical comparisons of high-altitude bees from the Nyika Plateau with others from Mt. Kilimanjaro (Tanzania), Mt. Kenya (Kenya) and the Drakensberg mountains revealed all of these bees to be larger in size than their lower-altitude counterparts (Radloff *et al.* 2000). Average of 8 metric characters which are more affected by the size of honeybee calculated. Sum of the means were obtained for each population. According the data Amol samples have smallest and Alborz samples have largest individuals. There is a positive correlation between size and altitude, that is, bees become increasingly larger with increasing altitude. Amol samples were collected from a location that is almost at see level. Alborz with an altitude around 2500 meter has the highest altitude among other locations. This correlation between size of individuals and altitude of locations demostrates effect of biogeographical conditions in morphological similarities of honeybees.

Mitochondrial DNA studies were performed on four Iranian honeybee populations. Restriction and sequence data obtained from these studies, showed the existence of only one haplotype in populations. The *Dra*I restriction exhibit 4 bands with sizes as 422, 64, 47, and 41 base pair which corresponds to C1 pattern. This pattern is similar to the results reported by Garnery 1993. C1 pattern also was observed in Turkish honeybees especially in populations of eastern and northeastern Turkey. The data obtained from mtDNA analysis did not suggest apparent distinction between *A. m. meda* and other Anatolian populations (Palmer *et al.* 2000).

The reason for homogeneity of mtDNA lineage C can be due to the short size and absence of length variability in COI-COII region. This reduces potential mtDNA lineage C for pattern variability (Garnery *et al.*1993*)*. The importing

of foreign queens by professional bee-keepers or experimental apairs may be the other reason for decreasing of variability in COI-COII region.

In microsatellite studies, loci A7, A24, A28, A43 and A113 respectively had 24, 9, 5, 10 and 13 alleles. Although samples of Alborz population did not work in microsatellite studies, polymorphism found in five microsatellite loci of other Iranian honeybee populations is higher than that Estoup et al. (1995) reported by studying nine populations from two evolutionary honeybee lineage C and M.

Average heterozygosities are also high between Iranian populations. Urmia with 0.701 and Amol with 0.695 have the highest average heterozygosity. This value is 0.693 for Viladereg and 0.611 for Sarein. Average heterozygosity for lineage C samples between 0.464 and 0.612, for linage M samples was calculated between 0.291 and 0.410, and for lineage A between 0.788 and 0.872 (Estoup *et al*. 1995).For A7 locus total value of heterozygosity in all populations was detected as 0,932 and A7 locus has the highest gene diversity among all loci. A28 locus with 0.285 has the lowest gene diversity between 5 loci.

Totally 32 alleles were detected that not reported in C lineage populations by Estoup *et al*. (1995) .In locus A7,18 alleles 103, 105, 109, 117, 119, 121, 125, 127, 129, 131, 133, 137,139, 141, 144, 159, 171 were found in Iranian populations that are not reported in lineage C .Also, Alleles 92, 94, 96, 100, 102 in locus A24, alleles 123, 131, 137 in locus A28, alleles 119,137, 142, 144, 146 in locus A43 and alleles 206, 212, 216, 218, 222, 232 were not reported in lineage C populations by Estoup *et al.*(1995). There are 12 alleles namely, in locus A7 alleles 114, 116, 118, 120, 122 in locus A24 allele 110, in locus A28 alleles 132, 141 in locus A43 alleles 128, 130 and in locus A113

alleles 202, 236 were not observed in Iranian honeybee samples but reported by Estoup *et al.*(1995) for C lineage. In this study we detected alleles not reported in the three loci investigated in Anatolian honeybee populations (Bodur 2001). These are alleles 103, 107, 109, 111, 113, 128 for locus A7, alleles 92, 94 for locus A24 and allele 206 for locus A113.

Amol with a total allele number of 40 (17 for locus A7, 6 for locus A24, 2 for A28, 5 for A43 and 10 for A113), has the highest number of alleles and located in south of Caspian Sea. Total allele number of Urmia, Sarein and Viladereg respectively are 33, 22 and 34. This difference in polymorphism also observed in average heterozygosity levels of populations which changes between 0.611 and 0.709. This value is more closer to average heterozygosity level of Anatolian samples than the average heterozygosity of C lineage. Average heterozygosity of Anatolian populations reported as between 0.784 and 0.817 (Bodur 2001), this value is between 0.464 and 0.612 for lineage C (Estoup *et al*. 1995). In all four populations, specific alleles were found. Highest number of specific alleles (16) were found in Amol samples. 6 alleles for locus A7 (103, 105, 107, 115,144, 171), 3 alleles for locus A24 (92, 94, 98), 5 alleles for locus A113 (206, 214, 220, 232, 234), and 1 allele for each of locus A28 (131), and locus A43 (119) loci. Specific alleles 115 (locus A7) and 220 (A113) with fractions of 0.20 and 0.27 have very high frequency. For each of Urmia, Sarein and Viladereg populations 3 specific allele and in Viladereg 6 specific alleles were detected: in Viladereg 137 for locus A137, 121 for locus A7, 102 and 108 for locus A24, 212 and 218 for locus A113, in Urmia 123 for locus A28, 146 for A43, and 109 for locus A7, in Sarein 113 and 159 for locus A7, and 100 for locus A24.

Five significant deviations from Hardy Weinberg equilibrium were detected out of 20 population-locus combinations at 0.05 significance level. For all

five loci combinations only one significant deviation from equilibrium are detected, between A28 and A43 loci in Sarein population. This means 3.2% of combinations are in linkage disequilibrium. With a significant *P* value ($P < 0.05$), such a proportion can be considered as insignificant since it is <0.05. So we can say that genotypes at one locus are independent from genotypes at other locus and no interference is expected to population genetic analysis coming from any possible linkage between the studied loci. The $F_{ST}$ indicates the power of the locus in differentiating the populations. $F_{ST}$ values are significantly different from zero in pair wise comparison of all four populations.

$F_{ST}$ values obtained for all populations and are changing between 0.0341 and 0.0870. The mean value for $F_{ST}$ (0.0538) was significantly different from zero. A24 locus ($F_{ST}$= 0.0870) has the highest $F_{ST}$ value and provides the best discrimination between Iranian populations. The least discrimination obtained with A133 locus ($F_{ST}$ = 0.0341). Pairwise $F_{ST}$ values between populations shows that, except Viladereg and Urmia populations others are significant at 0.05 level. These results strengthen the idea about importance of this region as gene center for other populations (Ruttner 1988; Kandemir *et al.* 2000). The $F_{IS}$ represents the reduction of the number of heterozygotes as a result of random mating. The negative value shows excess of heterozygotes and the positive value indicates excess of homozygotes. The $F_{IS}$ values are between -0.0642 and 0.3476 and the mean value is 0.1112. The locus A43 has the lowest value which represents excess of heterozygotes. The $F_{IT}$ value which represents reduction of heterozygosity varied from -0.0044 to 0.3699 with an average 0.159 being significantly different from zero. For locus A43 there is deviation in favor of heterozygotes however for other four loci there is a significant deviation in favor of homozygotes individuals.

P values and standard errors obtained by G-like test for each test and for each population pairs show there is a highly significant genic differentiation for each five microsatellite loci. Highest amount of genic differentiation between pairs of populations was detected for A7 locus. All the six matches are highly significant with P values <0.05. The lowest genic differentiation was observed in locus A28.

In the phenogram that was constructed by using microsatellite analysis for Iranian honeybee populations, Amol remains distinctly separate from others. After adding microsatellite data obtained form Hakkari samples, Urmia and Hakkari are clustered together. Amol again is clustered independently from other populations. Honeybee populations of Iran have not been studied extensively for their lineage identification. From microsatellite analysis we can conclude at least north western honeybee populations of Iran are similar to Hakkari population located in eastern Anatolia. MtDNA studies were used by Smith *et al.* (1997) and Palmer *et al.* (2000) to determine lineages in honeybees of Turkey. By considering mtDNA haplotypes, they decided Turkish honeybees are representing eastern Mediterranean (C) lineage. According to pattern of clustering after using microsatellite analysis we can conclude that these populations have been in interaction with *A. m. meda*. These results may show that Amol population has not effected to a large extent from other populations and better represents Iranian honeybee. This distinction is also observed in grouping of populations according to morphometric data, but have not decided as subspecies.

# REFERENCES


Black, W.C. 1993. PCR with arbitrary primers: Approach with care. *Insect Mol. Biol.* 2(1): 1-6

Bodur Ç. 2001 Microsatellite analysis in honeybee populations of Turkey. MSc dissertation, Middle East Technical University, Ankara, Turkey.

Borehman, M.M. and Roubik, D.W. 1987. Population change and control of Africanized honeybee (Hymenoptera: Apidae) in Panama Canal area. *Bull. Entomonol. Soc. Am*. 33: 34-38.

Contel, E.P.B.,Mestriner, M.A. and Martins, E. 1977. Genetic control and developmental expression of malate dehydrogenase in *Apis mellifera. Biochem. Genet.* 15 (9/10): 859-876.

Cornuet J. M., Garnery L. and Solignac M. 1991**.** Putative origin and function of the intergenic region between COI and COII of Apis mellifera L. mitochondrial DNA. Genetics, 128, 393-403

Daly, H.V. 1988. Overview of the identification of the Africanized honey bees. In Africanized Honeybee and Bee Mites,pp 245-250. Ellis Horwood limited, Chechester, England.

De La Rua P., Gallian J. Serrano J., Moritz R. F. A.(2002) Microsatellite analysis of non-migratory colonies of *Apis mellifera iberica* from south-eastern Spain. J. Zool. Syst. Evol. Research 40 164-168.

Dietz, A. 1986. "Evolution" Bee Genetics and Breeding (Ed. T.E. Rinderer). Academic Press, Orlando.

Estoup A, Solignac M, Harry M, Cornuet JM (1993) Characterization of $(GT)_n$ and $(CT)_n$ microsatellites in two insect species: *Apis mellifera* and *Bombus terrestris*. Nuc. Ac. Res. 21, 1427-1431.

Estoup A, Garnery L, Solignac M, Cornuet JM (1995) Microsatellite variation in (*Apis mellifera* L) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. Genetics 140, 679-695.

Franck P, Garnery L, Solignac M, Cornuet JM (2000) Molecular confirmation of a fourth lineage in honeybees from the Near East. Apidologie 31, 167-180.

Freeman, S. and Herron J.C., (2002) Evolutionary Analysis, 2nd edition. Prentice Hall Inc., Upper Saddle River, New Jersey.

Garnery L, Cornuet JM, Solignac M (1992) Evolutionary history of the honeybee *Apis mellifera* inferred from mitochondrial DNA analysis. Mol. Ecol. 1, 145-154.

Garnery L, Solignac M, Celebrano G, Cornuet JM (1993) A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic structure of *Apis mellifera* L. Experientia 49, 1016-1021.

Garnery L, Franck P, Baudry E, Vautrin D, Cornuet JM, Solignac M (1998) Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A.m. iberica*). II. Microsatellite loci. Genet. Sel. Evol. 30 (Suppl. 1), S49-S74.

Goudet J, Raymond M, De Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. Genetics 144, 1933-1940.

Guo SW and Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportions for multiple alleles. Biometrics 48, 361-372.

Haldane JBS (1954) An exact test for randomness of mating. J. Genet. 52, 631-635.

Hall HG, (1988) Characterization of the African honeybee genotype by DNA restriction fragments. In Africanized Honeybee and Bee Mites, pp 313-324 Ellis Horwood limited.

Hall, H.G. and Smith, D.R. 1991. Distinguishing African and European honeybee matrilines using amplified mitochondrial DNA. *Proc. Natl. Acad. Sci. USA 88: 4548-4552*

Kandemir İ, Kence A (1995) Allozyme variability in a central Anatolian honeybee (*Apis mellifera* L) population. Apidologie 26, 503-510.

Kandemir İ (1999) Genetic and morphometric variation in honeybee (*Apis mellifera* L) populations in Turkey. Ph.D. dissertation, Middle East Technical University, Ankara, Turkey.

Kandemir İ, Kence M, Kence A (2000) Genetic and morphometric variation in honeybee (*A. mellifera* L.) populations of Turkey. Apidologie 31, 343-356.

Louveaux J. (1967) Annual developmental cycles of bee colonies and its variations. Annes des Epiphyties 19 (1): 165-& 1968

McPherson, J.M., Eckstein, P.E., Scoles, G.J. and Gahadhar, A.A. 1993. Variability of the random amplified polymorphic DNA assay thermal cycles, and effect of primer and DNA concentration. *Mol. Cell. Probes.* 7: 293-299.

Mestriner, M.A. 1969. Biochemical polymorphism in bee (*Apis mellifera ligustica). Nature 223:188-189.*

Mestriner, M.A. and Contel, E.P.B. 1972. The P-3 and EST loci in the honey bee *Apis mellifera*.Genetics 72:733-738.

Nei M (1972) Genetic distances between populations. Am. Nat. 106, 283-292.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89,583-590.

Ota T (1993) DISPAN (Genetic Distance and Phylogenetic Analysis) software. Pennsylvania State University.

Pardue, M. L., Lowenhaupt, K., Rich, A. and Nordheim, A, (1987) *EMBO J.*, 6, 1781-1789.

Palmer MR, Smith DR, Kaftanoğlu O (2000) Turkish honeybees: genetic variation and evidence for a fourth lineage of *Apis mellifera* mtDNA. The Journal of Heredity 91, 42-46.

Raymond M and Rousset F (1994) GenePop ver. 3.0. Institut des Sciences de l'Evolution. Universite de Montpellier, France.

Raymond M and Rousset F (1995) An exact test for population differentiation. Evolution 49, 1280-1283.

Rohlf, J.F. 1992. Numerical Taxonomy multivariate analysis system. Version 1.70 NTSYS-pc Manual Book, Applied Biostatistics, New York.

Ruttner, F., Tassencourt, L., and Louveaux, J. 1978. Biometrical-statistical analysis of the geographical variability of *Apis mellifera L. Apidologie* 9:363-381.

Ruttner, F. (1986) "Geographic Variability and Classification" Bee Genetics and Breeding (Ed. Thomas Rindrer). Acadaemic press, pp.23-52.

Ruttner F (1988) Biogeography and Taxonomy of Honeybees. Springer-Verlag, Berlin Heidelberg, 284 p.

Ruttner F (1992) Naturgeschichte der honigbienen. München: Ehrenwirth, 357 p.

Saitou N and Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic tree. Mol. Biol. Evol. 4, 406-425.

Sheppard WS and Berlocher SH. 1985. New allozyme variability in Italian honeybees. *J. Hered* 76:45-48

Sheppard WS (1986) Genetic variation and differentiation in honeybees (*Apis*) Ph. D. Dissertation. University of Illinois at Urbana-Champaigni Illinois.

Sheppard WS, Arias MC, Meixner MD, Grech A (1997) *Apis mellifera ruttnerii*, a new honeybee subspecies from Malta. Apidologie 28, 287-293.

Smith DR (1991) Mitochondrial DNA and honeybee biogeography, 131-176. In DR Smith (eds.), Diversity in the genus *Apis*. Westview, Boulder, CO.

Smith DR, Slaymaker A, Palmer M, Kaftanoğlu O (1997) Turkish honeybees belong to the east Mediterranean lineage. Apidologie 28, 269-274.

Solignac M., Vaturin D., Loiseu A., Mougel F., Baudry E., Estoup A., Garnery L., Harbel M., Cornuet J M.Five hundred and fifty microsatellite markers for the study of the honeybee (*Apis mellifera* L.) genome (2003).Molecular Ecology Notes 3, 307-311.

Tahmasbi G. (1996) Morphometric and Biochemical study of Iranian honeybee populations Ph. D. Dissertation. University of Tarbiyat Modares.

Weir BS and Cockerham CC (1984) Estimatinf F-statistics for the analysis of population structure. Evolution 38, 1358-1370.

Weir BS (1990) Genetic data analysis. Sinauer Publ., Sunderland, MA.
Williams, J.G.K.,kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. 1990. DNA polymorphism amplified by arbitary primers are useful as genetic markers. *Nuc. Acids Res*. 18(22): 6531-5635.

Wilson, E.O 1971.Insect Societies. Harvard University Press, Cambridge.

# APPENDIX A

LIST OF REAGENTS

| Reagents | Brand | Catalogue number |
|---|---|---|
| Acrylamide/bis-Acrylamide | Sigma | A-2917 |
| Ammonium Persulfate | Sigma | A-9164 |
| Bovine Serum Albumin | MBI Fermentas | B14 |
| Dithiothreitol | Sigma | D-9779 |
| dNTP set | MBI Fermentas | RO181 |
| MgCl | MBI Fermentas | |
| PCR Buffer | MBI Fermentas | |
| Phenol | AppliChem | A1153,0500 |
| Primers | IDT | |
| Taq DNA Polymerase | MBI Fermentas | EP0402 |
| TBE Buffer | Sigma | T-4415 |
| TEMED | Sigma | T-7024 |
| Urea | AppliChem | A1049,5000 |

## APPENDIX B

## COMPOSITIONS OF SOLUTIONS

Table 1. Preparation of Wilson buffer

| |
|---|
| Add the followings: |
| 10 ml from 1 M Tris.Cl pH 8 stock solution |
| 200 µl from 0.5 M Ethylenediaminetetraacetic acid (EDTA) stock solution |
| 1 ml from 10% (w/v) Lauryl sulfate (SDS) stock solution |
| 0.771 g of Dithiothreitol (DTT) |
| 0.584 g of Sodiumchloride (NaCl) |
| Add distilled water to complete to 100 ml. |

Table 2. Solution B

| | |
|---|---|
| TBE 10X | 10 ml |
| $H_2O$ | 15 ml |
| Temed | 25 ml |

Table 3. 3.6% acrylamide gel solution

| | |
|---|---|
| Solution B | 20 ml |
| 40% Acrylamide solution (19:1) | 20 ml |
| Persulfate amoniom (10%) | 13 ml |
| $H_2O$ | 27 ml |

Table 4. 10% acrylamide gel solution

| |
|---|
| 75 ml from 40% acrylamide solution |
| 50 ml from 10x TBE |
| 240 g from urea |
| |
| Adjust the volume to 500 ml by distilled water |

Table 5. Six percent acrylamide / urea solution

| | |
|---|---|
| Slolution B | 20 ml |
| 40% Acrylamide solution (19:1) | 7,5 ml |
| persulfate amoniom (10%) | 13 ml |
| $H_2O$ | 39.5 ml |

Table 6. Loading buffer for polyacrylamide gel electrophoresis

| | |
|---|---|
| Formamid | 10 ml |
| Xylene cyanol FF | 10 mg |
| Bromophenol blue | 10 mg |
| 0.5 M EDTA (pH=8) | 200 µl |