Investigating the Phylogenetic Relationships of the Iranian Honey Bee with Other Bee Breeds Worldwide Using the Cytochrome Oxidase I Region

Mansoureh Sarhaddi Avval [1], Bahman Fazeli-Nasab [2, *], Anahita Pahlavan [3], Ali Reza Mirzaei [4]

[1] Department of Plant Protection, College of Agriculture, University of Zabol, Zabol, Iran
[2] Department of Agronomy and Plant Breeding, Agriculture Institute, Research Institute of Zabol, Zabol, Iran
[3] Department of Horticultural, Faculty of Agriculture, University of Zabol, Zabol, Iran
[4] Department of Agronomy and Plant Breeding, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran

*Corresponding Author: Department of Agronomy and Plant Breeding, Agriculture Institute, Research Institute of Zabol, Zabol, Iran. Email: bfazelinasab@gmail.com

Abstract

Background: Honey bees, classified within the insect class and widely distributed globally, consist of 26 identified subspecies across five evolutionary lines. Among these, the Iranian honey bee (*Apis mellifera meda*) is notable for its presence in Iran and neighboring regions. This subspecies exhibits unique biological traits, including a high reproductive capacity and effective propolis collection, distinguishing it from other local bee populations. Understanding genetic diversity within these populations is crucial for conservation and breeding efforts.

Objectives: The aim of this study is to determine the phylogenetic relationships and natural variations of the Iranian honey bee in comparison to other bee breeds worldwide using the cytochrome oxidase I (COI) region.

Methods: The study focuses on the identification of COI gene sequences in honeybees. The mitochondrial COI gene sequences related to the Iranian honey bee and eight populations of other honey bee breeds from around the world (24 ecotypes) were extracted and aligned from the genome database (NCBI). BLAST was used for sequence retrieval, and Clustal W was employed for alignment among species. Phylogenetic analysis was performed using MEGA software, applying the Neighbor Joining method to construct a phylogenetic tree based on the edited sequences. The percentage of nucleotide substitutions and replacements was calculated using the maximum likelihood method. Additionally, molecular diversity indices were computed using DnaSP software to assess genetic variation and conservation parameters among populations.

Results: The bioinformatics analysis of biological data revealed that transfer substitution occurred more frequently than crossover substitution. Additionally, the percentage of substitution in pyrimidine bases was higher than in purine bases. The dN/dS substitution ratio in the gene locus sequence across the examined ecotypes was estimated at 0.08, indicating purifying selection during the evolution of this gene locus. Furthermore, examination of the COI region of the mitochondrial genome identified a total of 4 haplotypes, 20 mutations, and 195 polymorphic loci within the entire studied population. The neutrality values, based on Tajima's test, indicated effects related to genetic drift, genetic bottlenecks, or balancing selection during the evolutionary history of the studied population. Phylogenetic analysis revealed nine distinct categories of evolutionary paths for the COI gene in the studied ecotypes. In the phylogenetic tree, the Iranian honey bee was placed in a separate cluster, exhibiting the highest and lowest phylogenetic distances relative to *Apis koschevnikovi* and *Apis mellifera*, respectively.

Conclusions: This research indicates that the conserved region of the mitochondrial cytochrome oxidase gene among various bee species is minimal, highlighting significant polymorphism and susceptibility to mutations. A total of 20 mutations and 195 polymorphic sites were identified, suggesting purifying selection processes that contribute to the gene's biological functions. The study also demonstrates that the COI region can effectively differentiate Iranian honey bee populations from others globally. Overall, these findings enhance our understanding of genetic diversity and evolutionary relationships within bee species, underscoring the importance of continued research for conservation efforts.

Keywords: Biodiversity, Cytochrome Oxidase I (COI), Haplotypes, Mitochondrial Genome, Phylogenetic Relationships

1. Background

The honey bee is an insect belonging to the class of insects and the order of diptera, found worldwide across various geographical regions (1). Based on morphological and molecular studies, 26 bee subspecies have been identified, which fall into five evolutionary lines: The Northern and Western European line (M), Central and Southern Africa (A), the Northern Mediterranean region and Eastern Europe (C), the Eastern Mediterranean, Near East, and Eastern Middle East (O), and finally the Eastern group of Ethiopia (Y) (2, 3). Among the eight subspecies of bees living in the Middle East, one subspecies, *Apis mellifera meda*, is found in Iran, known as the Iranian honey bee (1). In addition to Iran, this subspecies also occurs in northern Iraq, southeastern Turkey, and northern Syria (2, 3). Morphologically, it is similar to the Italian breed. In terms of biological characteristics, it is a breed with a high reproductive capacity, the ability to collect propolis, excellent wintering ability, and high aggression, all of which distinguish it from other subspecies in the region (4).

To study different populations of organisms, several methods have been widely adopted by systematic researchers, one of which is genetic diversity (5). Genetic diversity is a crucial element for genetic progress, preservation, evolution, and adaptation to changing and diverse environmental conditions (6). Phenotypic flexibility and genetic variability in traits used to identify species can lead to misidentification. However, the use of more accurate and faster systems for identifying living organisms, including molecular identifiers, can aid in the process of sample identification (7). Among molecular markers, the mitochondrial and nuclear genomes are both very useful in studying evolution and resolving relationships between different animal populations (6). The mitochondrial genome is often preferred over the nuclear genome in animals due to its multiple copies in cells, low recombination rate, haploid inheritance, and ease of extraction (8).

One of the most suitable mitochondrial genes for research using animal molecular markers is the mitochondrial cytochrome oxidase I (COI) gene (cytochrome oxidase c subunit one), which has evolved in such a way that it can distinguish even very closely related species (9). Cytochrome oxidase is the final enzyme in the mitochondrial respiratory chain, responsible for transferring electrons from cytochrome to oxygen. It is a protein complex found in the inner mitochondrial membrane, composed of 13 protein subunits, 10 of which are encoded by the nuclear genome and the remaining 3 by the mitochondrial genome (10). Cytochrome c oxidase is an enzyme present in all living organisms (11). This fact suggests that the enzyme has ancient origins, likely first arising in a primitive organism and subsequently being passed down to all organisms during evolution across generations (12).

In honey bees, genetic studies are primarily based on the analysis of both morphological (13) and molecular data (14). These two methods are currently used as powerful tools to estimate genetic diversity and determine phylogenetic relationships between different populations of bee subspecies worldwide. While morphological characteristics are useful for phylogenetic studies, they are highly sensitive to environmental factors and changing environmental conditions. In contrast, for phylogenetic studies, enzyme digestion analysis using mitochondrial DNA information provides a better and more reliable genetic marker (15). One advantage of mitochondrial DNA over morphological data is that mitochondrial DNA is more easily analyzed in the context of phylogenetic issues. Additionally, due to maternal inheritance, lack of recombination, and low efficiency of complete repair mechanisms, mtDNA evolves more rapidly than single-copy genes in nuclear DNA. Therefore, mtDNA is particularly useful for studying phylogenetic relationships and the historical changes among bee subspecies (16).

2. Objectives

In general, the first step in achieving targeted breeding is to determine the level of genetic diversity within populations. Based on the aforementioned content, and in order to identify genetic reserves and preserve breed diversity, the present study was conducted to investigate the genetic and phylogenetic structure of native Iranian honey bees and several important honey bee species worldwide.

3. Methods

3.1. Identification and Editing of Cytochrome Oxidase I Gene Sequences

To identify COI gene sequences, the DNA and mRNA nucleotide sequences for the COI gene in honey bees were identified. Additionally, using the BLAST tool in genome databases, the COI gene sequences for other target species were obtained (Table 1). To compare the sequences and determine genetic parameters, the COI gene sequence of the Iranian honey bee was aligned with those of other species using Clustal W.

3.2. Phylogeny and Determining the Evolutionary Path

The phylogenetic tree of the COI sequence was constructed for the studied species using MEGA software. After editing the sequences and removing non-homogeneous regions, the phylogenetic tree was drawn using the Neighbor Joining (NJ) method. In this method, the matrix (Q) is used, which includes all the branches, and selects the lowest value that represents the highest similarity between two branches, which is then used in the phylogeny tree. Bootstrapping values were obtained by resampling 1,000 times.

3.3. Molecular Diversity Indices

Molecular diversity indices such as the number of polymorphic sites, number of haplotypes, haplotype frequency, nucleotide diversity, total number of mutations, nucleotide differences between populations, and indices of conservation in protected areas, such as sequence conservation, minimum conservation length, conservation threshold limit, and neutrality statistics, were analyzed using Dnasp software.

3.3.1. Molecular Diversity Indices

3.3.1.1. Polymorphic Sites

These are positions in the DNA sequence where variations occur among individuals in a population. A higher number of polymorphic sites indicates greater genetic diversity.

3.3.1.2. Haplotypes and Haplotype Frequency

Haplotypes refer to a combination of alleles at multiple loci that are transmitted together on the same chromosome. Haplotype frequency is the proportion of a specific haplotype relative to the total number of haplotypes in the population, providing insights into genetic structure.

3.3.1.3. Nucleotide Diversity (π)

This measures the average number of nucleotide differences per site between two sequences. It is a crucial metric for assessing genetic variation within a population.

2.3.1.5. Total Number of Mutations

This refers to the cumulative count of mutations observed in a given sequence, which can signify evolutionary changes over time.

2.3.1.6. Nucleotide Difference Between Populations

This index quantifies the genetic distance between different populations, helping to understand evolutionary relationships.

2.3.1.7. Conservation Indices

Sequence conservation: This measures how much of a sequence remains unchanged across different species or populations, indicating evolutionary pressure.

Minimum conservation length and conservation threshold limit: These metrics help establish how much of a sequence must be conserved for it to be considered functionally important.

Neutrality Statistics: These tests (e.g., Tajima's D) assess whether allele frequencies in a population conform to expectations under neutral evolution, which can indicate past demographic events.

4. Results

4.1. Experiment and Analysis of Genetic Distance and Phylogenetic Tree

The results of the evolutionary diversity in the 26 studied species are presented in Table 2. Since the genetic distances between different species are calculated pairwise, the resulting numbers represent the nucleotide substitutions between the studied sequences. The genetic distance between two species is related to their phylogenetic correlation, making this

Table 1. Cytochrome Oxidase I Gene Sequence ID Numbers in Different Bee Species

No.	Scientific Name	ID Number in the NCBI Database
1	*Apis andreniformis Aan 4.1*	LC427579.1
2	*Apis andreniformis voucher AAT95*	MT670338.1
3	*Apis andreniformis voucher AAT97*	MT670340.1
4	*Apis cerana isolate 8*	OK253002.1
5	*Apis cerana indica voucher ACIOR1*	OP168349.1
6	*Apis cerana indica breed Honey bee*	MW093739.1
7	*Apis koschevnikovi Ak1_TB*	LC461206.1
8	*Apis koschevnikovi Ak1_HST*	LC461204.1
9	*Apis koschevnikovi Ak1_BL*	LC461205.1
10	*Apis florea voucher UBN3*	OK350423.1
11	*Apis florea voucher UBN1*	OK350421.1
12	*Apis florea voucher SMI9*	OK350397.1
13	*Apis dorsata isolate Jatoli*	OQ346377.1
14	*Apis laboriosa voucher ALM6*	MT679379.1
15	*Apis laboriosa voucher ALM5*	MT679378.1
16	*Apis laboriosa voucher ALM1*	MT679374.1
17	*Apis nigrocincta AnBK4.2*	LC534595.1
18	*Apis nigrocincta AnBK3.3*	LC534593.1
19	*Apis nigrocincta AnBK3.2*	LC534592.1
20	*Apis mellifera voucher HY42*	OP435367.1
21	*Apis mellifera voucher HY41*	OP435365.1
22	*Apis mellifera isolate Mellifera*	ON680902.1
23	*Apis mellifera isolate ACKU*	OM320635.1
24	*Apis mellifera meda isolate DUHOK29-9*	OP020350.1
25	*Apis mellifera meda isolate DUHOK29-7*	OP020348.1
26	*Apis mellifera meda isolate DUHOK29-6*	OP020347.1

index useful for determining the genetic distance and proximity between species.

Based on the Poisson correlation of the evolutionary matrix, the calculation and results indicated that the Iranian honey bee exhibits the highest genetic distance (1.615) with the *A. mellifera* isolate ACKU species, and the lowest genetic distance (0.012) with the *A. mellifera* isolate *Mellifera* species. The numbers obtained from the genetic distance matrix reflect the degree of nucleotide substitution between the sequences. The genetic distance between two species is closely related to their phylogenetic correlation, making this index useful for assessing the distance or proximity between species (17).

Phylogenetic analysis estimates how members of a family may have derived from one another during the course of evolution. These relationships are determined by studying mutations such as substitutions, deletions, insertions, and rearrangements that are subject to natural selection. The purpose of phylogenetic analysis

is to uncover a branching arrangement in the form of a tree that illustrates the best relationships between sequences. It also enables the discovery of gene functions, the tracing of gene origins, and the identification of an organism's relatives (17).

Evolutionary relationships between sequences are represented using diagrams called phylogenetic trees. The purpose of phylogenetic analysis is to discover an arrangement of branches in the form of trees that show the best relationships of sequences. Based on gene sequence cluster analysis, the studied species were classified into nine categories (Figure 1). Category one is related to *Apis koschevnikovi*; the second category is related to *Apis cerana indica*; the third category is related to *Apis laboriosa*; the fourth category is related to *Apis dorsata* and *Apis mellifera* species (excluding the *Mellifera* subspecies); the fifth category is related to *Apis andreniformis* species (excluding the Ana 4.1 subspecies) and Apis flora; the sixth category is related to *Apis*

Table 2. Matrix of Estimation of Evolutionary Diversity and Genetic Distance of Cytochrome Oxidase I Genes Among Studied Bee Species

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Specie 1																									
Specie 2	1.293																								
Specie 3	1.293	0.000																							
Specie 4	0.119	1.712	1.712																						
Specie 5	1.221	0.124	0.124	1.168																					
Specie 6	1.221	0.129	0.129	1.482	0.013																				
Specie 7	1.173	0.123	0.123	1.302	0.114	0.098																			
Specie 8	1.173	0.123	0.123	1.302	0.114	0.098	0.000																		
Specie 9	1.162	0.119	0.119	1.285	0.111	0.098	0.003	0.003																	
Specie 10	1.372	0.096	0.096	1.483	0.108	0.119	0.114	0.114	0.111																
Specie 11	1.372	0.096	0.096	1.483	0.108	0.119	0.114	0.114	0.111	0.000															
Specie 12	1.370	0.094	0.094	1.480	0.105	0.117	0.111	0.111	0.108	0.002	0.002														
Specie 13	1.306	0.120	0.120	1.390	0.117	0.103	0.124	0.124	0.127	0.134	0.134	0.132													
Specie 14	1.233	0.115	0.115	1.560	0.100	0.090	0.102	0.102	0.105	0.102	0.102	0.100	0.103												
Specie 15	1.233	0.115	0.115	1.560	0.100	0.090	0.102	0.102	0.105	0.102	0.102	0.100	0.103	0.000											
Specie 16	1.233	0.115	0.115	1.560	0.100	0.090	0.105	0.105	0.105	0.102	0.102	0.100	0.106	0.002	0.002										
Specie 17	0.113	1.253	1.253	0.074	1.240	1.167	1.138	1.138	1.128	1.237	1.237	1.234	1.288	1.133	1.133	1.133									
Specie 18	0.112	1.239	1.239	0.068	1.225	1.153	1.125	1.125	1.115	1.222	1.222	1.220	1.273	1.119	1.119	1.119	0.008								
Specie 19	0.112	1.239	1.239	0.068	1.225	1.153	1.125	1.125	1.115	1.222	1.222	1.220	1.273	1.119	1.119	1.119	0.008	0.000							
Specie 20	1.315	0.142	0.142	1.570	0.127	0.125	0.117	0.117	0.120	0.120	0.120	0.118	0.105	0.103	0.103	0.105	1.187	1.173	1.173						
Specie 21	1.310	0.144	0.144	1.576	0.129	0.125	0.117	0.117	0.121	0.122	0.122	0.120	0.108	0.105	0.105	0.108	1.187	1.173	1.173	0.002					
Specie 22	0.116	1.733	1.733	0.095	1.309	1.633	1.359	1.359	1.342	1.581	1.581	1.579	1.388	1.603	1.603	1.603	0.112	0.106	0.106	1.592	1.597				
Specie 23	1.311	0.147	0.147	1.711	0.127	0.125	0.113	0.113	0.115	0.119	0.119	0.117	0.105	0.107	0.107	0.109	1.187	1.173	1.173	0.000	0.002	1.751			
Specie 24	0.113	1.575	1.575	0.108	1.325	1.531	1.398	1.398	1.386	1.519	1.519	1.517	1.372	1.481	1.481	1.481	0.108	0.104	0.104	1.453	1.445	0.012	1.615		
Specie 25	0.113	1.575	1.575	0.108	1.325	1.531	1.398	1.398	1.386	1.519	1.519	1.517	1.372	1.481	1.481	1.481	0.108	0.104	0.104	1.453	1.445	0.012	1.615	0.000	
Specie 26	0.113	1.575	1.575	0.108	1.325	1.531	1.353	1.353	1.341	1.519	1.519	1.517	1.372	1.481	1.481	1.481	0.108	0.104	0.104	1.453	1.445	0.012	1.615	0.000	0.000

nigrocincta; the seventh category is related to *Apis mellifera meda* and *Mellifera Apis mellifera* isolate species; and the eighth and ninth categories are related to *Apis cerana* isolate 8 and *Apis andreniformis* Ana 4.1 species, respectively. The highest and lowest phylogenetic distances for the Iranian honey bee are with *Apis koschevnikovi* and *Mellifera Apis mellifera* isolate, respectively.

4.2. Nucleotide Substitutions

The results of the substitution percentages for COI gene nucleotide sequences, shown in Table 3, indicate that the highest substitution values are associated with pyrimidine bases. Specifically, the conversion of thymine to cytosine was 9.8%, while the conversion of cytosine to thymine was 31.51%. These values were lower for purine bases, with the conversion of adenine to guanine and guanine to adenine at 2.74% and 7.84%, respectively.

4.3. Replacement Rate

Based on the results in Table 2, the dN/dS ratio for the COI gene was calculated for different bee species, with the numerical value of the ratio (dN/dS) being 0.08. If this ratio is greater than one, it indicates positive selection; if it is less than one, it indicates purifying
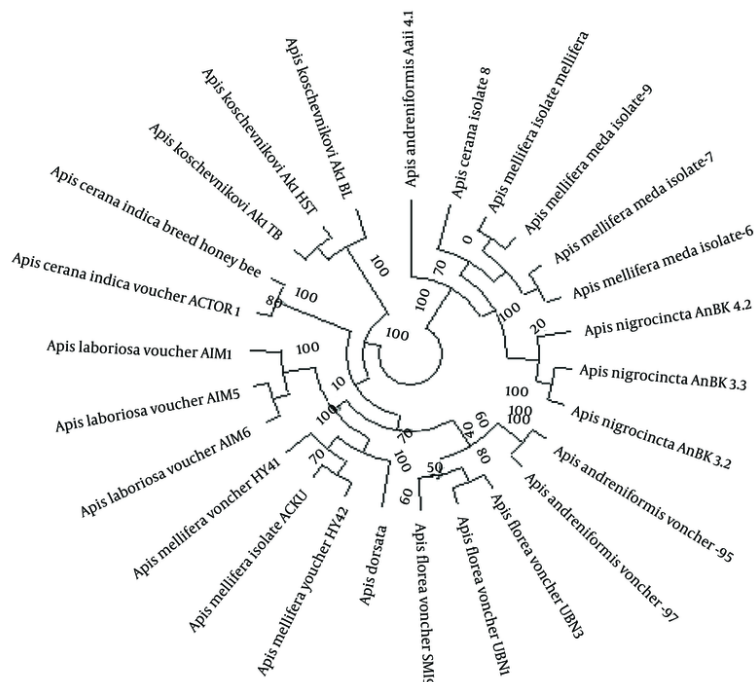
Figure 1. Phylogenetic tree of nucleotide sequences of cytochrome oxidase I (COI) genes in studied bee species (0.2 nucleotide substitution rate per position).

Table 3. The Percentage of Cytochrome Oxidase I Gene Nucleotide Sequence Substitutions in the Examined Species Based on the Maximum Likelihood Method [a]

Nucleotides	A	T/U	C	G
A	-	22.7992	1.3910	2.74 [b]
T	18.3493	-	9.8 [b]	0.3126
C	3.6015	31.5107 [b]	-	0.2617
G	7.84 [b]	1.1113	0.2892	-

[a] This table shows the percentage of substitution from a row to a column. The numbers outside the diagonal represent crossover substitutions (italicized).
[b] The numbers on the diagonal represent transition substitutions.

selection; and if it is equal to one, it indicates neutral selection during gene evolution ([18], [19]). In this study, the dN/dS ratio was found to be less than one, suggesting purifying selection during the evolution of the COI gene ([Table 4]).

4.4. Genetic Diversity Indices

The nucleotide diversity and haplotype diversity in terms of base pairs for the studied bee populations were estimated at 0.259 and 0.966, respectively. The average

nucleotide difference between ecotypes was calculated as 83.19 bp. Additionally, in the analysis of genetic diversity indices for the COI region of the mitochondrial genome in the native Iranian bee ecotype, compared with the nucleotide sequences of this region in the other studied bee ecotypes, 20 point mutations were identified. These mutations were distributed differently along the length of the mitochondrial genome. In research conducted in Pakistan and China, the COI region of the mitochondrial genome ([20]) was also used to study the genetic structure of the honey bee (*A.*

Table 4. Parameters for Determining the Process of Natural Selection at the Cytochrome Oxidase I Gene Locus

Parameters	Numerical Value	Standard Deviation
dN	0.030950	0.008807
dS	0.386167	0.041641
dN/dS	0.080146	0.000154

Abbreviations: dN, nucleotide changes that cause amino acid changes; dS, nucleotide changes that do not affect the resulting amino acid; of dN/dS the process of natural selection.

Table 5. Genetic Diversity Indices for the Nucleotide Sequence of the Cytochrome Oxidase I Region of the Mitochondrial Genome Among the Studied Bee Ecotypes

Indices	Theta-W	S	H	Hd	Pi	Eta	K
Values	0.159	195	4	0.966	0.259	20	83.19

Abbreviations: Theta-W, mutation rate calculated from all polymorphic sites; S, number of polymorphic loci; H, number of haplotypes; Hd, haplotype frequency; Pi, nucleotide diversity; Eta, total number of mutations; K, nucleotide difference between populations or species (nucleotide divergence).

Table 6. Neutrality Statistics and Indices of the Analysis of the Protected Regions of the Cytochrome Oxidase I Region of the Mitochondrial Genome Among the Studied Bee Ecotypes

Indices	D Tajima Test	CT	MWL	C
Values	1.174	0.49	80	0.396

Abbreviations: CT, protection threshold limit; MWL, minimum protection length; C, sequence protection.

mellifera) population. The results indicated that, in the entire studied population, 44 polymorphisms, 2.429 inter-population nucleotide differences, and 23 mutations were detected (Table 5).

According to Table 6, the analysis of indices for the protected areas of the COI region of the mitochondrial genome in the studied population revealed the following values: Minimum protection threshold = 0.396, protection length = 80, and sequence protection = 0.49. Neutrality tests, including the Tajima's D test, were performed to identify any deviation from the null hypothesis of neutral evolution and to examine the effects of natural selection on genes across different populations. The numerical value for Tajima's D was estimated to be 1.174 in this study.

5. Discussion

Previous research on animals and various genes has concluded that the cytochrome oxidase gene is a useful and reliable marker for differentiating populations (21-23). In one study (24), COI genes were used to investigate the phylogenetic characteristics of honey bee populations. The results demonstrated the suitability of

the COI gene for studying and differentiating various honey bee populations in Iran. Another study conducted by previous researchers (25) in Thailand used genetic comparisons of mitochondrial genes COI and COII to differentiate populations of the small honey bee (*A. florea*) and the small honey bee (*A. andreniformis*). The results showed that these genes are effective in distinguishing the populations of these two species. The mitochondrial COI gene was also employed in research to differentiate populations of bees collected from the northwestern Himalayan region, where the results indicated that the gene region plays an important role in separating populations and subspecies of honey bees in this area (26). In another study by previous researchers (27), the COI mitochondrial gene was used to distinguish different populations of the Apis genus. The findings confirmed that the COI gene, as a molecular marker, can be effectively used to analyze the ancestral status and evolution of Apis species and has a superior ability to differentiate between bee species.

In this study, the investigation of phylogenetic relationships showed that the genetic diversity of the evaluated ecotypes does not align with their geographical diversity. Although, in most cases, the

study populations were classified based on their geographical origin, it was observed that, in some instances, two populations from geographically close areas were placed into completely separate groups in the population classification based on molecular data. For example, the ecotype of *Apis mellifera meda* from Iran and the ecotype of *Apis andreniformis* from Southeast Asia were classified into separate groups in the cluster analysis, despite both populations belonging to similar climates. Such results are also evident in the cluster analysis of previous research (24), the grouping of small honey bee populations using COI and COII mitochondrial genes in earlier studies (2, 20), the grouping of different Iranian honey bee populations using the mitochondrial ND2 gene in previous research (1), the study of the genetic diversity of bee populations in Kerman province using ISSR markers (16), and the investigation of phylogenetic relationships in Iranian honeybees using both morphological and molecular markers.

To explain cases of inconsistency between groupings based on molecular data and geographical origin, it can be suggested that these groupings may reflect the unique genetic backgrounds of these populations and the presence of specific features in their genome sequences. Additionally, placing populations from geographically distant areas into the same group based on molecular data may be due to factors such as germplasm transfer, the insufficient number of markers used in genetic diversity studies, or the broad distribution of a plant species within the studied geographical area. In other words, the geographical distance between populations cannot be considered conclusive evidence of the distance or proximity of their genetic backgrounds (28-30).

Four microsatellite loci of honey bee populations in northwestern Iran have been investigated in various studies, including the study of morphometric characteristics of honey bee (*Apis mellifera* L.) populations in Kermanshah province (16, 28, 31), the study of morphological diversity of the small honey bee in Iran using statistical methods such as cluster analysis and minimum variance (32), and an investigation of mitochondrial genes ND1 and ND5 to explore the separation of Iranian honey bees from global subspecies (33). Additionally, the genetic diversity of bee populations in Southeast Europe has been studied using

22 microsatellite markers (34), and the genetic diversity of honey bee populations in the Tomsk region of Russia was analyzed using COI and COII genes (35). Furthermore, the genetic diversity of Argentine honey bee populations has been examined using COI and COII genes (36). All these studies have shown that there is significant genetic diversity within the species, with notable intraspecies diversity at the DNA level. According to these researchers, factors such as colony migration, migratory beekeeping, and the random mating of queens with males from different colonies play important roles in creating diversity within honey bee populations.

The substitution values obtained in this study are consistent with those reported by other researchers, who found the highest substitution rate to be in pyrimidine bases (37). Generally, the cause of these changes can be attributed to cytosine methylation (38).

The dN/dS method is one of the key indicators used to detect the process of natural selection during gene evolution. In this method, the results of nucleotide changes that lead to amino acid changes (dN) are compared to nucleotide changes that do not affect amino acids (dS) (18, 19).

Populations that have undergone expansion, or experienced a significant increase in size, or that have been subject to directional selection, typically show negative and significant D values. On the other hand, positive and significant values of D indicate the effects of genetic drift, genetic bottlenecks, or balancing selection during the evolutionary history of the population. Genetic drift leads to increased homogeneity and the loss of alleles, particularly recessive alleles. Ultimately, these factors result in reduced diversity and a decreased ability of populations to adapt to environmental changes, a process that is especially common in smaller and more homogeneous populations (18, 39).

5.1. Conclusions

In conclusion, the results of this study indicate that the conserved region of the mitochondrial cytochrome oxidase gene among the species studied represents only a small portion of the entire sequence. This suggests a high level of polymorphism in this gene, as well as its susceptibility to nucleotide changes and mutations. A total of 20 mutations were identified during the

evolution of this gene locus. Additionally, 195 polymorphic sites were observed, supporting the idea of purifying selection acting on this gene. This may play a crucial role in understanding the biological function of the gene, as it may lead to the emergence of new proteins and functions. Finally, by comparing the nucleotide sequences of Iranian honey bee populations with those of other honey bee populations from different regions of the world in gene banks, it was determined that the COI region is capable of distinguishing Iranian honey bee populations from other populations.

The study provides significant insights into the genetic diversity and phylogenetic relationships among various bee species, particularly the Iranian honey bee. The analysis reveals a notable genetic distance between the Iranian honey bee and other species, indicating its unique evolutionary path. The findings highlight the importance of mitochondrial COI gene sequences in understanding genetic variation and natural selection processes. Furthermore, the results underscore the necessity for ongoing research to preserve bee biodiversity and inform effective conservation strategies. Overall, this research contributes valuable data to the field of apidology and enhances our understanding of honey bee evolution.

Footnotes

Authors' Contribution: Conceptualization: M. S. A., B. F. N.; data curation: M. S. A., B. F. N.; formal analysis: M. S. A., B. F. N.; investigation: M. S. A., B. F. N.; methodology: M. S. A., B. F. N.; project administration: M. S. A., B. F. N.; resources: M. S. A., B. F. N., A. M.; validation: M. S. A., B. F. N.; visualization: M. S. A., B. F. N., A. P.; writing–original draft: All authors; writing–reviewing & editing: All authors.

Conflict of Interests Statement: The authors declare no potential conflict of interest.

Data Availability: All the data are embedded in the manuscript.

Funding/Support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Bahador Y, Mohammadabadi M, Khezri A, Asadi M, Medhati L. [Study of Genetic Diversity in Honey Bee Populations in Kerman Province using ISSR Markers]. *Res Animal Product.* 2016;**7**(13):192-86. FA. https://doi.org/10.18869/acadpub.rap.7.13.192.

2. Garnery L, Solignac M, Celebrano G, Cornuet JM. A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic structure of Apis mellifera L. *Experientia.* 1993;**49**(11):1016-21. https://doi.org/10.1007/bf02125651.

3. Safdari Shahroudi M, Mehrabani H, Pakdel A, Nejati-Javaremi A, Nehzati-Paghalee GA. [PCR-RFLP Analysis of Two mtDNA Gene Segments in Iranian Honey Bee Population]. *J Agricultural Biotechnolo.* 2013;**4**(1):21-6. FA.

4. Hammond A, Rowland T, Mills DS, Pilot M. Comparison of behavioural tendencies between "dangerous dogs" and other domestic dog breeds - Evolutionary context and practical implications. *Evol Appl.* 2022;**15**(11):1806-19. [PubMed ID: 36426126]. [PubMed Central ID: PMC9679229]. https://doi.org/10.1111/eva.13479.

5. Brito RM, Arias MC. Mitochondrial DNA characterization of two Partamona species (Hymenoptera, Apidae, Meliponini) by PCR+RFLP and sequencing. *Apidologie.* 2005;**36**(3):431-7. https://doi.org/10.1051/apido:2005020.

6. Mohammadabadi MR. Role of Clostridium Perfringens in Pathogenicity of Some Domestic Animals. *J Adv Agriculture.* 2017;**7**(3):1117-21. https://doi.org/10.24297/jaa.v7i3.6325.

7. Bovo S, Utzeri VJ, Ribani A, Taurisano V, Schiavo G, Fontanesi L. A genotyping by sequencing approach can disclose Apis mellifera population genomic information contained in honey environmental DNA. *Sci Rep.* 2022;**12**(1):19541. [PubMed ID: 36379985]. [PubMed Central ID: PMC9666642]. https://doi.org/10.1038/s41598-022-24101-z.

8. Charistos L, Hatjina F, Bouga M, Mladenovic M, Maistros AD. Morphological Discrimination of Greek Honey Bee Populations Based on Geometric Morphometrics Analysis of Wing Shape. *J Apicultural Sci.* 2014;**58**(1):75-84. https://doi.org/10.2478/jas-2014-0007.

9. Hebert PD, Ratnasingham S, deWaard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Biol Sci.* 2003;**270 Suppl 1**(Suppl 1):S96-9. [PubMed ID: 12952648]. [PubMed Central ID: PMC1698023]. https://doi.org/10.1098/rsbl.2003.0025.

10. Eimanifar A, T. Kimball R, L. Braun E, Fuchs S, Grunewald B, Ellis JD. The complete mitochondrial genome of Apis mellifera meda (Insecta: Hymenoptera: Apidae). *Mitochondrial DNA B Resour.* 2017;**2**(1):268-9. [PubMed ID: 33473795]. [PubMed Central ID: PMC7800517]. https://doi.org/10.1080/23802359.2017.1325342.

11. Kosakyan A, Heger TJ, Leander BS, Todorov M, Mitchell EA, Lara E. COI barcoding of Nebelid testate amoebae (Amoebozoa: Arcellinida): extensive cryptic diversity and redefinition of the Hyalospheniidae Schultze. *Protist.* 2012;**163**(3):415-34. [PubMed ID: 22130576]. https://doi.org/10.1016/j.protis.2011.10.003.

12. Medeiros-Santana L, Dohanik VT, Serrao JE. Differential expression of aquaporin genes during ovary activation in the honeybee Apis mellifera (Hymenoptera: Apidae) queens. *Comp Biochem Physiol B Biochem Mol Biol.* 2021;**253**:110551. [PubMed ID: 33359142]. https://doi.org/10.1016/j.cbpb.2020.110551.

13. Rahimi A, Mirmoayedi A, Kahrizi D, Zarei L, Jamali S. Morphometric diversity and phylogenetic relationships among Iranian honey bee (Apis mellifera meda Skorikow, 1829) populations using morphological characters. *Sociobiol.* 2017;**64**(1). https://doi.org/10.13102/sociobiology.v64i1.1179.

14. Rangel J, Giresi M, Pinto MA, Baum KA, Rubink WL, Coulson RN, et al. Africanization of a feral honey bee (Apis mellifera) population in South Texas: does a decade make a difference? *Ecol Evol.* 2016;**6**(7):2158-69. [PubMed ID: 27069571]. [PubMed Central ID: PMC4782243]. https://doi.org/10.1002/ece3.1974.

15. Techer MA, Clémencet J, Turpin P, Volbert N, Reynaud B, Delatte H. Genetic characterization of the honeybee (Apis mellifera) population of Rodrigues Island, based on microsatellite and mitochondrial DNA. *Apidologie.* 2014;**46**(4):445-54. https://doi.org/10.1007/s13592-014-0335-9.

16. Rahimi A, Mirmoayedi A, Kahrizi D, Zarei L, Jamali S. Molecular genetic diversity and population structure of Iranian honey bee (Apis mellifera meda) populations: implications for breeding and conservation. *J Plant Dis Protection.* 2022;**129**(6):1331-42. https://doi.org/10.1007/s41348-022-00657-w.

17. Huson DH, Steel M. Distances that perfectly mislead. *Syst Biol.* 2004;**53**(2):327-32. [PubMed ID: 15205056]. https://doi.org/10.1080/10635150490423809.

18. Fazeli Nasab B. [Biological Evaluation of Coronaviruses and the Study of Molecular Docking, Linalool, and Thymol as orf1ab Protein Inhibitors and the Role of SARS-CoV-2 Virus in Bioterrorism]. *J Ilam Univ Med Sci.* 2021;**28**(6):77-96. FA. https://doi.org/10.29252/sjimu.28.6.77.

19. Pahlavan A, Solouki M, Fakheri B, Fazeli Nasab B. [Using Morphological and Phytochemical Traits and ITS (1, 4) and rbcl DNA Barcodes in the Assessment of Different Malva sylvestris L. Genotypes]. *J Med plants By-Products.* 2021;**10**(1):19-35. FA. https://doi.org/10.22092/jmpb.2020.343520.1232.

20. Rizwan M. High Mitochondrial Diversity of Apis Mellifera under Coi Gene from China and Pakistan. *Appl Ecol Environ Res.* 2018;**16**(3):2933-45. https://doi.org/10.15666/aeer/1603_29332945.

21. Tavassoli M, Seidi S, Malekifard F. [Molecular Identification and Phylogenetic Analysis of Fleas Ctenocephalides Canis and Pulex Irritans in West and Northwest of Iran Based on Cytochrome Oxidase I]. *Armaghane Danesh.* 2021;**26**(4):511-26. FA. https://doi.org/10.52547/armaghanj.26.4.511.

22. Camilli MP, Kadri SM, Alvarez MVN, Ribolla PEM, Orsi RO. Zinc supplementation modifies brain tissue transcriptome of Apis mellifera honeybees. *BMC Genomics.* 2022;**23**(1):282. [PubMed ID: 35395723]. [PubMed Central ID: PMC8994358]. https://doi.org/10.1186/s12864-022-08464-1.

23. Maiwald F, Haas J, Hertlein G, Lueke B, Roesner J, Nauen R. Expression profile of the entire detoxification gene inventory of the western honeybee, Apis mellifera across life stages. *Pestic Biochem Physiol.* 2023;**192**:105410. [PubMed ID: 37105637]. https://doi.org/10.1016/j.pestbp.2023.105410.

24. Najafzadeh D, Nazemi-Rafie J, Rostamzadeh J. Evaluation of phylogenetic characteristics of dwarf honeybee populations (Apis florea) using COI, intergenic region and COII. *Cell Mol Res (Iran J Biol).* 2019;**32**(2):242-54. FA.

25. Deowanish S, Wattanachaiyingcharoen W, Wongsiri S, Oldroyd BP, Leepitakrat S, Rinderer TE, et al. Biodiversity of dwarf honey bees in

Thailand. *Proc. 7th Int. Conf. Trop. Bees.* Chiang Mai, Thailand. 2001. p. 97-103.

26. Bhat S, Paschapur AU, Subbanna A, Stanley J, Gupta J, Mishra K. Genetic Divergence and Phylogeny of North-Western Indian Himalayan Population of Honey Bee (Hymenoptera: Apidae) Inferred from Mitochondrial DNA Sequences. *Indian J Ecol.* 2022;**49**(4):1440-8. https://doi.org/10.55362/ije/2022/3682.

27. Ren XiaoXiao RX, He XingJiang HX, Gong XueYang GX, Zhao WenZheng ZW, Liu YiQiu LY, Dong Kun DK. Phylogenetic analysis of Apis (Hymenoptera: Apidae) based on COI gene sequences and an evolutionary analysis of its dancing and nesting behavior. *Acta Entomologica Sinica.* 2018;**61**(12):1453-61.

28. Rahimi A, Mirmoayedi A, Kahrizi D, Abdolshahi R, Kazemi E, Yari K. Microsatellite genetic diversity of Apis mellifera meda skorikov. *Mol Biol Rep.* 2014;**41**(12):7755-61. [PubMed ID: 25103023]. https://doi.org/10.1007/s11033-014-3667-7.

29. Beaurepaire AL, Truong TA, Fajardo AC, Dinh TQ, Cervancia C, Moritz RF. Host Specificity in the Honeybee Parasitic Mite, Varroa spp. in Apis mellifera and Apis cerana. *PLoS One.* 2015;**10**(8). e0135103. [PubMed ID: 26248192]. [PubMed Central ID: PMC4527838]. https://doi.org/10.1371/journal.pone.0135103.

30. Munoz I, Henriques D, Jara L, Johnston JS, Chavez-Galarza J, De La Rua P, et al. SNPs selected by information content outperform randomly selected microsatellite loci for delineating genetic identification and introgression in the endangered dark European honeybee (Apis mellifera mellifera). *Mol Ecol Resour.* 2017;**17**(4):783-95. [PubMed ID: 27863055]. https://doi.org/10.1111/1755-0998.12637.

31. Gomeh H, Nazemi Rafei J, Maroufpour M. [Morphological Characteristics Comparison of Population of Honeybees (Apis mellifera L.) in Kermanshah Province]. *J Appl Res Plant Protection.* 2016;**5**(2):187-97. FA.

32. Fatemi SS, Harkinezhad T, Farshine Adl MB, GHaredaghi L. [Morphological diversity of Iranian small honey bee (Apis Florea) population based PCA and minimum variance methods]. *J Livestock Res.* 2017;**6**(1):15-25. FA. https://doi.org/10.22077/jlr.2018.500.1002.

33. Solatani H, Nazemi-Rafie J, Harighi B. [Differentiation of Iranian Honeybees (Apis mellifera meda) from Commercial Honeybee Subspecies using ND1 and ND5 Genes]. *Res Animal Product.* 2021;**12**(31):157-68. FA. https://doi.org/10.52547/rap.12.31.157.

34. Martimianakis S, Klossa-Kilia E, Bouga M, Kilias G. Phylogenetic relationships of Greek Apis mellifera subspecies based on sequencing of mtDNA segments (COI and ND5). *J Apicultural Res.* 2011;**50**(1):42-50. https://doi.org/10.3896/IBRA.1.50.1.05.

35. Ostroverkhova NV, Konusova OL, Kucher AN, Kireeva TN, Vorotov AA, Belikh EA. Genetic diversity of the locus COI-COII of mitochondrial DNA in honeybee populations (Apis mellifera L.) from the Tomsk region. *Russian J Genetics.* 2015;**51**(1):80-90. https://doi.org/10.1134/S102279541501010X.

36. Porrini LP, Quintana S, Brasesco C, Maggi MD, Porrini MP, Garrido MP, et al. Current genetic diversity of managed and commercially produced Apis mellifera colonies in Argentina inferred by wing geometric morphometrics and COI-COII mtDNA locus. *Apidologie.* 2022;**53**(5):61. https://doi.org/10.1007/s13592-022-00970-1.

37. Shakouri M, Ghavam Mostafavi P, Pourkazemi M, Fatemi SM. [Phylogenetic Relationship of Yellowtail Scad (Atule mate) in the Persian Gulf and Oman Sea Using Cytochrome Oxidase I Gene]. *J Animal Biol.* 2022;**14**(3):221-34. FA.

38. Bakker FT, Culham A, Gomez-Martinez R, Carvalho J, Compton J, Dawtrey R, et al. Patterns of nucleotide substitution in angiosperm cpDNA trnL (UAA)-trnF (GAA) regions. *Mol Biol Evol.* 2000;**17**(8):1146-55. [PubMed ID: 10991703]. https://doi.org/10.1093/oxfordjournals.molbev.a026397.

39. Larsson H, Kallman T, Gyllenstrand N, Lascoux M. Distribution of long-range linkage disequilibrium and Tajima's D values in Scandinavian populations of Norway Spruce (Picea abies). *G3 (Bethesda).* 2013;**3**(5):795-806. [PubMed ID: 23550126]. [PubMed Central ID: PMC3656727]. https://doi.org/10.1534/g3.112.005462.