

Molecular characterization of Turkish honey bee populations (*Apis mellifera*) inferred from mitochondrial DNA RFLP and sequence results*

Fulya ÖZDİL¹, Mehmet Ali YILDIZ², H. Glenn HALL³

¹ Department of Animal Science, University of Selçuk, 42075 Konya, Turkey

² Department of Animal Science, University of Ankara, Dışkapı 06110 Ankara, Turkey

³ Department of Entomology and Nematology, University of Florida, Gainesville, Florida, 32611, USA

Received 5 November 2008 – Revised and accepted 16 February 2009

Abstract – To identify the evolutionary lineage of honey bee colonies in Turkey, the mtDNA of 244 colonies from 20 locations was analyzed. Several polymorphic restriction sites showed that they belonged to the Mediterranean C lineage. *DraI* digestion of the *CoxI*–*CoxII* intergenic region produced four fragment patterns, one first seen in this study. From 37 colonies from 16 different locations in Turkey and two colonies from Iran, the intergenic region was sequenced. Previously, from among all honey bee populations of the C lineage, eight haplotypes had been described from this mtDNA region, three of which were found here. In addition, eight new haplotypes were found. A deletion in one of these haplotypes accounts for the novel *DraI* RFLP pattern. Most of the novel haplotypes were in a subgroup of lineage C, yet none of these had been found in previous studies of Turkish honey bees. The geographical distribution of some haplotypes suggests that they may be characteristic of subspecies native to Turkey.

Apis mellifera / mtDNA / DNA sequence / genetic variability / Turkey

1. INTRODUCTION

From morphometric and molecular studies, the 26 subspecies of the honey bee, *Apis mellifera* L., are grouped into five evolutionary lineages: M from northern and western Europe and northern Africa, A from southern and central Africa, C from the northern Mediterranean region and eastern Europe, O from the eastern Mediterranean and the Near and Middle East region, and Y from the east African country of Ethiopia (Ruttner, 1988; Hall and Smith, 1991; Garnery et al., 1992; Arias and Sheppard, 1996; Franck et al., 2000, 2001).

Within Turkey, wide ranges of climates and habitats are found, and several honey bee subspecies and ecotypes have been described (Ruttner, 1988). Except for the northeast and southeast, nearly all of Turkey is occupied by the Anatolian honey bee (*A. m. anatoliaca* Maa). Caucasian (*A. m. caucasica* Gorbachev) and Iranian (*A. m. meda* Skorikov) honey bees are in the northeast and the southeast, respectively. Several ecotypes of the Anatolian honey bee exist, but they are not isolated, and high migratory activity impacts the purity of the ecotypes and the geographical differences among them.

Honey bee lineages can be distinguished by restriction and sequence analyses of the mitochondrial DNA (mtDNA) region between the cytochrome oxidase subunits I and II genes (*CoxI*–*CoxII* intergenic region) (Hall and Smith, 1991; Cornuet et al., 1991; Franck

Corresponding author: F Özgül,
fulyaozdil@selcuk.edu.tr

* Manuscript editor: Walter S. Sheppard

The department where the study was conducted: University of Ankara, Faculty of Agriculture, Biometry & Genetics, 06110 Ankara-Turkey.

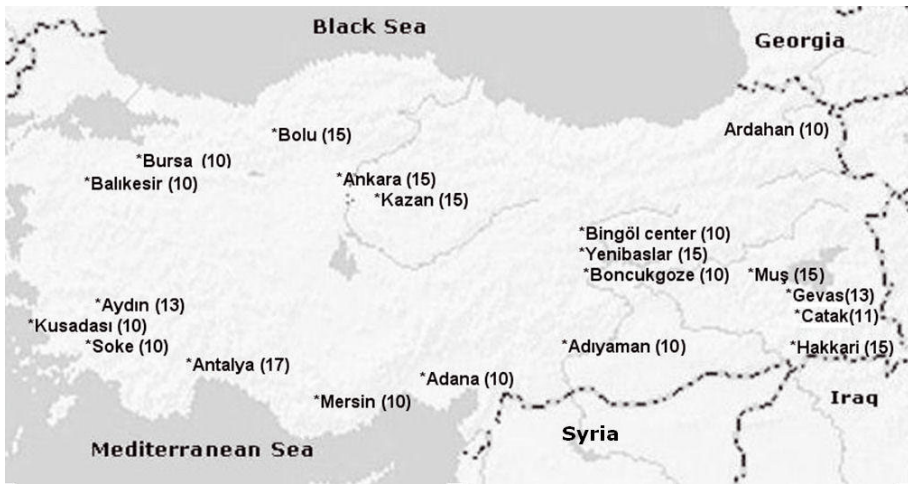


Figure 1. Sampling locations in Turkey. The number of colonies is given within parenthesis.

et al., 2000). *DraI* restriction has revealed more than 50 RFLP patterns in lineages A and M but only five fragment size patterns in lineage C (Franck et al., 2000, 2001; De la Rúa et al., 2004; Sušnik et al., 2004; Kandemir et al., 2006; Kozmus et al., 2007). Restriction fragment patterns of this region tend to be simpler in lineage C, because of the absence of variable repeats found in the other lineages (Hall and Smith, 1991; Garnery et al., 1992, 1993; Franck et al., 2000, 2001; De la Rúa et al., 2004). From nucleotide sequences, six haplotypes have been found in the C lineage of European bees: C1, C2a and C2b in *A. m. ligustica* Spinola, *A. m. carnica* Pollmann and *A. m. caucasica* populations, respectively (Franck et al., 2000); C2c and C2d in *A. m. carnica* and *A. m. macedonica* Ruttner populations, respectively (Sušnik et al., 2004) and C2e in *A. m. carnica* populations from Serbia (Kozmus et al., 2007). The difference between the C1 and C2 groups of haplotypes is a single cytosine base insertion in C1 at position 3428 (Franck et al., 2000). These six haplotypes have a total of three *DraI* fragment patterns. Two additional *DraI* fragment patterns were discovered in honey bees in Turkey (Kandemir et al., 2006).

In this study of Turkish honey bee populations, RFLPs at several mtDNA regions showed that they belonged to the C lineage.

The CoxI-CoxII intergenic region was digested with *DraI*, and this region was sequenced. Novel haplotypes were found which equal the number discovered previously in this mtDNA region, from among all honey bee populations of the C lineage. Some haplotypes may be characteristic of subspecies found within Turkey, but this possibility is debated in light of results obtained by other investigators.

2. MATERIALS AND METHODS

2.1. Sampling and DNA extraction

A total of 244 honey bees each representing a different colony was collected from 20 widespread locations in Turkey (Fig. 1). Two colonies of the *A. m. meda* subspecies came from the west part of Iran, Tabriz and Urmia.

Workers were collected in 95% ethanol and subsequently air-dried. Total DNA was extracted from each bee's thorax according to Hall (1990).

2.2. mtDNA RFLP analysis

From all 244 bees, the following mitochondrial regions were PCR-amplified and digested with the respective enzymes: the large ribosomal subunit (16S rRNA) gene with *EcoRI*, the cytochrome oxidase

subunit I (CoxI) gene with *HincII*, the CoxI–CoxII intergenic region (starting at position 3090), which included part of the CoxI gene, with *XbaI* (Hall and Smith, 1991); the cytochrome-b (Cyt-b) gene with *BglIII* (Crozier et al., 1991). The mtDNA region (starting at position 3363) including the tRNA^{leu} gene, the CoxI–CoxII intergenic region and the 5' end of the CoxII subunit gene was amplified and digested with *DraI* (Garnery et al., 1993). The mtDNA nucleotide positions are from the sequence first obtained by Crozier and Crozier (1993). The fragments were separated by gel electrophoresis: *DraI* digests in 10% acrylamide and the other enzyme digests in 2% agarose.

2.3. Sequence analysis

From 37 colonies from 16 different locations in Turkey and from two colonies from the western part of Iran, the amplified CoxI–CoxII intergenic region, used for *DraI* digestion (previous paragraph), was sequenced on an ABI PRISM 310 automated sequencer using standard protocols. Sequences were aligned with the computer program Clustal X (Thompson et al., 1997).

3. RESULTS

3.1. lrRNA, CoxI–CoxII, Cyt-b and CoxI genes

An *EcoRI* restriction site in the lrRNA gene, an *XbaI* site within the CoxI gene (from the amplified CoxI–CoxII intergenic region which contained part of the CoxI gene) and a *BglIII* site in the Cyt-b gene were present in all our Turkish honey bee samples. *HincII* restriction sites in the CoxI gene were absent. This combination of the presence and absence of restriction sites is characteristic of the C Mediterranean lineage.

3.2. CoxI–CoxII intergenic region

Table I lists the 11 haplotypes found in this study. The designations of the new haplotypes follow those of Franck et al. (2000). Of the other five haplotypes found previously in the C lineage, the four given a “C” designation are

included in the table. The *DraI* restriction fragment sizes and Genbank accession numbers of each haplotype are listed, and the positions of the nucleotide polymorphisms are provided. The numbers of colonies from the different locations that carry the haplotypes are included.

In *A. m. ligustica* mtDNA, *DraI* cuts at three sites within the amplified CoxI–CoxII intergenic region to generate 47, 41, 64 and 420 bp fragments (corresponds to the linear order in the genome) which are characteristic of the C Mediterranean lineage (Garnery et al., 1993; Franck et al., 2000). Three additional haplotypes were seen here, two of which (47, 40, 64, 420 and 47, 39, 64, 420 bp fragments) have been reported in *A. m. carnica*, *A. m. caucasica* and *A. m. anatoliaca* subspecies by others (Garnery et al., 1993; Franck et al., 2000; Kandemir et al., 2006). The other one (47, 41, 64, 419 bp fragments) has not been reported previously. The single nucleotide deletions at positions 3428, 3442 and 3522 (Tab. I) can account for the length differences in the *DraI* fragment patterns found in our samples.

Of the colonies that had the CoxI–CoxII intergenic region sequenced, the mtDNA of 27 from Turkey was assigned to the C1 group (with a cytosine at position 3428), and the mtDNA of the remaining ten from Turkey and two from Iran was assigned to the C2 group (the nucleotide deleted). In this study, eight new haplotypes were found, six within the C1 group and two within the C2 group (Tab. I). Among the eight haplotypes discovered previously, one within the C1 group and two within the C2 group were also found.

In the C1a haplotype from *A. m. ligustica* populations, nucleotides T and C are at positions 3632 and 3767 respectively (Franck et al., 2000). However, 38 of the 39 colonies tested here, whether in the C1 or the C2 groups, had a T→C transition at position 3632 and a C→T transition at position 3767 (Tab. I). Some haplotypes, but only in group C2, had been reported previously to have these nucleotide substitutions (Franck et al., 2000; Sušnik et al., 2004).

One colony from Bolu had the C1a haplotype, but the other 26 colonies, within the C1 group, had one of six haplotypes different from those already described. Fifteen

Table I. Haplotypes, restriction fragment lengths and polymorphic nucleotides of the CoxI–CoxII intergenic region¹.

MtDNA nucleotide positions ²		3428	3442	3467	3514	3522	3567	3575	3587	3632	3662	3767	3769		
Haplotypes	GenBank Accession Numbers	<i>Dra</i> I fragments													Locations (number of colonies sequenced)
C1a ³	FJ357798	47/41/64/420	C	T	C	T	A	A	C	G	T	T	C	C	Bolu (1)
C1b (new)	FJ357799	47/41/64/420	C	T	C	T	A	A	C	G	C	T	T	C	Ten locations (15) ⁷
C1c (new)	FJ357800	47/40/64/420	C	~	C	T	A	T	C	G	C	T	T	C	Adıyaman (5)
C1d (new)	FJ357801	47/41/64/420	C	T	C	T	A	T	C	G	C	A	T	C	Bingöl / Yenibaşlar (2)
C1e (new)	FJ357802	47/41/64/420	C	T	C	T	A	A	C	G	C	C	T	C	Aydın / Center (1)
C1f (new)	FJ357803	47/41/64/420	C	T	C	T	A	A	T	A	C	T	T	C	Bingöl / Center (1)
C1g (new)	FJ357804	47/41/64/419	C	T	C	T	~	A	C	G	C	T	T	C	Aydın / Kuşadası (2)
C2a ³	FJ357805	47/40/64/420	-	T	A	T	A	A	C	G	C	T	T	C	Van / Çatak (2)
C2b ³		47/40/63/421	-	T	C	A	A	A	C	G	C	T	T	C	
C2c ⁴		47/40/64/420	-	T	C	T	A	A	C	G	T	T	C	C	
C2d ⁴		47/40/64/420	-	T	C	T	A	A	C	G	C	T	T	C	
C2e ⁵		47/40/64/420	-	T	C	T	A	A	C	G	C	T	T	C	
C2f ⁶	FJ357806	47/39/64/420	-	~	C	T	A	A	C	G	C	T	T	C	Hakkari (6)
C2g (new)	FJ357807	47/39/64/420	-	~	C	T	A	T	C	G	C	T	T	C	Iran (2) ⁸
C2h (new)	FJ357808	47/40/64/420	-	T	C	T	A	T	C	G	C	T	T	T	Kazan (2)
														Total (39)	

¹In this study bold haplotypes were found, bold fragments and polymorphisms were novel, character “~” shows deletions found. ²Crozier and Crozier, 1993. ³Franck et al., 2000. ⁴Sušnik et al., 2004. ⁵Kozmus et al., 2007. ⁶Kandemir et al., 2006. ⁷Antalya (2), Aydın / Söke (2), Aydın / Center (1), Balıkesir (1), Bolu (1), Bursa (1), Mersin (2), Muş (2), Van / Çatak (1), Van / Gevaş (2). ⁸Iranian samples were collected from Tabriz and Urmia.

colonies, from 10 locations, had two base substitutions at positions 3632 and 3767 (haplotype C1b) (Tab. I). Additional nucleotide differences gave rise to the remaining five new C1 haplotypes.

A deletion at position 3442 was found in all five colonies from Adıyaman (haplotype C1c), in all six colonies from Hakkari (haplotype C2f) and in two samples from Iran (haplotype C2g). Two colonies from Ankara-Kazan had base substitutions at positions 3567 and 3769 (haplotype C2h), similar to published sequences of *A. m. caucasica* (Garnery et al., 1992).

4. DISCUSSION

Several major studies of Turkish honey bee populations have been conducted (Smith et al., 1997; Palmer et al., 2000; Kandemir et al., 2006). Our results coincide with some findings from the previous work. However, we discovered a significant number of novel haplotypes along with some of the other haplotypes reported previously within the CoxI-CoxII intergenic region, which provided contrasting results in the frequencies and distributions of different mtDNA types within Turkey.

With restriction digests, nearly all Turkish colonies analyzed previously by the other investigators and all colonies in this study were found to belong to the C Mediterranean lineage. This finding was expected because the Turkish populations consist mainly of *A. m. anatoliaca* and *A. m. caucasica*. A few colonies belonging to the O lineage were found in the southernmost province of Hatay near the Syrian border (Palmer et al., 2000; Kandemir et al., 2006), an area we did not sample.

As in the other studies, we utilized the mitochondrial CoxI-CoxII intergenic region, because it is particularly valuable in distinguishing honey bee lineages. Palmer et al. (2000) obtained sequences of this region. Kandemir et al. (2006) identified haplotypes based on *DraI* RFLPs. We identified haplotypes by both approaches, from *DraI* restriction and from nucleotide sequences. The sequences from our samples were very similar to those pub-

lished of *A. m. caucasica* and *A. m. anatoliaca* (Garnery et al., 1992, 1993; Franck et al., 2000).

Franck et al. (2000) described the different haplotypes of the CoxI-CoxII intergenic region found in the major honey bee lineages. We have followed the pattern of haplotype designations that they established. The C lineage is further divided into C1 and C2 groups by the presence of a cytosine at position 3428 or its absence. This nucleotide was first found to be absent in Caucasian bee samples (Garnery et al., 1992). Previously, from all honey bee populations of the C lineage, only one haplotype from within the C1 group and seven from within the C2 group had been described (Franck et al., 2000; Sušnik et al., 2004; Kandemir et al., 2006; Kozmus et al., 2007). The one from the C1 group and two from the C2 group were found in this study. Additionally, eight new haplotypes were found, six within the C1 group and two within the C2 group (Tab. I). Seven of the novel haplotypes were found among our Turkish samples and one in both Iranian samples. All of the Turkish samples studied by Palmer et al. (2000) belonged to the C2 group. From the CoxI-CoxII intergenic region, they broadly categorized three mtDNA types: Anatolian, Caucasian and Syrian. From this mtDNA region, Kandemir et al. (2006) found four RFLP patterns in their Turkish bee samples that all belonged to the C2 group. Two patterns were novel, and two (corresponding to C2a and C2b) had been identified previously in other honey bee populations. In our study, we also found C2a and one of the two patterns that they had originally found (TrDra-1). To be consistent with Franck et al., 2000, we designated the haplotype with this RFLP pattern as C2f.

Kandemir et al. (2006) found the RFLP pattern TrDra-1 (corresponding to C2f) to be the most common, present in 80 to 100% of colonies in the different regions. We identified C2f in only six colonies in Hakkari, near the Iraqi border, an area where they did not sample. Haplotype C1b, not found in previous honey bee studies, was the most common haplotype in our samples, present in about 40%. The C1 haplotypes as a group were present in 73%. It is puzzling that we did not find C2f to

be more widespread and that other investigators did not find C1 *DraI* RFLP patterns.

These differences seen in the distribution and abundance of the haplotypes could be due to problems in haplotype identification. Although all our 244 colony samples were digested with *DraI*, only from the 39 sequenced could we be certain of fragment sizes that differed by only a single nucleotide. Even if fragment lengths can be established with certainty, the identities of many of the mitochondrial haplotypes cannot be determined by RFLP patterns alone. For example, the TrDra4 haplotype, found by Kandemir et al. (2006), corresponding to the previously found C2a, has the same pattern as four other haplotypes distinguished by their sequences. Kandemir et al. (2006) submitted sequences to Genbank representing each of the four haplotypes found within the CoxI-CoxII intergenic region, but this is not how the population samples were analyzed. As mentioned above, they found haplotype TrDra1 to be most abundant and widespread, but we found that fragment pattern (C2f) to be more rare and localized.

According to Ruttner's (1988) morphometric studies, the borders drawn for Iranian bees include the cities of Van, Bingöl, Muş, Adıyaman and Hakkari. The nucleotide deletion at position 3442 (haplotypes C1c, C2f and C2g) was present in all the samples from Adıyaman, Hakkari, and Iran, and only in these samples. Hakkari is in the extreme southeast part of Turkey near the Iranian and Iraqi border. This distribution suggests that this deletion may be characteristic of the Iranian subspecies *A. m. meda*. However, this possibility would not be true if the TrDra-1 RFLP pattern found in the more widespread colonies by Kandemir et al. (2006) were identical to haplotype C2f. We found haplotype C1b only in regions occupied by *A. m. anatoliaca* populations, but common and widespread within those areas. Perhaps, this haplotype is characteristic of this subspecies, but the failure of these other investigators to detect C1b, after extensive sampling, has yet to be explained. Clearly, much remains to be learned about the honey bee populations of Turkey, a country that is obviously rich in honey bee mtDNA variability.

ACKNOWLEDGEMENTS

This work was performed by Fulya Özdil in partial fulfillment of the Ph. D. degree in Biometry and Genetics at University of Ankara, Turkey. The authors would like to thank Dr. Çetin Fıratlı (Ankara University), Dr. H. Vasfi Gençer (Ankara University) and Dr. Cengiz Erkan (Yüzüncü Yıl University) for providing honey bee samples and Raquel McTiernan, Hasan Meydan and Yasemin Gedik for their help with the laboratory work. This project was supported by three grants: DPT-2003K12019015-5 from T.R. Prime Ministry State Planning Organization (Project coordinator: Dr. M.A. Yıldız), TUBITAK TOVAG-103V070/VHAG-2046 and TUBITAK TOVAG-108O200 from the Scientific and Technological Research Council of Turkey (Project coordinator: Dr. M.A. Yıldız).

Caractérisation moléculaire des populations d'abeilles (*Apis mellifera*) turques à l'aide du RFLP-ADN et du séquençage.

génétique population / variabilité génétique / ADN mitochondrial / séquence ADN / Turquie

Zusammenfassung – Molekulare Charakterisierung türkischer Honigbienenpopulationen (*Apis mellifera* L.) mittels mitochondrialer DNA RFLP und Sequenzierung. Die mitochondriale DNA (mtDNA) von 244 Honigbienenvölkern aus 20 Regionen der Türkei wurde hinsichtlich der folgenden Restriktionsprofile untersucht: (1) *XbaI* in der intergenen Region der Cytochromoxidase-Untereinheiten I und II (CoxI–CoxII), die einen Teil der CoxI Gene enthält; (2) *BglIII* im Cytochrom-b Gen (Cyt-b); und (3) *EcoRI* im Gen der grossen ribosomalen RNA Untereinheit (18S rRNA). Jeweils eine *BglIII*, *EcoRI* und *XbaI* Schnittstelle lagen in den Cyt-b, 18S rRNA und CoxI Genen. Eine *HincII* Schnittstelle fehlte im CoxI Gen. Diese Polymorphismen sind charakteristisch für die mediterrane Linie (C) der Unterarten der Honigbiene. Für die intergene Region CoxI–CoxII fanden wir auch einen Verdauungsmittel *DraI*, wobei die vier Restriktionsfragmentmuster aus Längenunterschieden von eins bis zwei Basenpaaren in zwei der vier Fragmente resultierten. Drei dieser Muster waren bereits für *A. m. ligustica*, *A. m. carnica* und *A. m. caucasica* beschrieben worden und eine für Bienen aus der Türkei. Des weiteren fanden wir ein neues in dieser Untersuchung erstmals beschriebenes Muster. Für Bienen von 37 Völkern aus 16 unterschiedlichen Regionen der Türkei und von zwei Völkern aus dem westlichen Iran wurde die CoxI–CoxII intergene Region sequenziert. Die Sequenzen zeigten

eine hohe Ähnlichkeit mit veröffentlichten Sequenzen von *A. m. caucasica* und *A. m. anatoliaca*. Von den bisher für Populationen der C-Linie beschriebenen acht Haplotypen in dieser mtDNA-Region fanden wir in unserer Untersuchung drei. Ausserdem fanden wir acht neue Haplotypen. Eine Deletion in einem dieser Haplotypen charakterisiert das neue *DraI* RFLP-Muster. Die geographische Verteilung der Haplotypen last darauf schliessen, dass sie entweder für *A. m. anatoliaca* oder *A. m. meda* charakteristisch sind.

***Apis mellifera* / mtDNA / DNA Sequenz / genetische Variabilität / Türkei**

REFERENCES

- Arias M.C., Sheppard W.S. (1996) Molecular phylogenetics of honey bee subspecies (*Apis mellifera* L.) inferred from mitochondrial DNA sequence, *Mol. Phylogenet. Evol.* 5, 557–566.
- Cornuet J.M., Garnery L., Solignac M. (1991) Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA, *Genetics* 1128, 393–403.
- Crozier Y.C., Koulianos S., Crozier R.H. (1991) An improved test for Africanized honeybee mitochondrial DNA, *Experientia* 47, 968–969.
- Crozier R.H., Crozier Y.C. (1993) The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization, *Genetics* 133, 97–117.
- De la Rúa P., Hernández-García R., Pedersen B.V., Galián J., Serrano J. (2004) Molecular diversity of honeybee *Apis mellifera iberica* L. (Hymenoptera: Apidae) from western Andalusia, *Archiv. Zootec.* 53, 195–203.
- Franck P., Garnery L., Solignac M., Cornuet J.-M. (2000) Molecular confirmation of a fourth lineage in honeybees from the near east, *Apidologie* 31, 167–180.
- Franck P., Garnery L., Loiseau A., Oldroyd B.P., Hepburn H.R., Solignac M., Cornuet J.-M. (2001) Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data, *Heredity* 86, 420–430.
- Garnery L., Cornuet J.-M., Solignac M. (1992) Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis, *Mol. Ecol.* 1, 145–154.
- Garnery L., Solignac M., Celebrano G., Cornuet J.-M. (1993) A simple test using restricted PCR amplified mitochondrial DNA to study the genetic structure of *Apis mellifera*, *Experientia* 49, 1016–1021.
- Hall H.G. (1990) Parental analysis of introgressive hybridization between African and European honeybees using nuclear DNA RFLPs, *Genetics* 125, 611–621.
- Hall H.G., Smith D.R. (1991) Distinguishing African and European honeybee matrilineages using amplified mitochondrial DNA, *Proc. Natl. Acad. Sci. USA* 88, 4548–4552.
- Kandemir İ., Kence M., Sheppard W.S., Kence A. (2006) Mitochondrial DNA variation in honey bee (*Apis mellifera* L.) populations from Turkey, *J. Apic. Res. Bee World* 45, 33–38.
- Kozmus P., Stevanović J., Stanimirović Z., Stojić V., Kulišić Z., Meglič V. (2007) Analysis of mitochondrial DNA in honeybees (*Apis mellifera*) from Serbia, *Acta Vet. (Beograd)* 57, 465–476.
- Palmer M.R., Smith D.R., Kaftanoğlu O. (2000) Turkish honeybees: Genetic variation and evidence for a fourth lineage of *Apis mellifera* mtDNA, *J. Hered.* 91, 42–46.
- Ruttner F. (1988) *Biogeography and Taxonomy of Honeybees*, Springer-Verlag, Berlin, 284 p.
- Smith D.R., Slaymaker A., Palmer M., Kaftanoğlu O. (1997) Turkish honeybees belong to the east Mediterranean mitochondrial lineage, *Apidologie* 28, 269–274.
- Sušnik S., Kozmus P., Poklukar J., Meglič V. (2004) Molecular characterization of indigenous *Apis mellifera carnica* in Slovenia, *Apidologie* 35, 623–636.
- Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res.* 25, 4876–4882.