

## Phylogeography and Population Genetics of Honey Bees (*Apis mellifera*) From Turkey Based on COI-COII Sequence Data

by

Cesar D. Solorzano<sup>1</sup>, Allen L. Szalanski<sup>1\*</sup>, Meral Kence<sup>2</sup>, Jackie A. McKern<sup>1</sup>,  
James W. Austin<sup>3</sup> & Aykut Kence<sup>2</sup>

### ABSTRACT

A study that involved DNA sequencing of COI-COII intergenic region of the mitochondrial DNA genome of *Apis mellifera* honey bees from Turkey was conducted to determine the population genetics and phylogeographic structure of this species from seven distinct areas of Turkey. From the 132 honey bees subjected to DNA sequencing, a total of 12 mitotypes of *A. mellifera* "C" lineage were observed, of which only one mitotype, C13, had been reported previously. The most common mitotype, C12, accounted for 47% of the *Apis mellifera* "C" lineage samples and was found in 13 of the 22 sampled locations. This mitotype was also the basal ancestral mitotype based on TCS spanning tree analysis. The greatest amount of genetic diversity was observed in Bursa, where 4 mitotypes of the *A. mellifera* "C" lineage were unique to this location. Wright's F-statistics revealed that Artvin and Bursa were the most genetically distinct locations relative to the other sampled locations. Applying a molecular clock, Turkish *A. mellifera* "C" lineage mitotypes have been diverging for approximately 10,000 to 16,500 yr. based on phylogenetic analysis. In addition, two *A. m. syriaca* samples were observed from Hatay, Turkey. Phylogenetic analysis which included other *A. mellifera* subspecies confirms the subspecies relationships of *A. mellifera* "C" lineage, and *A. m. syriaca*. This study corroborates other studies that show Turkey to be a reservoir of genetically distinct populations of *A. mellifera* "C" lineage, which can be useful for developing genetic conservation strategies for *A. mellifera*.

Keywords: *Apis mellifera*, COI-COII intergenic region, genetic variation, Turkey, phylogeography

---

<sup>1</sup>Department of Entomology, University of Arkansas, Fayetteville, AR, USA 72701

<sup>2</sup>Department of Biology, Middle East Technical University, Ankara, Turkey 06531.

<sup>3</sup>Department of Entomology, Texas A&M University, College Station, TX, USA 77843.

\*Corresponding author. email:aszalan@uark.edu.

## INTRODUCTION

*Apis mellifera* L. is native to Europe, Africa, and Asia (including Saudi Arabia, Iran and the Ural Mountains of Russia). Currently the species is widely distributed across the world due to multiple migrations and introductions (Ruttner 1988). *A. mellifera* includes about two dozen subspecies that are present in different regions of the world. These subspecies have been classified into four main lineages: C (Carnica group that includes *A. m. carnica* and *A. m. ligustica*); M (North and Western European honey bees that include *A. m. mellifera*, *A. m. iberica*, and *A. m. intermissa*); A (African group that includes *A. m. scutellata*, *A. m. capensis*, *A. m. lamarckii*, *A. m. litorea*, *A. m. adansonii*, and *A. m. unicolor*); and the O group (Oriental group which groups *A. m. anatolica*, *A. m. caucasica*, *A. m. syriaca*, *A. m. pomonella*, and *A. m. cypria*) (Ruttner 1992). At the molecular level, these subspecies are genetically divergent based on nuclear and mitochondrial DNA sequence data. They each have specific behavioral and morphometrical characteristics, and can interbreed because they are the same species.

In the Old World, studies of honey bee genetic diversity have been conducted in Turkey (Kandemir *et al.* 2006, Bodur *et al.* 2007), Africa (Franck *et al.* 2001), and Australia (Chapman *et al.* 2008). It is particularly important to study their genetic diversity to determine patterns of migration, selection and diversification at local levels. Several studies have generated mitochondrial DNA (mtDNA) information that has provided insights into subspecies inhabiting particular regions. In a study conducted on seven regions of Turkey, according to *Dra I* profiles of the COI-COII region, pattern C predominated (97.9%) which corresponds to *A. m. carnica* / *A. m. ligustica* (Kandemir *et al.* 2006). Profile A was observed in Hatay in 6 of 24 colonies (1.8%) which corresponded to the African subspecies (Kandemir *et al.* 2006); however, this study was limited by the fact that only one restriction enzyme was used, which reduced the amount of genetic variation that could be detected.

The objective of our study was to determine the population genetic and phylogeographic structure extent of honey bees from Turkey based on COI-COII mtDNA sequence data, and to determine if there are any genetically distinct populations which may be used to develop genetic conservation strategies of *A. mellifera*.

## MATERIALS AND METHODS

Adult worker specimens were collected from seven regions of Turkey and stored in 100% ethanol until processed for DNA extraction (Table 1, Fig. 1). Voucher specimens are deposited at the Arthropod Museum, Department of Entomology, University of Arkansas, Fayetteville, AR, USA.

Genomic DNA from individual honey bee thoraces was extracted using the Qiagen DNeasy extraction kit (QIAGEN, Valencia, CA) according to the manufacturer's protocol and per Szalanski & McKern (2007). A 570 to 640 bp region of the COI-COII intergenic region was PCR amplified in a Techne T-412 thermal cycler (Techne Inc, Burlington, NJ) using primers E2 and H2 (Garney *et al.* 1993). PCR was conducted following the profile of 35 cycles of 94°C for 45 s, 46°C for 45s, and 72°C for 45s. Amplicons were separated in 1% agarose gel electrophoresis and photo documented using a BioDoc-It™ Imaging System (UVP, Inc., Upland, CA). PCR products were purified using Microcon-PCR Filter Units (Millipore, Bedford, MA), and sent to the University of Arkansas Medical Sciences DNA Sequencing Core Facility (Little Rock, AR) for direct sequencing in both directions. Sequences new to this study were deposited in GenBank as accession numbers FJ037776 to FJ037787.

DNA sequences were aligned with CLUSTAL W (Thompson *et al.* 1994) using Bioedit v5.0.7 (Hall 1999). Number of mitotypes and their frequencies were determined both visually and with the program DNAsp version



Fig. 1. Sampling locations in seven geographical regions of Turkey for *Apis mellifera* "C" lineage, and *A. m. syriaca* and the number of colonies (n) from each sample location.

4.10.9 (Rozas *et al.* 2003). DNAsp also estimated the following variables: haplotypic diversity ( $Hd$ ) and its variance, equation 8.4 and 8.12 of Nei (1987), nucleotide diversity ( $\pi$ ) and its variance, equation 10.5 and 10.7 of Nei (1987), Nei's  $Nm$  and Wright's  $F_{ST}$  values. Nucleotide diversity was interpreted as the average proportion of nucleotide differences between all possible pairs of sequences in the sample (Hartl & Clark 1997), mean number of pairwise nucleotide differences ( $K$ ) equation A3 (Tajima 1983), number of polymorphic sites ( $S$ ), and the parameters  $\theta_s$  and  $\theta_g$ . The parameter  $\vartheta$  is the proportion of nucleotide sites that are expected to be polymorphic in any suitable sample from this region of the genome (Hartl & Clark 1997). To test for neutral mutation, Tajima's  $D$  (Tajima, 1989), and Fu and Li's  $D^*$  and  $F^*$  (Fu & Li 1993) were calculated. Tajima's  $D$  was calculated using the value of  $\theta_s$  based on the number of segregating sites. To examine demographic stability, Fu's  $F_s$  statistic (based on mitotype distribution) was used.

Levels of gene flow were determined through the effective number of migrants ( $Nm$ ) between locations in DNAsp using Nei (1982). Genealogy of mitotypes was determined using the method of Templeton *et al.* (1992), which represents the evolutionary steps between mitotypes, using TCS v 1.21 (Clement *et al.* 2000). The distance matrix option of PAUP\* 4.0b10 (Swofford, 2002) was used to calculate genetic distances according to the Tajima-Nei model (Tajima & Nei 1984).

## RESULTS AND DISCUSSION

Twelve mitotypes of *A. mellifera* "C" lineage were found throughout the six regions of Turkey sampled, and were designated as C12-C24 (Tables 1-2). In addition one mitotype of *A. m. syriaca* was observed from Hatay. Among the 12 "C" lineage mitotypes, a total of 15 nucleotide sites were polymorphic (Table 2). The most abundant mitotype was C12 (49%) and occurred in 12 locations (Table 1). Two mitotypes, C15 and C24, only occurred once. All of the mitotypes observed in this study have not been previously reported with the exception of mitotype C13, which is present in the Marmara, Black Sea, and Central Anatolia regions, and was previously reported as *TrDra-3* by Kandemir *et al.* (2006).

Genetic divergence among mitotypes was calculated and ranged from 0.24 to 0.38%. Based on a mtDNA molecular clock value of 2.3% genetic

divergence per million years (Brower 1994), the divergence time among mitotypes ranged from 10,000 yr to 16,521 yr. Haplotypic diversity was highest in Gökçeada and Bursa (Table 1). Nucleotide diversity ( $\pi$ ) was also low, and so were the average number of pairwise nucleotide differences ( $K$ ), and theta per site and per gene, whose values were the same, therefore they are represented in the same column (Table 3).

The hypothesis of neutral mutation that was tested with multiple statistics, where populations have a constant size, proved true and therefore no assumptions on population growth or selection can be made (Table 3). Gene flow calculated thru the  $N_m$  was highest between Sakarya and Kirklareli, and between Hiresun, Sakarya and Bursa (Table 4). Wright's  $F_{ST}$  statistic revealed that Artvin is the most genetically distinct population, followed by Bursa and Sakarya (Table 4). This gives genetic support that these populations need to be considered for conservation of honey bee genetic diversity in Turkey.

Table 1. Sampling locations and mitotype frequencies (n) of *Apis mellifera* from Turkey.

Region (n)	Location	Mitotype (n)
Marmara (79)	Kirklareli	C12(10)
	Canakkale(Gökçeada)	C12(3), C13(4), C14(4)
	Bursa	C12(1), C18(17), C19(3), C20(3),
	Sakarya	C21(4)
Aegean (2)	Izmir	C12(1)
	Mugla	C12(1)
	Black Sea (26)	Giresun
Artvin		C11(16)
Duzce		C12(1)
Central Anatolia (2)	Ankara	C19(1)
	Sivas	C13(1)
South East Anatolia (7)	Diyarbakir	C24(1)
	Silvan	C12(1)
	Batman	C19(1)
	Bitlis	C11(1)
	Sirnak	C12(1)
	Gaziantep	C19(1)
	Adiyaman	C12(1)
Medditerranean (13)	Hatay	C11(1), C12(8), C22(2), ams1(2)
Eastern Anatolia (2)	Erzincan	C11(1)
	Hakkari	C11(1)

Table 2. Variable nucleotide sites and mitotype frequency (n) from 12 mitotypes of *A. mellifera* "C" lineage from Turkey. \*From Kandemir *et al.* (2006)

Mitotype	85	105	152	181	205	209	211	212	213	214	215	217	225	300	454	n	GenBank #
C11	-	C	T	T	T	A	A	A	C	A	A	A	G	T	A	19	FJ037776
C12	-	.	.	.	A	.	.	.	.	.	.	.	.	.	.	64	FJ037777
C13	-	.	A	.	A	.	.	.	.	.	.	.	.	.	.	6	AY618915*
C14	-	.	.	.	A	.	.	.	.	.	.	.	.	C	.	4	FJ037778
C15	-	A	.	.	A	.	.	.	.	.	.	.	.	.	.	1	FJ037779
C17	-	.	.	.	A	.	.	.	.	.	.	.	.	.	G	4	FJ037780
C18	-	.	.	.	A	.	T	.	.	.	.	.	.	.	.	17	FJ037781
C19	-	.	.	.	A	.	.	.	.	.	.	.	A	.	.	6	FJ037782
C20	-	.	.	.	A	.	.	.	.	.	.	.	.	.	.	3	FJ037783
C21	-	.	.	.	A	.	.	.	.	.	.	.	.	.	.	4	FJ037784
C22	A	.	.	.	A	.	T	.	.	.	.	.	.	.	.	8	FJ037785
C24	-	.	.	.	A	T	.	.	.	.	.	.	.	.	.	1	FJ037786

TCS spanning network of mitotypes revealed that mitotype C12 was the basal mitotype, which was only one base pair different from eight other mitotypes (Fig. 2). Mitotype C12 can be considered as the basal ancestral mitotype because it is located at the base of the genealogy. Mitotype C21 was the most distantly related to C12, with five base pair differences, and only occurred in Bursa (Marmara region). Furthermore, Palmer *et al.* (2000) suggest that the same restriction digest patterns were not observed from Anatolia proper, implicating some geographic isolation between northwest Anatolia and the rest of the country. This seemingly geographic isolation may act as natural preservation of the diversity of honeybees throughout western Anatolia.

The large amount of mtDNA mitotype genetic diversity of Turkish honey bees that we found has been observed in microsatellite analysis of honey bees from Turkey as well (Bodur *et al.* 2007). They found that the populations from the Black Sea (Artvin)

Table 3. Summary statistics for mtDNA polymorphisms in *Apis mellifera* from seven locations in Turkey.

Location	N	S	H	Hd ± sd	π	K	ϕ <sub>s</sub>	D <sup>+</sup>	F <sup>+</sup>	F <sub>s</sub> <sup>+</sup>	S	D
Arvin	16	0	1	0.00±0.00	0.00	0.00	0.00	-	-	-	-	-
Haray	11	3	3	0.473±0.162	0.00089	0.836	0.00120	-0.33034	-0.49428	-0.659	0.90	-0.79
Giresun	9	0	1	0.00±0.00	0.00	0.00	0.00	-	-	-	-	-
Gökçeada	11	2	3	0.727±0.068	0.00178	1.018	0.00120	0.99697	1.25816	0.623	0.668	1.50
Sakarya	30	5	5	0.405±0.103	0.00065	0.505	0.00133	-1.47512	-1.61523	-1.885	0.967	-1.19
Bursa	28	10	5	0.608±0.092	0.00035	3.013	0.00046	0.60275	0.38831	0.098	0.843	-0.36
Kirklareli	10	0	1	0.00±0.00	0.00	0.00	0.00	-	-	-	-	-
Total	130	14	12	0.71±0.01	0.01	1.20	0.54	-0.87	-1.12	-3.18	0.98	-1.10

NOTE: N = number of sequences; S = number of polymorphic sites; H = number of mitotypes; Hd = nucleotide diversity; π = nucleotide diversity; K = mean number of pairwise nucleotide differences; ϕ<sub>s</sub> = theta per site and ϕ<sub>s</sub> = theta per gene were the same; D<sup>+</sup> and F<sup>+</sup> statistics (Fu and Li 1993); F<sub>s</sub><sup>+</sup> statistic; Strobeck S statistic (Strobeck 1987); D = Tajima's (1989) statistic; \* p > 0.10 for all values.

and Marmara (Kirklareli) had the lowest levels of gene flow based on pairwise Nm values. Our mtDNA results support the genetic distinctiveness of the Black Sea populations. However, unlike their findings, our population from Kirklareli was not genetically distinct, and all sampled colonies consisted of mitotype C12. This discrepancy between mitochondrial and nuclear markers may be reflective of the maternal inheritance of mitochondrial DNA, and the Kirklareli population may be distinct for nuclear markers due to the polyandrous matings of queens with genetically distinct drones.

The only other studies on mtDNA genetic diversity of honey bees in Turkey was conducted by Kandemir *et al.* (2006), and Palmer *et al.* (2000). The Kandemir *et al.* (2006) study involved PCR-RFLP analysis of the COI-COII marker using the restriction enzyme *Dra* I. A total of 7 different *Dra* I digest profiles were observed, with two profiles accounting for over 90% of the observed genetic diversity. In addition, individuals of each profile were subjected to DNA sequencing and only one profile TrDra-3 (GenBank AY618915) corresponded to one of our mitotypes, C13. The discrepancy between our results and theirs can be attributed to the greater sensitivity of DNA sequencing to detect genetic variation, their use of only one restriction enzyme for PCR-RFLP, and the fact that only 3

of our 7 sampled populations overlapped with theirs. The study by Palmer *et al.* (2000) involved PCR-RFLP of the COI-COII region of 84 *A. mellifera* colonies from across Turkey. Combined restriction site and sequence data of the non-coding region revealed four distinct mitotypes. Three of the mitotypes were of the eastern Mediterranean lineage while the fourth one that they observed from bees collected near Hatay, represented a new lineage within the range of *A. m. syriaca*. This agrees with our finding of *A. m. syriaca* from the same geographical area.

This study has provided useful insights into the levels of genetic variation and the phylogenetic relationships of honey bee populations from Turkey based on mtDNA COI-COII sequence data. Other studies based on morphology have been less revealing (Güler & Kaftanoğlu 1999). Our findings support

Table 4. Gene flow ( $N_m$ ) of *Apis mellifera* from seven locations in Turkey calculated with Nei (1982) (below axis), and observed  $F_{ST}$  values (above axis).

Location	Artvin	Hatay	Giresun	Gökçeada	Sakarya	Bursa	Kirklareli
Artvin	--	0.67	1.00	0.71	0.80	0.53	1.00
Hatay	0.23	--	0.08	0.22	0.07	0.19	0.08
Giresun	0.00	6.22	--	0.30	0.07	0.08	0.00
Gökçeada	0.31	2.27	1.94	--	0.23	0.27	0.30
Sakarya	0.25	8.05	19.50	0.00	--	0.07	0.06
Bursa	0.13	11.50	17.13	1.78	8.86	--	0.15
Kirklareli	0.00	5.88	0.00	1.84	18.00	15.83	--

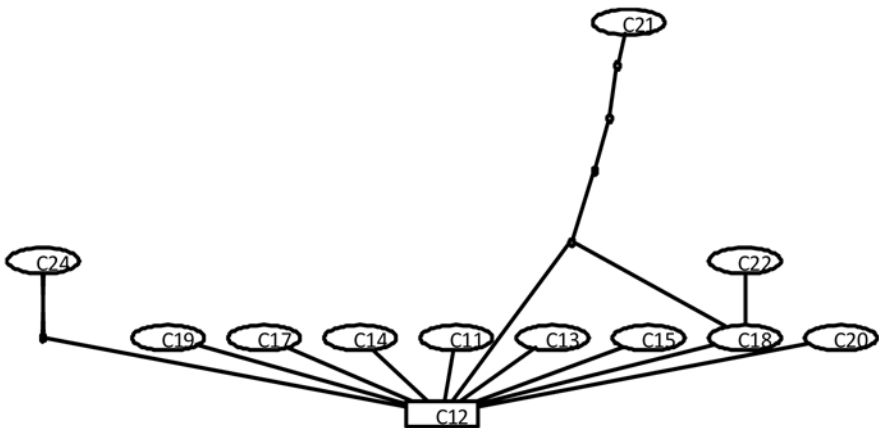


Fig. 2. Genealogical relationships among mitotypes of *Apis mellifera* “C” lineage from Turkey estimated by TCS (Clement *et al.* 2000). A unit branch represents one mutation and small ovals indicate mitotypes that were not observed.



previous studies by Bodur *et al.* (2007), implicating Turkey as a resource of genetic diversity of *A. mellifera* and implores the need for conserving these genetic resources. Fragmentation of available land resources through urbanization, current apicultural migratory habits of honey producers, and the identification of genetically distinct populations which demonstrate high degrees of genetic diversity must be considered for future genetic preservation of this group in Turkey.

## ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Ministry of Agriculture and Rural Affairs of Turkey, beekeepers and colleagues for their assistance in the collection of honey bee specimens in Turkey. This research was supported in part by the University of Arkansas, Arkansas Agricultural Experiment Station, the Center for Urban & Structural Entomology, Texas A&M University, and the Middle East Technical University, Department of Biology, Ankara, Turkey.

## REFERENCES

- Bodur, C., M. Kence & A. Kence. 2007. Genetic structure of honeybee, *Apis mellifera* L. (Hymenoptera: Apidae) populations of Turkey inferred from microsatellite analysis. *J. Apicul. Res.* 46: 50-56.
- Brower, A.V.Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl. Acad. Sci. USA* 91: 6491-6495
- Chapman, N.C., J. Lim & B.P. Oldroyd. 2008. Population genetics of commercial and feral honey bees in western Australia. *J. Econ. Entomol.* 101: 272-277.
- Clement, M., D. Posada & K.A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. *Molec. Ecol.* 9: 1657-1660.
- Franck, P., L. Garnery, A. Loiseau, B.P. Oldroyd, H.R. Hepburn, M. Solignac & J.-M. Cornuet. 2001. Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. *Heredity* 86: 420-430.
- Fu, Y.X. & W.H. Li. 1993. Statistical tests of neutrality of mutations. *Genetics* 133: 693-709.
- Garnery, L., M. Solignac, G. Celebrano & J.M. Cornuet. 1993. A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic structure of *Apis mellifera* L. *Experientia* 49: 1016-1021.
- Güler, A. & O. Kafanoğlu. 1999. Türkiye'deki önemli Balarısı (*Apis mellifera* L.) irk ve ekotiplerinin morfolojik özellikleri-I. *Tr. J. Vet. Anim. Sci.* 23: 565-570.

- Hall, T.A. 1999. Bioedit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41: 95-98.
- Hartl, D.L. & A.G. Clark. 1997. Principles of Population Genetics. Sunderland, MA, Sinauer Associates, Inc.
- Kandemir I., M. Kence, W.S. Sheppard & A. Kence. 2006. Mitochondrial DNA variation in honey bee (*Apis mellifera* L.) populations from Turkey. J. Apicul. Res. 45:33-38.
- Nei, M. 1987. Molecular Evolutionary Genetics, Columbia University Press, New York.
- Palmer, M. R., D. R. Smith & O. Kaftanoglu. 2000. Turkish Honeybees: Genetic variation and evidence for a fourth lineage of *Apis mellifera* mtDNA. Heredity 91: 42-46.
- Rozas, J., J.C. Sanchez-DelBarrio, X. Messeguer & R. Rozas. 2003. DnaSP, DNA polymorphism analyzes by the coalescent and other methods. Bioinformatics 19: 2496-2497.
- Ruttner, F. 1988. Biogeography and Taxonomy of Honey bees. Springer-Verlag; Berlin Heidelberg, Germany, 284pp.
- Ruttner, F. 1992. Naturgeschichte der Honigbienen, Ehrenwirth, Munich.
- Swofford, D.L. 2002. PAUP: Phylogenetic analysis using parsimony (and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Szalanski, A.L. & J.A. McKern. 2007. Multiplex PCR-RFLP diagnostics of the Africanized honey bee (Hymenoptera: Apidae). Sociobiol. 50: 939-945.
- Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105: 437-460.
- Tajima, F. & Nei, M. 1984. Estimation of evolutionary distance between nucleotide sequences. Mol. Biol. Evol. 1: 269-285.
- Tajima, F. 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.
- Templeton, A.R., K.A. Crandall & C.F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132: 619-633.
- Thompson, J.D., D.G. Higgins & T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

