Genetic Evidence for Honey Bees (*Apis mellifera* L.) of Middle Eastern Lineage in the United States

by

Roxane Magnus & Allen L. Szalanski*

ABSTRACT

Honey bees, *Apis mellifera* L. are the principle managed pollinator of agriculture and horticulture crops in the United States. *Apis mellifera* is not native to the United States and the first record of this species in the United States was during the early to mid 17th century when European settlers brought it to the United States. The mitochondrial DNA COI-COII intergenic region of *A. mellifera* exhibits a high degree of genetic variability within and among *A. mellifera* lineages and is useful for differentiating lineages as well as detecting unique mitotypes. We conducted a study of the genetic diversity of honey bees from central and south central United States from primarily feral populations. Of the 469 samples from 14 states subjected to DNA sequencing we found evidence of four mitotypes from the ‘O’ lineage: O5, O5d, O5”b, and O2. Only one of these mitotypes, O2, has been previously observed (in Lebanon). Within the feral population, this lineage accounted for 5% of the observed mitotypes. Of the 24 ‘O’ lineage samples mitotype O5 was the most common and accounted for 52% of the total observed ‘O’ mitotypes. Bayesian and maximum parsimony (MP) phylogenetic analysis revealed that O2, O5, and O5d were more closely related to those found in Libya (O5b, O5a, O4a, O4b), Lebanon (O1b, O2, O3), and Egypt (O1c). However, O5”b appears to have no close relationship to any of the other mitotypes. The existence of the Middle Eastern ‘O’ lineage in the south central and central United States suggests that further molecular genetic studies of the honey bee population is needed for utilizing and conserving the genetic variation which most likely exists in the Unites States. Furthermore, this study also suggests that feral honey bees are surviving despite the introduction of the varroa mite in the 1980s which reduced the feral and managed populations.

Department of Entomology, University of Arkansas, Insect Genetics Research Laboratory, Fayetteville, AR, USA 72701.

*Corresponding author: email:aszalan@uark.edu
Keywords: *Apis mellifera*, COI-COII intergenic region, Middle Eastern honey bee, mtDNA, O lineage

**INTRODUCTION**

There is historical evidence that at least eight different subspecies of the western honey bee, *Apis mellifera* L., representing three different lineages: Eastern European ‘C’, Near East and Middle East ‘O’, and Western European ‘M’ (Schiff & Sheppard 1993, Sheppard 1989a, b, USDA-ARS 1967), have been introduced to the United States before the arrival of Africanized bees. *A. m. mellifera* L., the “Dark Bee” of Northern and Western Europe, ‘M’ lineage was the first subspecies to be introduced in the 1600’s and was the primary honey bee in the United States until the importation of *A. m. ligustica* Spinola, ‘C’ lineage started in the mid 1800s (Schiff *et al.* 1994, Sheppard 1989a, 1989b). The distribution and origin of the Near East and Middle Eastern ‘O’ lineage extends along the Mediterranean including Libya, Tunisia, Egypt, Syria, Iran, Iraq, Cyprus and Turkey (Shaibi *et al.* 2009, Franck *et al.* 2000, Kandemir *et al.* 2000, 2006, Arias & Sheppard 1996, Ruttner 1988, 1992).

Subspecies from the Middle Eastern ‘O’ lineage were introduced in the 1880s and 1890s and by the end of the century, importation of this lineage stopped due to its less desirable behavioral traits relative to the other lineages that were being imported (Sheppard 1989a, b). In 1922, the US Honey Bee Act was enacted to end the importation of honey bees in order to prevent the introduction of honey bee diseases (Anonymous 2009a). The subspecies most commonly used by the beekeeping industry in the United States today are *A. m. ligustica* Spinola and *A. m. carnica* Pollman (Sheppard 1989b) both of which belong to the ‘C’ lineage (Ruttner 1988, 1992). *A. m. caucasica* Gorbatschev is also used (Sheppard 1989b), and while it was originally placed in the ‘O’ lineage (Ruttner 1988), molecular systematic studies have placed it in the ‘C’ lineage (Smith 1991, Garnery *et al.* 1992).

Before the introduction of Africanized honey bees in the United States, feral honey bees were a mix of the ancestors from Eurasia which were mostly of the ‘C’ lineage and also had a mixture of the ‘M’ and ‘O’ lineage. After Africanization, feral bees were mostly of the African race based on morphology with a mixture of the ‘M’ and ‘C’ lineage but exhibited less of the ‘O’ lineage (Whitfield *et al.* 2006).
Recently, hobbyists, sideliners, and a few commercial beekeepers have become interested in feral colonies for their breeding programs because of their ability to survive despite pests, pathogens, and diseases that plague them (McNeil 2009a, b, c, Webster 2009). These colonies have not been manipulated by beekeepers for a long period of time and may have adapted to the various environmental stresses that threaten their survivability (Villa et al. 2008, Loper et al. 2006, Seeley 2007). Genetic studies can be conducted on feral populations in order to identify those that belong to distinct lineages which reveal mitotype differences that can be a source of genetic variation for beekeepers. Studies have shown that genetically diverse honey bee colonies are more productive, more fit, have increased colony growth and are less susceptible to severe infections (Mattilla & Seeley 2007, Tarpy 2003). With the recent loss of many honey bee colonies due to Colony Collapse Disorder (CCD), the lack of genetic diversity may be one of the contributing factors.

Previous studies conducted on the genetic variation of managed and feral honey bee populations in the United States focused primarily in the southern (Pinto et al. 2007, 2003, Coulson et al. 2005, Schiff et al. 1994), western and southeastern (Delaney et al. 2008, 2009, Schiff & Sheppard 1995, 1996) regions using PCR-RFLP, cytochrome b, microsatellite, and allozyme analysis. Genetic variation of honey bees by DNA sequencing of the mtDNA COI-COII region has mostly been conducted in other countries such as Turkey (Ozdil et al. 2009, Solorzano et al. 2009), Mexico (Kraus et al. 2007), South America (Collet et al. 2006, Ferreira et al. 2009, Prada et al. 2009), Africa (Franck et al. 2001), and Australia (Chapman et al. 2008) which resulted in mitotype variation. Delaney et al. (2008) DNA sequenced 13 samples from 11 U.S. southern states that were collected from 1980 to 1992. This has been the only study conducted in the United States in which samples had been DNA sequenced using the same marker. Pinto et al. (2007) sequenced five samples from three southern states; however, cyt b was used as the molecular marker. In all the previous studies conducted on the genetic variation in the United States, evidence of the Middle Eastern ‘O’ lineage has not been observed.

The objective of this study was to determine the genetic variation of feral and managed honey bee colonies in the central and south central United States using DNA sequencing analysis in order to determine if the ‘O’ lineage of *Apis mellifera* is present in the United States.
MATERIALS AND METHODS

Adult worker specimens were collected from Arkansas, Missouri, Mississippi, Nebraska, New Mexico, Oklahoma, Texas, Utah, Louisiana, South Dakota, Florida, Wisconsin, Ohio, and California, as part of a survey for genetic diversity of A. mellifera in southern and central United States (RM unpublished data) (Table 1, Fig. 1).

Specimens were preserved in 70-100% ethanol until DNA extraction. We only extracted DNA from a single worker from each bee colony since mitochondrial DNA (mtDNA) is identical for all members of the colony. After allowing the specimen to dry on a paper towel, DNA was extracted from individual thoraces using the Puregene DNA isolation kit D-5000A (Gentra, Minneapolis, MN). A 680 to 1200 bp region of the mitochondrial COI-COII intergenic region was amplified in a Techne T-412 thermal cycler (Techne Inc., Burlington, NJ) using E2 (5’-GGC AGA ATA AGT GCA TTG-3’) and H2 (5’-CAA TAT CAT TGA TGA CC-3’) PCR primers (Garnery et al. 1993). PCR was conducted with 2 µl of the extracted DNA following the profile of 35 cycles of 94°C for 45s, 46°C for 45s, and 72°C for 45s. Amplicons were separated by gel electrophoresis in 1% agarose and photo documented using a BioDoc-it™ Imaging System (UVP, Inc., Upland, CA). Amplified DNA was 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. DNA sequences new to this study were deposited in GenBank as accession numbers FJ743632, FJ743633, GQ856212 and GQ856213.

DNA sequences were aligned using CLUSTAL W (Thompson et al. 1994) by using DNA sequences from this study and additional ones from GenBank (Fig. 2). Maximum parsimony (MP) analysis on the alignments was conducted for the best-fitting model using PAUP* 4.0b10 (Swofford 2002). Due to the large size variation in the COI-COII region of *Apis mellifera* (629-1223 bp), the MP analysis was unrooted, gaps were treated as missing characters, and no outgroup taxa were used. The reliability of trees was tested with a bootstrap test (Felsenstein 1985). Parsimony bootstrap analysis included 1,000 replicates using the Branch and Bound algorithm of PAUP* (Swofford 2002). In addition, Bayesian phylogenetic analysis was conducted with the GTR+G model using Bayesian Evolutionary Analysis Sampling Trees (BEAST) v 1.4.1 (Rambaut & Drummond 2007). For determining Bayesian inference, four Markov chains ran for $10^5$ generations with a burn-in of 0 ($2 \times 10^4$) were used to reconstruct the consensus tree. Genetic distances, with gaps excluded, were calculated according to the Tajima-Nei model (Tajima & Nei 1984) using the distance matrix option in PAUP* (Swofford 2002).

**RESULTS**

Of the 469 *A. mellifera* samples subjected to DNA sequencing analysis (R.M. unpublished data), 24 samples were found to have four mitotypes characteristic of the Near East and Middle Eastern ‘O’ lineage based on our phylogenetic analysis (Table 1, Fig. 2). These mitotypes were found in seven States (Table 1, Fig. 1). Of these four mitotypes, we observed three new mitotypes not previously documented on GenBank; O5d, O5”b, and O5, with another mitotype, O2, being identical to GenBank accession FJ477996 from Lebanon (Franck et al. 2001).

The size of the amplicon for the mitotypes were 629 bp (O1), 836 bp (O5), 823 bp (O2) and 1223 bp (O5”b). This size variation was due to insertions/deletions along the P and Q non-coding regions in cytochrome oxidase II. A BLAST search revealed that mitotype O5 was most similar to O5a (GenBank EU785974, 99% match), O5d was most similar to O5 (GenBank FJ743633,
98% match, Fig. 2), and O5”b was most similar to O5’ (GenBank EU785977, 95% match). Tajima-Nei sequence distances (Tajima & Nei 1984) among ‘O’ lineage mitotypes found in the United States ranged between 0.174% to 0.241% (gaps excluded). Mitotype O5 was the most common, n = 13, and accounted for 52% of the total observed ‘O’ mitotypes (Table 1). Mitotype O5d and O5”b were found in Arkansas, mostly in feral colonies. However, O5”b was found in one managed hive. Mitotype O2 was found in managed and feral colonies in California as well as in Utah. However, the managed colonies were originally derived from swarms or from feral colonies. In Mississippi, O5”b and O5 was only found in the feral populations. Mitotype O5 was found in the feral populations of New Mexico as well as in Oklahoma and Texas.

Alignment of the *A. mellifera* ‘O’ lineage mtDNA COI-COII sequences along with representatives of the three other lineages resulted in a total of 1253 characters. Of these characters, 20 (1.6%) were variable and 39 (3.1%) were parsimony informative (gaps excluded). Maximum parsimony analysis resulted
in a single most parsimonious tree (Fig. 2) (Length = 86, CI = 0.733) using the Branch and Bound search algorithm in PAUP*. Bayesian and maximum parsimony analysis revealed four distinct monophyletic clades representative of the African ‘A’ mitotype, Western European ‘M’, and the Eastern European ‘C’ mitotype (Fig. 2). Within the ‘O’ lineage, four distinct clades were observed. Mitotype O5”b formed a distinct clade, relative to all of the other ‘O’ lineage mitotypes. Mitotype O2 formed a common clade with O1b and O4b from Lebanon and Libya, while mitotypes O5 and O5d formed a common clade with five mitotypes from Libya, Egypt and Lebanon.

DISCUSSION

This is the first study to document the existence of the *A. mellifera* ‘O’ lineage in the United States based on mtDNA COI-COII DNA sequence data. The last known existence of it is based on historically documented evidence when it was introduced in the late 1800s (Sheppard 1989a, b). Previous studies conducted on the genetic variation of managed and feral colonies in the U.S.
did not reveal this lineage (Delaney et al. 2008, 2009, Schiff et al. 1994, Schiff & Sheppard 1995, 1996). Even though the studies by Delaney et al. (2008, 2009) used the same marker as ours, they relied primarily on PCR-RFLP data. DNA sequencing of this same region can distinguish a single nucleotide difference resulting in greater genetic variation detection (Ozdil et al. 2009, Shaibi et al. 2009, Franck et al. 2001, Palmer et al. 2000). As evidence for the greater sensitivity of DNA sequence data, we have observed, by DNA sequencing of the same mtDNA intergenic region, ten mitotypes within the African ‘A’ lineage of which five had not been previously described (ALS & RM unpublished data).

This study provides evidence that there may be a geographical clinal distribution of the ‘O’ lineage mitotype which they may have adapted to. For example, mitotype O2 is only found in the western states (UT and CA), mitotype O5 was observed in the southwestern states (OK, NM and TX), and mitotype O5d and O5”b were observed in Arkansas and Mississippi (Figs. 1, 2). There is substantial climate differentiation between these regions ranging from arid and semi-arid with some variation of temperatures among the seasons to a more temperate weather pattern which includes more humidity (Anonymous 2009b). The molecular marker that we used determines maternal ancestry and not adaptive traits. Further research will need to be conducted on individual populations of this lineage in order to observe adaptive trait differences. However, another possible explanation for the phylogeography of these mitotypes is that they may have been initially introduced to different regions of the country.

Further molecular genetic research on honey bee populations in the United States will be important in order to determine if this lineage is of the Old World honey bee which was introduced by the early settlers or if it has been recently introduced from approved countries that have been, since 2004, allowed to export to the U.S. germplasm and adult queen honey bees and packaged bees (worker bees, drones, with a queen) of A. m. mellifera (Anonymous 2009c). Although A. m. mellifera is the noted subspecies, without genetic molecular confirmation, it is not known what lineage is actually being imported.

Our study does provide genetic evidence that feral bees are surviving in the U.S. despite the introduction of Varroa destructor (Anderson & Trueman) (Villa et al. 2008; Loper et al. 2006, Seeley 2007) as well as other pests, pathogens,
and diseases that plague them. The fact that this bee lineage was introduced into the United States nearly 130 years ago and has not been commercially available for nearly 100 years but is present in feral populations, as well as being managed by a few beekeepers, gives support to the idea that these bees may represent a distinct genetic source which can be used for breeding purposes. It further shows that there may be sufficient genetic variation in the U.S. to sustain the beekeeping industry which has been recently affected by CCD. Studies currently being conducted on the genetic variation of several U.S. queen breeder colonies have thus far resulted in only 6 mitotypes, all of which are of the Eastern European ‘C’ lineage (RM & ALS unpublished data). These preliminary results show that there may be a possible bottleneck within the commercial beekeeping industry. Further research is needed by conducting field experiments on honey bees to see if there are any trait differences within other lineages, as well as studies on bees from hobbyists and sideliner beekeepers that collect swarms as well as feral colonies which may contain greater genetic diversity.

ACKNOWLEDGEMENTS

We thank Ed Levi of the Arkansas Plant Board, Richard A. Grantham, Carrie Owens, Marion Ellis, John Wagner, Clarence Collison, Danielle Downey, Grant Gillard, Tom Glenn and Lisa Bradley. We also thank Jackie McKern for assisting with DNA extractions. This research was supported in part by the University of Arkansas Agricultural Experiment Station.

REFERENCES


Schiff, N.M. & W.S. Sheppard. 1996. Genetic differentiation in the queen breeding population of the western United States. Apidologie. 27: 77-86.


