GENETIC CHARACTERIZATION OF U.S. HONEY BEE POPULATIONS

By

Deborah A. Delaney

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Entomology

August 2008

To the faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of Deborah A. Delaney find it satisfactory and recommend that it be accepted.

_____
Chair

_____

_____

_____

ACKNOWLEDGMENT

- I would like to thank my committee; Dr. Walter S. Sheppard, Dr. Thomas Unruh, Dr. Michael Webster and Dr. Carol Anelli, for guiding me through this process. I also want to thank Dr. Marina Miexner and Dr. Jamie Strange for all their advice and support. Dr. Richard Zack has been a great advocate and extremely supportive throughout this research. This work could not have happened without the assistance of the western commercial queen breeders and southeastern commercial queen breeders. Finally I would like to give thanks to the bees for sacrificing their lives for the sake of science

GENETIC CHARACTERIZATION OF U.S. HONEY BEE

POPULATIONS

Abstract

By Deborah A. Delaney, Ph.D.
Washington State University
August 2008

Chair: Walter S. Sheppard

The history of honey bee importation into North America began in the 17th century. Between 1622 and 1859 eight subspecies were brought into North America. Small numbers of queens representing these subspecies were initially brought over and are the genetic ancestors of the populations that remain in the United States today. In 1987 *Varroa destructor,* a parasitic brood mite, was found in the United States. It has been responsible for the near elimination of the once flourishing feral honey bee population, and has annually reduced the commercial honey bee populations. The reduction of honey bee numbers by *Varroa* caused a second genetic bottleneck of the U.S. populations. A third genetic bottleneck can be attributed to current queen breeding practices. Queen breeders typically use a small number of breeder queens to make a large number of replacement queens for managed colonies. My data shows that 473 breeder queens were used to make replacement queens for 1/3 of all managed colonies in the United States.

These three population bottleneck events underlie our concerns for the genetic diversity of the honey bee in the United States. To better understand changes in the

genetic diversity of honey bee populations, feral populations collected between 1980 and 1992 and commercial populations from the southeastern and western United States collected in 1993-1994 and again in 2004-2005 were analyzed using mitochondrial and microsatellite DNA.

The results indicate that feral populations collected between 1980 and 1992 were genetically different from the managed populations, and the feral populations probably served as a source of allelic diversity for both the managed populations. The genetic composition of the two managed populations was different. The western managed population experienced a significant loss of original alleles over the ten year sample period, likely due to genetic drift and or bottlenecks caused by *Varroa destructor* and breeding practices. However, there was also a simultaneous gain of "new" alleles into both the western and southeastern managed populations probably reflecting contributions from additional introductions of Russian or Carniolan strains or from Africanized honey bees.

TABLE OF CONTENTS

Page

Dedication

This dissertation is dedicated to my husband and the hard but amazing journey we both traveled during this research, and to my boys, Levi and Farren who kept things in perspective with lots of love and laughter.

General Introduction

The genus *Apis* contains close to a dozen recognized species, with all but one endemic to eastern Asia. The exceptional "western" honey bee, *Apis mellifera* L., had an original distribution allopatric to the rest of the genus *Apis* that included Africa, Europe and central and western Asia. Within this expansive range, more than 2 dozen subspecies of *A. mellifera* have been recognized, based on multivariate analysis of morphology (Ruttner, 1988; Sheppard et al. 1997;Engel, 1999; Sheppard and Meixner, 2003). Although intraspecific classification of *A. mellifera* is based on morphology, substantial differences in behavior and physiology also occur among subspecies, associated with their adaptation to divergent climatic and ecological conditions.

Humans have been instrumental in increasing the range of the honey bee. Honey bee importation into North America began in the early 17th century and by 1622 the Dark bee of western Europe (subspecies *Apis mellifera mellifera*) had established feral populations (Sheppard, 1989). These populations expanded in advance of European settlers, such that Native Americans considered the local presence of honey bees to foretell the impending arrival of European settlers and referred to the insect as "white man's flies" (Jefferson, 1788). No additional introductions of honey bees are known to have been made until 1859 (Sheppard, 1989). However, between 1859 and 1922, seven additional subspecies from Europe, Africa and western Asia were introduced into the United States, with varying measures of commercial success (Sheppard, 1989; Schiff and Sheppard, 1993).

Of the eight subspecies brought into the country, only three found favor with the beekeeping community and remain available today as selected "strains" from bee breeders. These subspecies (and the commercial designations under which their presumptive descendents are commonly sold) included: *Apis mellifera ligustica,* (Italian honey bees), *Apis mellifera carnica* (Carniolan honey bees) and *Apis mellifera caucasica* (Caucasian honey bees).

The degree of similarity between the strains of honey bees currently sold and the original subspecies that were introduced remains largely unknown. In 1929, Alpatov compared morphometric variation in U.S. "Italian" honey bee populations to *A. m. ligustica* from Italy and reported that beekeeper selection in the U.S. had led to higher levels of yellow coloration in the U.S. strain compared to the original subspecies (Alpatov, 1929).

In the years following the Alpatov study, Italian honey bees became the predominant strain used by U.S. beekeepers. However, analysis of feral honey bee populations in the U.S. showed that measurable genetic influences remain from early introductions of other European and African subspecies. For example, the mean frequency of mtDNA haplotypes typical of *A. m. mellifera* was 37% in 692 feral colonies sampled from the southern U.S. and as high as 67% in feral colonies from Arizona (Schiff et al. 1994). Although *A. m. mellifera* was the first subspecies introduced into North America in the 17th century, it has not been commercially available for many decades. Similarly, about 2% of the feral colonies sampled in these studies had a mtDNA haplotype characteristic

of *A. m. lamarckii*, a subspecies endemic to the Nile River valley in Egypt and Sudan that was introduced into the U.S. in 1869 (Sheppard, 1989; Schiff et al. 1994).

Past analyses of commercial breeding populations revealed little influence from *A. m. mellifera* mtDNA, suggesting that queen producers had maintained control of the maternal lineages of their selected strains (Schiff and Sheppard, 1995; Schiff and Sheppard, 1996). However, one limitation of the mtDNA test employed in these studies was the inability of the marker to discriminate among subspecific maternal origins of *A. m. ligustica*, *A. m. carnica* and *A. m. caucasica.*.

Allozyme analysis of feral populations from three states (North Carolina, Georgia and Alabama) in 1991-92, revealed a higher frequency of the malate dehydrogenase 80 allele (Mdh^{80}) compared to commercial populations (Schiff et al. 1994, Schiff and Sheppard 1995, 1996). *Mdh80* occured in high frequency in the Old World species *A. m. mellifera* (Badino et al. 1984; Sheppard and Berlocher, 1984; Cornuet et al. 1986; Schiff et al. 1994) and the finding of higher frequencies of the allele in feral stocks likely reflected the historical influence of *A. m. mellifera*. Commercial populations in the U.S. were characterized by a low frequency of this allele, as are their presumptive ancestral subspecies, *A. m. carnica* and *A. m. ligustica*.

The genetic composition of honey bee populations from the two main queen breeding regions of the U.S. has been analyzed using mitochondrial and allozyme methods (Schiff and Sheppard, 1995, 1996). In a mtDNA analysis of queen mother colonies from the southeastern U.S., Schiff and Sheppard assessed 142 breeder colonies from 22 apiaries

and found that 4 % exhibited representative M lineage mtDNA haplotypes known from *A. m. mellifera.* Representatives of the C mtDNA haplotypes, characterizing *A. m. carnica (C1), A. m. ligustica (C2),* and *A. m. caucasica* (C3) , were found in 96% of the breeder colonies. These results differed significantly from those found in the feral population where M lineage haplotypes persisted. The low frequency of *A. m. mellifera* mtDNA haplotypes in the commercial managed population supported the hypothesis of restricted maternal contributions from the feral into the managed populations (Schiff and Sheppard, 1995). Analysis of malate dehydrogenase allele frequencies also suggested that managed populations were more homogeneous than the feral population. The similarity of the findings from both mtDNA and the allozyme studies suggest that commercial and feral populations have experienced some barriers to gene flow.

From the western queen producing area, Schiff and Sheppard genetically characterized 178 breeder queen colonies from 22 apiaries (1996). Relative to the feral populations sampled, low heterogeneity was also found in the western breeding population, Two colonies out of 178 breeder colonies had a M lineage mtDNA haplotype that likely originated from *A. m. mellifera*, while the other 176 colonies a mtDNA haplotype associated with the C lineage subspecies *A. m. carnica* and *A. m. ligustica.*. Malate dehydrogenase allele frequencies were significantly different from those reported from both the feral and southeastern commercial population. Overall, these studies showed that the feral populations were distinct from both commercial managed queen breeding populations. The two geographically isolated commercial queen breeding populations

were also genetically different from each other. These authors suggested that the feral population could serve as a reservoir for genetic variability useful for breeding and that genetic differences between queen producing populations in the southeastern and western U.S. could provide a source of genetic diversity useful to the commercial queen producing industry.

The purpose of this research project was to assess the current genetic composition of US honey bee populations and to compare them to the composition of a parallel set of honey bee samples collected in 1993-1994. I investigated whether there have been measurable changes in genetic variability in the U.S. commercial queen breeding population in the past decade and whether genetic differences between the two queen producing regions still remain.

Chapter 1 characterizes extant western and southeastern commercial breeding populations in the U.S. using mitochondrial DNA and microsatellite markers. These populations are compared to a previously collected set of honey bee samples for the same genetic markers. This chapter will be submitted to the Annals of the Entomological Society of America. Drs. Marina Miexner, Nathan M. Schiff and W.S. Sheppard will also be authors on this paper due to their contributions to this work.

Chapter 2 evaluates the influence of feral honey bee populations on commercial populations based on analyses of microsatellite and mtDNA variation. This chapter also examines whether the ensuing loss of feral populations due to *Varroa* differentially affected the two commercial breeding populations. This manuscript is being prepared for

Molecular Ecology and will have multiple authors; Drs. Marina Miexner, Nathan M.

Schiff and W.S. Sheppard.

Bibliography

Alpatov W. (1929) Biometrical studies on variation and races of the honey bee *Apis mellifera* L.. Rev Biol **4**, 1-57.

Badino G., Celebrano G., Manino A. (1984) Population genetics of Italian honeybee (*Apis mellifera ligustica* Spin.) and its relationships with neighbouring subspecies. Boll. Mus. Reg. Sci. Nat. Torino **2**, 571-584.

Cornuet J., Daoudi A., Chevalet C. (1986) Genetic pollution and number of matings in a black honey bee (*Apis mellifera mellifera)* population. Theor. Appl. Genet. **73**, 223-227.

Cornuet J., 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.

DeJong D. (1990) Mites: *Varroa* and other parasites of brood, pp.200-218 in R. A. Morse and Nowogrodzki (eds.) Honey bee pests, predators and diseases. 2nd edition Cornell University Press, Ithaca, NY

Engel M. S., (1999) The taxonomy of recent and fossil honey bees (Hymenoptera: Apidae; *Apis*). J. Hym. Res., **8**, 165-196.

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

Estoup A., Largiader C., Perrot E., Chourrout D. (1996) Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. Mol. Mar. Biol. Biotechnol. **5**, 295-298.

Finley J., Camazine S., Frazier M. (1996) The epidemic of honey bee colony losses during the 1995-1996 season. Amer. Bee Journal 136, 805-808.

Franck P., Garnery L., Solignac M., Cornuet J. (1998) The origin of West European subspecies of honey bees (*Apis mellifera*): New insights from microsatellite and mitochondrial data. Evolution **52**, 1119-1134.

Franck P., Garnery L., Loiseau A., Oldroyd B., Hepburn H., Solignac M., Cornuet J. (2001) Genetic diversity of the honey bee in Africa: microsatellite and mitochondrial data. Heredity **96**, 420-430.

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

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

Harrison J., Hall G. (1993) African-European honeybee hybrids have low nonintermediate metabolic capacities. Nature **363**, 258-260.

Jefferson T. (1788) *Notes on the State of Virginia.* Prichard and Hall, Philadelphia. 244pp.

McGregor S. (1976) Insect pollination of cultivated crop plants, Agricultural Handbook No. 496 Agricultural Research Service, US Dept. of Ag.

Morse R., Calderone N. (2000) The value of honey bees as pollinators of U.S. crops in 2000. Cornell University, Ithaca, NY, March 1-31pp.

Raymond M., Rousset F. (1995) Population genetics software for exact test and ecumenism. J. Hered. **86**, 248-250.

Ruttner F. (1988) Biogeography and Taxonomy of the Honeybees. Springer, Heidelberg 1-284.

Sheppard W., Berlocher S. (1984) Enzyme polymorphisms in *Apis mellifera mellifera* from Norway. J. Apic. Res. **23**, 64-69.

Sheppard W. (1988) Comparative study of enzyme polymorphism in the United States and European honey bee (Hymenoptera: Apidae) populations. Ann. Entomol. Soc. Amer. **81**, 886.

Sheppard W. (1989) A history of the introduction of honey bee races into the United States: Part I and II. Amer. Bee Journal **129**, 617-619, 664-667.

Sheppard W., Arias M., Grech A., Meixner M. (1997) *Apis mellifera ruttneri*, a new honey bee subspecies from Malta. Apidologie **28**, 287-293.

Sheppard W., Meixner M. (2003) *Apis mellifera pomonella*, a new honey bee subspecies from Central Asia. Apidologie **34**, 367-375.

Schiff N., Sheppard W. (1993) Mitochondrial DNA evidence for the 19th century introduction of African honey bees into the United States. Experientia **49**, 530-532.

Schiff N., Sheppard W., Loper G., Shimanuki H. (1994) Genetic diversity of feral honey bee (Hymenoptera: Apidae) populations in the southern United States. Ann. Entomol. Soc. Amer. **87**, 842-848.

Schiff N., Sheppard W. (1995) Genetic analysis of commercial honey bees (Hymenoptera: Apidae) from the Southeastern United States. J. Econ. Entomol. **88**, 1216-1220.

Schiff N., Sheppard W. (1996) Genetic differentiation in the queen breeding population of the western United States. Apidologie **27**, 77-86.

Solignac M., Vautrin D., Loiseau A., Mougel F., Baudry E., Estoup A., Garnery L., Haberl M., Cornuet J. (2003) Five hundred and fifty microsatellite markers for the study of the honeybee (*Apis mellifera* L.) genome. Mol. Ecol. Notes **3**, 307-311.

Weir B. (1988) Statistical Analysis of DNA sequences. J. Nat. Cancer Institute **80**, 395-406.

LIST OF TABLES

LIST OF FIGURES

THE GENETIC CHARACTERIZATION OF COMMERCIAL HONEY BEE (*Apis mellifera* L.) POPULATIONS IN THE UNITED STATES USING MITOCHONDRIAL AND MICROSATELLITE MARKERS

Abstract

By Deborah A. Delaney, Ph.D.
Washington State University
August 2008

Chair: Walter S. Sheppard

Honey bee populations introduced into North America probably experienced three major genetic bottleneck events that led to potential reductions in genetic diversity. The first certainly occurred during the sampling of subspecies for importation, when settlers introduced a small number of subspecies representing only 1/3 of the known subspecies diversity. *Varroa destructor,* a parasitic brood mite, likely caused a second bottleneck, as its outbreak was associated with the virtual disappearance of feral honey bee populations and annual reductions of commercial populations by as much as 50%. A third bottleneck may be ongoing and stems from current queen breeding practices that use a small number of breeding mothers to supply annual replacement queens for 1 million managed colonies. Due to these bottleneck events, concerns arose about sustaining genetic diversity in these populations.

In this study we examined samples from two commercial breeding areas, the western commercial breeding population (WCBP) and the southeastern commercial breeding population (SCBP) sampled in the 1993-1994 and again in 2004-2005. The

goal of this study was to characterize the genetic composition of these populations and to measure potential changes in genetic diversity and composition across the sampling period. Diversity levels within and between the two commercial breeding areas in the U.S. were analyzed using the *Dra*I RFLP of the COI-COII mitochondrial region and 10 polymorphic microsatellite loci.

The mitochondrial DNA haplotypes C1 and C2, characteristic of the most popular bee strains (Italians and Carniolans, respectively) sold in the U.S., were the dominant haplotypes at both sample dates. The frequency of *Apis mellifera mellifera* M haplotypes, M4, M7, and M7', decreased during the ten year span. An A1 haplotype characteristic of Africanized bees was found in the SCBP from 2005. Microsatellite analysis showed there was a loss of alleles in both the WCBP and SCBP but these losses were not found to be significant due to simultaneous gains of new alleles into these populations between 1993 and 2005. Genetic differences that occurred between the WCBP and SCBP were still detectable a decade later, suggesting that these populations could be useful sources of diversity for each other in the future.

Introduction

Honey bees are the most economically valuable pollinators of agricultural crops worldwide. One third of the total human diet is dependent on plants pollinated by insects; predominately honey bees (McGregor, 1976). In North America, honey bees pollinate more than 90 crops, about 14.6 billion dollars annually (Morse and Calderone, 2000). Production of almonds in California is a $2 billion enterprise(National Research Council, 2007) and is almost entirely dependent upon honey bees. However, the honey bee, *Apis mellifera* L., is not native to the Americas. Across an endemic range of Europe, Africa and western and central Asia, the honey bee evolved and adapted to a large variety of climatic and ecological conditions. Currently, 26 different subspecies are recognized within the endemic range, with classification based on morphology (Ruttner 1992). From this large pool of honey bee subspecies a modest subset was exported to North America early in the seventeenth century.

The first recorded importation occurred in 1621, when the Virgina Company sent ships full of seeds, fruit trees and various animals, including bees (Horn, 2005). The majority of honey bee importations occurred between 1859 and 1922. In 1922, in response to the Isle of Wight disease, a law was passed which prohibited the importation of adult honey bees into the U.S. (Sheppard, 1989). Between 1859 and 1891 seven additional subspecies were brought into the U.S. (Table 1). However, only three subspecies found favor with the beekeeping community (*Apis mellifera carnica, Apis mellifera caucasica* and *Apis mellifera ligustica)* and remain available as selected strains today in the U.S. In 1990 Africanized honey bees expanded their range to North America, representing

descendents of the ninth subspecies (*Apis mellifera scutellata)* to colonize the U.S. (Sugden &Williams, 1990).

Honey bee importation to the U.S. was based on only a partial sampling of the original diversity contained within the species (only 9 out of 26 subspecies were brought to the U.S.). Not only was this a modest sampling of the species-wide diversity, but within most sampled subspecies only a few queens of each were initially imported to the New World (Sheppard, 1989).

There are currently two geographically distinct commercial queen producing regions in the U. S. producing roughly equal numbers of queens for sale. The western queen producing region is primarily located in central California with some operations in southern California (Fig.1). The second queen producing area is in the southeastern U.S., with the majority of operations located from Florida westward to Texas (Fig.1). Commercial honey bee breeding populations are maintained by queen producers who typically select for traits desirable to the beekeeping industry, including honey production, colony growth, colony survivorship and temperament. Coloration is also a major criterion used for the selection of different strains (Delaney et al. unpublished data).

As a whole, the honey bee breeding industry uses a small number of queen mothers (less than 600) to produce nearly 1 million replacement queens for beekeepers in the U.S. (Schiff and Sheppard 1995, 1996). The genetic basis for these replacement queens is

determined by the number of queen mothers used by queen producers and the total genetic diversity of males to which they mate. These 1 million daughter queens are sufficient to requeen close to one third of the total estimated population of 3.2 million managed colonies in the U.S. (National Research Council, 2007). Thus, breeding practices create the potential for another genetic bottleneck for U.S. honey bee populations and a further loss of diversity.

Queen producers from the western portion of the U.S. market two main strains of queens, the Italian honey bee originally derived from *Apis mellifera ligustica,* and the Carniolan honey bee, derived from *Apis mellifera carnica.* Due to the open aerial mating behavior of honey bees, the queens produced are some admixture of the different gene pools that are present at these central California sites.

The southeastern queen producing region also produces *A. m. ligustica* and *A. m. carnica*. However, the presence of feral populations may have affected the genetic make up of this southeastern population (Schiff and Sheppard 1995). Genetic remnants of *Apis mellifera mellifera,* the “dark bee” from northern Europe, were still present in the southeastern managed population in 1993 in low percentages, reflecting genetic contributions from the feral population. Schiff et al. (1994) found that the 1991-1992 feral population from this region contained over 30 percent mtDNA representative of A.m.m. Thus, the feral population in 1991-1992 represented a somewhat separate "gene pool”, that reflected the historical introduction and spread of *A. m. mellifera*.

In 1987, a parasitic mite (*Varroa destructor*) became established in the U.S. and rapidly spread (Wenner and Bushing, 1996). Untreated colonies often died within 2 years and feral populations were, for the most part, eliminated (Ritter, 1988; Sanford, 2001; Seeley, 2007). Large annual winter losses of managed colonies also occurred due to the mite, although beekeepers could usually recover numbers by making colony splits in the spring (Korpela et al. 1992). The apparent mite-induced loss of the feral honey bee and mite-associated overwintering losses of managed colonies represents another possible loss of genetic diversity for U.S. honey bee populations, both from loss of A.m.m. genes and possibly reduced effective population sizes of breeding stocks.

This research addresses three main questions: 1) Has there been a change in genetic diversity in the western commercial breeding population (WCBP) and southeastern commercial breeding population (SCBP) during the past decade? 2) Do genetic differences that were detected over ten years ago between the WCBP and SCBP still remain? 3) Did the disappearance of feral populations disproportionately affect the SCBP? The bee breeding populations were chosen for this study because the daughter queens produced and sold by queen producers annually are sufficient to requeen about 1/3 of the managed colonies in the U.S. Therefore the genetic composition of the queen breeding population is representative of the managed honey bee population in the U.S. Results from this work will provide data applicable to discussions of the need for further importation of new honey bee germplasm into the U.S.

Materials and Methods

Collection protocol

Western commercial breeding population

In 1993 and 1994 samples of adult honey bee workers were collected from 178 colonies from 21 commercial apiaries in the western United States (Figure 1) (Schiff and Sheppard 1996). A sample consisted of 200-300 workers from each queen mother colony. Of these 178 colonies 176 were reanalyzed in the current study. All samples were stored at -80ºC until analysis. In May of 2004, the western queen producing area was resampled, with 212 colonies sampled from fifteen queen producing operations. Of these fifteen queen producing operations, eleven had been sampled in the 1993-1994 sampling period. Two to three hundred adult workers were collected from each of the 212 colonies.

Southeastern commercial breeding population

In 1993 and 1994 honey bee workers were collected from 185 colonies from 22 different queen producers in the southeastern United States (Figure 1) (Schiff and Sheppard 1995). All samples were stored in -80º C temperatures. Of these 185 colonies, 155 colonies were reanalyzed in the current study to ensure that microsatellite and mtDNA measurements are comparable with recent samples. In June of 2005, the southeastern queen producing region was sampled. Adult honey bee workers were collected from 132 colonies from 20 queen producing operations. Of these 20 operations, eleven had been sampled in the 1993-1994 sample period. Two samples were also collected from a

commercial beekeeper concerned with the temperament of particular stock. A sample consisted of 200-300 adult workers collected directly from the hive.

Sampled adult workers were daughters of the queen mothers used by queen producers during the sampling season, and represented the genetic stock currently being sold to beekeepers. After collection the samples were stored in 95 % ethanol. Survey data on the numbers of queens produced, the selection criteria and source of stock were gathered from each queen producer.

Molecular analysis

DNA extraction

Total DNA was extracted from half the thorax or the middle leg of a single adult honey bee worker from each queen mother colony and placed in 150 µl of 10% chelex and 5 µl proteinase K solution (Walsh et al. 1991). These samples were placed in a thermocycler for 1 h at 55º C, 15 min at 99º C, 1 min at 37º C and 15 min at 99º C. The extracted DNA was stored at -80º C until used for polymerase chain reaction (PCR) amplification.

Mitochondrial DNA analysis

Mitochondrial DNA analysis was performed on the intergenic region between the COI and COII genes. This region was amplified from one worker bee from each colony using PCR and primers E2 (5'-GGCAGAATAAGTGCATTG-3') and H2 (5'-CAATATCATTGATGACC-3') (Cornuet et al.1991; Garnery et al. 1992). Each sample reaction was performed in 25 µl containing 2.7µl *Taq* buffer (Promega), 1.5 mM $MgCl_2$,

25 nmoles of each dNTP, 25 pmoles of E2 and H2 primers, 0.001mg bovine serum albumin (BSA), 11.8µl of distilled water, 0.6 unit of Promega *Taq* polymerase and 1.25 µl of DNA extraction. The conditions for PCR amplification were 30 cycles at 92º C for 30 s, 1.5 min at 47º C and 2 min at 63º C.

To determine the size of the amplified region, a 6µl aliquot of PCR product was electrophoresed on a 1.4% agarose gel. Amplified fragments were visualized for size scoring using ethidium bromide. The remaining 19µl from each sample were digested for 6 h at 37º C with 4 units of restriction enzyme *Dra* I. The restricted DNA fragments were separated on a 10% acrylamide gel and stained with ethidium bromide. Fragment sizes were estimated and mtDNA haplotype names within honey bee "mitochondrial lineages" were assigned to each sample based upon Garnery et al. (1993). The *Dra* I test differentiates up to 50 distinct mtDNA haplotypes within the three mitochondrial lineages of *A. mellifera* (Frank et al. 2000). AMOVA analysis was performed among and within populations and pairwise Fst values for each population and the overall Fst value were calculated using the software program ARLEQUIN 3.1 (Excoffier, L.G. Laval, and S.Schneider, 2005). Populations were defined as four regions: SCBP93, SCBP05, WCBP94 and WCBP04. Chi-squared tests were performed to test for changes in haplotype frequencies between sampling dates in WCBP of 1994 and 2004 and the SCBP of 1993 and 2005.

Microsatellite analysis

Ten variable microsatellite loci were characterized for all samples: A7, A24, A28, A88, A113, B124 (Estoup et al.1995), Ap43 (Garnery et al. 1998), Ap55, Ap66 and Ap81

(Solignac et al. 2003). Amplification of the ten loci was split into two multiplex reactions. Amplification of extracted DNA was performed using 10 µl PCR reactions containing 1x Promega reaction buffer and 1.5 units *Taq* polymerase (Promega, Madison WI), 3mM dNTP mixture, 1.0-4.0 mM of florescent dye-labeled primer, 0.001 mg bovine serum albumin, 1 µl DNA and $MgCl_2$ The loci in plex one, A24, A28, A88, Ap66 and B124 had a final concentration of 1.5mM $MgCl_2$. The loci in plex two, A7, A113, Ap43, Ap55 and Ap81 had a final concentration of 1.2 mM $MgCL_2$. All reactions were amplified at 95º C for one 7 min cycle, 30 cycles of 95º C for 30 sec, 54º C for 30 sec, 72º C for 30 sec and 72º C for a 60 min cycle. The amplifications were processed using an Applied Biosystems 3730 automatic sequencer. Microsatellite fragment sizes were scored using GeneMapper™ software (Applied Biosystems).

Allele frequencies and single locus (and multiple locus genotype) frequencies for each of the four populations (WCBP 1994, WCBP 2004, SCBP 1993 and SCBP 2005) were summarized in GENETIX™. Allelic richness (a_g), which is calculated by adding the number of alleles seen in a population to the expected number, given the number of genes examined in the population and the allele frequencies observed over a set of populations, was calculated using HP-RARE 1.0 (Kalinowski, 2005). The number of private alleles, alleles unique to a particular population, in each of the populations was also calculated in HP-RARE 1.0 (Kalinowski, 2005). ARLEQUIN 3.1 (Excoffier and Schneider, 2005) was used to calculate pairwise FST values by locus and over all loci. The number of alleles at each locus and the loss and gain of alleles at each locus across the two sampling periods were calculated from GENETIX™ allele frequency data. Paired-T tests

(SPSS™) on the number of alleles at each locus (n = 10) for each population were used to test for significant losses or gains in diversity between 1993-1994 and 2004-2005. Exact tests for Hardy-Weinberg equilibrium for each locus, genotypic linkage disequilibrium among loci, and genetic structure over all loci using F statistics, were computed in GENEPOP (web version 3.4; Raymond and Rousset 1995). Population structure was also analyzed using a Bayesian model-based clustering method provided in STRUCTURE (Pritchard et al. 2000). This program infers population structure of populations using allele frequencies of unlinked markers (microsatellites). The parameter set was programmed for independent alleles and individuals to have a mixed ancestry. The program was asked to place individuals from all four commercial populations into 1,2,3 and 4 groups and the analyses consisted of 10^5 burn-in replicates and a run length 10^5 replicates. The best estimate of K or number of populations was determined by looking at the values of log Pr(X/K) and the value of α.

The software program BOTTLENECK (Piry et al. 1999) was used to test whether bottlenecks could be detected in the commercial breeding populations from 1993-1994 to 2004-2005 in the west and southeast. This program determines reductions in effective population size by calculating excesses in heterozygosity using three kinds of statistical tests: sign test, standardized differences test and the Wilcoxon's signed rank test. This program has been evaluated against other genetic diversity indices and is moderately sensitive in detecting genetic changes from source populations but is not as sensitive as some other genetic indices for assessing the magnitude or severity of the bottleneck event (Spencer et al. 2000).

Results

Collection survey

In 1994, the 21 western queen breeders used 295 breeder queens to produce 406,800 queens. Thus the mean production from each breeder queen was 1379 daughters/queen (Schiff and Sheppard 1996). In 2004, the western queen breeders used fewer queen mothers (218) to produce 430,000 daughter queens. Thus, each breeder queen produced an average of 1972 daughter queens.

The southeastern queen production numbers were similar. In 1993, the 22 southeastern queen breeders used 308 breeder queens to produce 483,900 queens for sale. On average each breeder queen produced 1571 daughter queens (Schiff and Sheppard 1996). In 2005, the southeastern queen breeders used 255 breeder queens to produce 439,500 marketable queens, making the mean production 1724 daughters/queens.

Mitochondrial DNA

Three mtDNA haplotypes were detected in the WCBP sampled in 2004 (Table 2 a), and four mtDNA haplotypes were seen from samples from a decade earlier. The most common mtDNA haplotype for both years was C2 (52% in 1994 and 57% in 2004). The haplotype frequencies were significantly different between 1994 and 2004 ($\chi^2 = 21.7$, df = 4 and $P = 0.0002$). The frequency of the C1 mtDNA haplotype increased during the ten year span from 38% in 1994 to 43% in 2004. The C3 mtDNA haplotype occurred in 6%

of the sampled colonies in the 1994 WCBP, but was not detected in 2004. The frequency of the M3 mtDNA haplotype was 3% in 1994 and 1% in 2004.

Six different mtDNA haplotypes (C1, C2, M3, M7, M7' and A1) occurred in samples collected from the SCBP in 2005 (Table 2 b). The haplotype frequencies were significantly different between 1993 and 2005 in the SCBP ($\chi^2 = 17.4$, df = 8 and $P = 0.03$). Frequencies of mtDNA haplotypes from the two sampling periods (1993 and 2005) again show the dominant haplotype increasing: C1 increased from 45% to 56% from 1993 to 2005. The M3 and A1 mtDNA haplotypes were absent in 1993, but were seen at 1% and 2%, respectively in the 2005 SCBP. The frequency of mtDNA haplotype C2 decreased from 47% in 1993 to 40% in 2005. In 1993, M7 and M7' mtDNA haplotypes made up 2 and 3% of the SCBP, respectively, while they both occurred in about 1% frequency in the 2005 population. Two samples analyzed from a commercial beekeeping operation in the south had the A1 mtDNA haplotype.

Based on mtDNA, Fst values were not significantly different from zero within the two regions WCBP (1993 v 2004, Fst = 0.003) and SCBP (1994 v 2005, Fst = 0.012). A higher level of genetic divergence was found between the SCBP of 2005 and the WCBP from 2004 (Pairwise Fst values = 0.044). The Fst value is significantly different from zero ($P = 0.001$). The overall Fst value was 0.04 ($P = 0.001$). The Analysis of molecular variance showed that most of the variation was from within populations (Sum of squares = 334) rather than among populations (Sum of squares = 11). The percentage of variation within populations was 96.2 whereas, 3.8 among populations.

Microsatellites

WCBP

In 1994, the number of alleles per locus ranged between 4.9 (Ap66) and 25.4 (A7) in the WCBP, and in 2004 the numbers of alleles per locus in the WCBP ranged from 5.5 (A88) to 18.5 (A7) (Table 3.). The mean number of alleles per locus for the WCBP from 1994 was 12.8, and 11.5 in 2004. The mean expected heterozygosity values for the WCBP from1994 and 2004 were 0.667 and 0.633, respectively. The mean observed heterozygosity values for the WCBP from 1994 and 2004 were 0.575 and 0.531. The number of private alleles per locus for 1994 and 2004 in the WCBP is given in Table 3. A single locus, Ap66, showed an increase in the number of private alleles in the west in 1994 from 0.00 to 6.01 in 2005. However, reductions in the number of private alleles were found for nine out of ten loci. Loci A7 and A81 experienced large losses of private alleles over the ten year span in the WCBP (8.70 to 1.77 for locus A7 and 7.82 to 1.79 for locus A81).

In the 1994 WCBP a total of 128 alleles were found in the 10 screened loci. Of the 128 alleles detected in 1994, only 92 were found in the samples taken in 2004 for the same population, a loss of 36 alleles. Therefore, 25% of the alleles found in the WCBP in 1994 were not present in 2004. This represents a significant loss, based on a paired t-test, ($p = 0.004$) of the original alleles. All of the 36 alleles lost during the sample period were found in very low frequency (below 0.03) in 1994. Twenty-two new alleles were detected in the 2004 WCBP. In combination with the 92 previously known alleles, a total of 114 alleles were thus detected in the 2004 WCBP. The decline from 128 total alleles

in 1994 to 114 total alleles in 2004 was not significant based on a paired t-test (df = 9 and $P = 0.3$). Eight of these 22 new alleles are known to occur in established U.S. populations (Feral 1980-1990, SCBP 1993); 14 alleles were new to U.S. honey bee populations.

The allele frequencies in WCBP from 1994 compared to 2004 were different and highly significant in 7 (A28, Ap66, B124, A7, Ap113, Ap43 and A81) out of 10 loci, based on probability test (or Fisher exact test) as described by Raymond and Rousset (1995). All pairs of populations across 6 loci (A28, Ap66, B124, Ap113, Ap43 and A81). had significantly different genotypic distributions between the two sampling dates, WCBP 1994 and 2004.

SCBP

In 1993, the number of alleles per locus for the SCBP ranged from 3.4 (locus A28) to 15.2 (locus A7) (Table 3). In 2005, the number of alleles per locus for the SCBP varied from 4.7 (locus A28) to 13.6 (locus A7) (Table 3.). The mean number of alleles per locus for 1993 and 2005 samples was 10.2, and 9.9, respectively. The mean expected and observed heterozygosity for all 1993 samples and loci were 0.6424 and 0.5702, respectively, and for the 2005 samples and loci 0.6223 and 0.5293, respectively.

In 1993 the SCBP had a total of 102 alleles. Of the 102 alleles detected in 1993, only 77 were found in the samples taken in 2005 for the same population, a loss of 25 alleles.

The 25 alleles that were lost during the sample period were found in very low frequencies (below 0.009) in 1993. Twenty-two new alleles were detected in the 2005 SCBP. In combination with the 77 previously known alleles, a total of 99 alleles were detected in 2005 SCBP. Of those twenty-two alleles, 19 were present in other established U.S. honey bee populations (Feral 1980-1990, WCBP 1994). Three alleles were unique to the 2005 SCBP and undetected in previous populations. The SCBP experienced a significant gain in alleles from 1993 to 2005 based on a paired t-test (df = 9 and $P = 0.02$).

Allelic richness (corrected for population sizes) and the amount of private alleles per locus for the SCBP are shown in Table 3. Seven out of ten loci gained more private alleles during the ten year sample period for example, Ap66 increasing from 1 in 1993 to 3.5 in 2005. Three loci show a reduction in the number of private alleles, notably A7 and Ap43, which dropped from 4.5 to 2.9 and 5.7 to 3.1., respectively. Once again these results are supported by allele frequency data (Table 5.), and can be explained by the acquisition of new alleles into the 2005 commercial breeding population.

The population differentiation test (Raymond and Rousseau, 1995) indicated that genic differentiation occurred at 6 loci (A24, B124, Ap113, Ap43, Ap55 and A81) between the SCBP of 1993 and 2005. Genotype distributions were also found to be significantly different between the SCBP of 1993 and 2005 for the same 6 loci.

WCBP v SCBP

The genic and genotypic differentiation tests (below) across all loci were significant when comparing WCBP and SCBP in both 1993-1994 and 2004-2005. This suggests that the two commercial breeding populations are still genetically different from one another.

The analysis for linkage disequilibrium found no significant deviation from equilibrium among the 45 locus pairs at the 5 % level for any of the commercial populations. In both commercial populations (west and southeast) at both sampling times (1993-1994 and 2004-2005), significant departures from Hardy-Weinberg equilibrium were detected using the "exact HW test". The *U*- test for heterozygote deficiency showed these deviations were caused by a deficiency of heterozygotes.

The output of the Bayesian-Markov assignment algorithm implemented in STRUCTURE supported the differences found in the allelic composition of the WCBP and SCBP. The probability of K = 2 was most strongly supported (LnP(D) = -21552, Var [LnP(D)] = 446.4 and a1 = 0.05)(Figure 2 a. and 2 b.), where K=2. Population pairwise Fst values (Table 4) ranged from 0.002 to 0.017 for the 6 unique pairwise comparisons. The within region F_{ST} values of 0.002 and 0.005 during the ten year span indicate little genetic divergence occurred within SCBP and WCBP, during that time span. The Fst values between the WCBP and SCBP had a 10 % difference (0.012 in 1993-1994 and 0.017 in 2004-2005) indicating genetic differences between the two queen producing areas still remain (Table 4).

The statistical tests used to measure excesses in heterozygosity as implemented in the program BOTTLENECK did not detect the occurrence of any bottlenecks for the WCBP or SCBP under the IAM (Infinite alleles model). Changes in allelic diversity between the two sampling periods were not significant (based on a paired T test). However when the newly gained alleles that were not seen in U.S. populations before 2004 and 2005 were removed from the analysis there was a significant loss of alleles (P= 0.01, paired T-test) from 1994 to 2005 for both populations combined.

Discussion

Mitochondrial DNA

In the early 1990s the production queens advertised for sale in the U.S. were mainly Italian, Carniolan or Caucasian, with the exception of several commercial strains of variable genetic origin, such as Buckfast, Starline and Yugo (all typified by C lineage mtDNA). This was supported by the mtDNA results for the 1994 WCBP and the SCBP. Based on the 2004-2005 interviews, 16 out of 36 queen producers now obtain their breeder stock from universities or government programs that have developed specific genetic lines. However, the majority of the queen mothers used by the industry are still from producer-maintained Italian and Carniolan stock. The reliance the beekeeping industry has put on the production of these two strains is mirrored by changes in the mitochondrial haplotypes frequencies, at least for the WCBP breeding population where both C1 and C2 haplotypes increased. However, mitochondrial variation in the two commercial breeding populations, when pooled together, did not significantly change over the course of ten years. High frequencies of the C haplotypes characterized the

commercial populations produced in both regions. Differences within the two regions were characterized by losses of various M haplotypes that existed in low frequencies in WCBP and SCBP in 1993 and 1994 and the addition of an A haplotype found in SCBP in 2005.

In the WCBP, increases in frequency of the C1 (*Apis mellifera ligustica)* and C2 (*Apis mellifera carnica)* mtDNA haplotypes (Kraus et al. 2007) were evident over the ten year period between sampling. The loss of the C3 mtDNA haplotype was probably due to drift and the selective pressure to maintain these other two strains of honey bees. The presence of the M3 mtDNA haplotype typically found in western Europe and representative of *A. m. mellifera* shows that genetic remnants from early introductions still remained in the WCBP in 1994 and the SCBP of 2005. However, the M3 haplotype frequency has continued to decline as less and less queen breeders are selecting queen mothers from local stocks and are using stocks supplied by universities which are limited to bees with C haplotypes representing *A. m. ligustica* and *A. m. carnica*. Also both the western and southeastern queen breeders used fewer queen mothers in 2004-2005 than they did in 1993-1994.

The mtDNA results in the SCBP are similar to the WCBP. More M haplotypes were detected in the SCBP than in the WCBP, probably due to contributions from the once prevalent feral population that existed in that region. Several of these M haplotypes, M3, M7 and M7', are known from western Europe, specifically the Iberian Peninsula (Franck et al. 1998, 2001). However, as occurred in the WCBP, there was also a reduction of M

haplotype patterns from 6% in 1993 to 1% in 2005 in the SCBP (Table 2.), presumably caused by reduced use of local stock s and greater reliance on university stocks or simply sampling drift due to using fewer queen mothers, leading to reductions by chance of these uncommon haplotypes (Genetic drift can also cause haplotype frequencies to increase).

The A1 haplotype (haplotype associated with Africanized bees) pattern found in the SCBP signifies the presence of the African mitochondrial lineage in U.S. honey bee populations. Although the geographic origin of the A1 haplotype is still uncertain, it was reported that the frequency of A1 increases from south to north in Africa, forming a clinal gradient (Moritz et al. 1994, Garnery et al. 1995 and Franck et al. 2001). This latitudinal gradient was supported by Collet et al. (2006), who reported that the frequency of A1 increased towards northern and northeastern Brazil. Sheppard et al. (1999) reported that colonies analyzed from Argentina characterized by the A1 haplotype may have originated from the subspecies *Apis mellifera intermissa,* which is found in northern Africa and is known to have been introduced into the U. S. independently of the African bee, *Apis mellifera scutellata*, which escaped in Brazil.

Also an additional sample collected from a non-breeding commercial bee operation in the southeast had an A1 haplotype. The presence of African mtDNA and nuclear DNA markers in the southern U.S. suggests that there has been introgression of Africanized alleles into U.S. honey bee populations (Whitfield et al. 2006 and Delaney et al.unpublished data). When queen breeders use queen mothers with African mtDNA the dissemination of African haplotypes across the U.S. is inevitable. The queen breeder

operation that contained the A1 haplotype used 8 breeder queens that year to make 6,000 daughter queens. Assuming that they all made equal numbers of daughters, the A1 breeder queen would have produced 750 A1 daughters for sale to beekeepers throughout the U.S.

Microsatellite DNA

There has been no significant loss in overall genetic diversity in the WCBP or the SCBP over the past ten years; however, there has been a change in the allelic composition of these two populations. The WCBP and the SCBP both lost alleles over the ten year span, which could be due to drift, sampling error or queen breeding practices. However, the gain of new alleles into both of these populations due to sampling error or new sources such as Africanized populations, Russian stocks, or other unknown importations, helped to ameliorate the loss of alleles within these populations. Due to the fact that there was only a 50 % sampling overlap of queen breeders between the two sample times (1993-1994 and 2004-2005), allelic losses and gains were tested for significance in the 11 overlapping operations using paired t-tests. When only the data from the 11 queen breeding operations that were sampled in both time periods were analyzed there was no significant loss or gain in alleles in the WCBP between 1994 and 2004 or between both the WCBP and SCBP between 1993-1994 and 2004-2005. However, there was a significant gain ($p = 0.009$) in alleles in SCBP between 1993 and 2005.

Microsatellite data revealed no significant differences in the variance of allele frequencies between the two populations (WCBP and SCBP) based on pairwise Fst

values (0.041 for 1993-1994 comparison and 0.044 for 2004-2005), but pairwise Fst values were 10 % higher in comparison to within population divergence (0.005 and 0.012) between 1993-1994 and 2004-2005. The software program STRUCTURE could best separate sampled individuals into two populations. Significant allelic and genotypic differentiation between the WCBP and SCBP revealed that different amounts and combinations of alleles made up these two populations. Therefore, there are changes in the genetic composition between the WCBP and SCBP after ten years. These differences may reflect the unique historical conditions under which the west and southeast incorporated their mix of source populations and honey bee subspecies, the lack of extensive E-W interchange between queen producers and the geographical barriers of the desert southwest that may play a role in retaining genetic distinctiveness between these two queen breeding regions.

The WCBP experienced a greater loss of original alleles than the SCBP. Reductions in the amount of private alleles occurred in 9 out of ten loci in the WCBP samples. Only three loci showed a reduction in the number of private alleles in the SCBP samples. The SCBP resilience to allelic loss might be attributed to exposure to the feral gene pool. The SCBP shared a higher proportion of alleles, (89%) with the once established feral populations compared to WCBP (77%) (Delaney et al.,unpublished data). Therefore the historical presence and proximity of feral honey bee populations to the SCBP may have helped reduce the loss of alleles in this population. The arrival and presence of Africanized bees and other new honey bee stock in southeast could also be responsible for the gain of alleles into the SCBP.

The importance of maintaining sufficient levels of genetic variation in managed honey bee populations is crucial to the long term health and stability of those populations. Three proposed bottleneck events may have affected genetic diversity within U.S. honey bee populations. The goal of this research was to measure changes in genetic diversity in two commercial queen breeding populations sampled ten years apart using mitochondrial and microsatellite DNA markers. The overall genetic diversity of the WCBP and SCBP did not change over the ten year span, however, the allelic makeup of these populations did change. The WCBP and SCBP were genetically different from one another, and the SCBP actually gained alleles during the ten year span.

Populations exposed to bottlenecks or reductions in effective breeding size are vulnerable to many genetic consequences. Garza et al. (2001) identified some of these consequences as increased identity by descent and a reduction in genetic variation. These changes can result in negative outcomes for populations such as extinction, reduced reproductive function, exposure to rare deleterious alleles and the inability to evolve adaptations to new diseases and parasites (Keller et al. 2001, Freeman and Herron, 2004).
The affects of two genetic bottlenecks i.e., early importation of only a few subspecies and individuals, and the reduction of populations due to *Varroa destructor*) were not detected in this study. Due to the lack of samples from the 19th century when honey bee subspecies were actively being imported into North America, there is no way to measure the extent of genetic loss from this early bottleneck. However, molecular analysis of honey bee samples from the Old World, representative of the imported subspecies, could

be useful in assessing losses in diversity. The absence of honey bee samples from populations before the establishement of *Varroa destructor* also creates difficulties in determining the effects this parasitic mite has had on the genetic diversity of commercial U.S. honey bee populations. The third proposed bottleneck imposed by current queen breeding practices was evident numerically, but a significant loss of alleles was not detected using microsaellite markers. The molecular markers used may not be suitable for detecting allelic loss at the population level and within a ten year span. The significant loss of existing alleles in the WCBP between 1994 and 2004 should not be overlooked, but should be the impetus driving the future monitoring of these populations with markers better designed for assessing differences at the population level.

Bibliography

Belkhir K.P., Borsa P., Chikhi L., Raufaste N. and F. Bonhomme. 2002. Genetix 4.04, logiciel sous windows TM pour la genetic des populations. Laboratoire Génome, Populations, Interactions, Université de Montpellier, Montpellier, France

Collet T., Ferreira K.M., Arias M.C., Soares A.E.E., and M.A. Del Lama. 2006. Genetic structure of Africanized honeybee populations (*Apis mellifera* L.) from Brazil and Uruguay viewed through mitochondrial DNA COI-COII patterns. Heredity **97**: 329-335.

Cornuet J-M., Garnery L., and M. Solignac. 1991. Putative Origin and Function of the Intergenic Region Between COI and COII of *Apis mellifera* L. Mitochondrial DNA. Genetics **1128**: 393-403.

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

Excoffier L.G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online **1**:47-50.

Franck P., Garnery L., Celebrano G., Solignac M., and J.-M. Cornuet. 2000. Hybrid origins of honeybees from Italy (*Apis mellifera ligustica)* and Sicily (*A. m. sicula*). Molecular Ecology **9**:907-921.

Franck P., Garnery L., Loiseau A., Oldroyd B.P., Hepburn H.R., and M. Solignac. 2001. Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. Heredity **86**: 420-430.

Freeman S., and J. Herron. 2004. Evolutionary Analysis 3rd Ed. Pearson Education Inc.: Upper Saddle River, New Jersey. pp. 801

Garnery L., Cornuet J.-M., and M. Solignac. 1992. Evolutionary history of the honeybee *Apis mellifera* inferred from mitochondrial DNA analysis. Molecular Ecology **1**:145-154.

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

Garnery L., Mosshine E. H., Oldroyd B.P., and J-M Cornuet. 1995. Mitochondrial DNA variation in Moroccan and Spanish honey bee populations. Molecular Ecology **4**: 465-471.

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

Garza J.C. and E.G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. Molecular Ecology **10**: 305-318.

GeneMapper™ Software ver. 3.5. 2003. Applied Biosystems. Foster City, California

Horn T. 2005. Bees in America: How the Honey Bee Shaped a Nation.The University Press of Kentucky: Lexington. pp. 352.

Kalinowski S.T. 2005. HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Molecular Ecology Notes **5**: 187-189.

Keller L.F., Jeffery K.J., Arcese P., Beaumont M.A., Hochachka W.M., Smith J.N.M., and M.W. Bruford. 2001. Immigration and the Ephemerality of a Natural Population Bottleneck: Evidence from Molecular Markers. Proceedings of the Royal Society of London B **268**: 1387-1394.

McGregor S. 1976. Insect pollination of cultivated crop plants, Agricultural Handbook No. 496 U.S. Government Printing Office: Washington, D.C. pp. 411.

Moritz R.F.A., Cornuet J-M, Kryger P., and L. Garner. 1994. Mitochondrial DNA variability in South African honeybees (*Apis mellifera* L.) Apidologie **25**:169-178.

Morse R., and N. Calderone. 2000. The value of honey bees as pollinators of U.S. crops in 2000. Cornell University, Ithaca, NY, March 31pp.

National Research Council (2007) Status of pollinators in North America. The National Academies Press, 500 Fifth Street, NW Washington DC 20001. pp 307.

Piry S., Luikart G. and J-M Cornuet. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity **90**: 502-503.

Pritchard J.K., Stephens M., and P. Donnelly. 2000. Genetics **155**: 945.

Raymond M., and F. Rousset. 1995. Population genetics software for exact test and ecumenism. J. Hered. **86**: 248-250.

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

Ritter W. 1988. Varroa jacobsoni in Europe, the tropics, and subtropics. In: Africanized Honey bees and bee mtes. (Needham, G. R., Page, R. E., Delfinado-Baker, M., Bowman, C. E., eds.) John Wiley, New York, 349-359.

Ruttner F. 1988. Biogeography and taxonomy of honey bees. Springer-Verlag: Berlin

Schiff N., Sheppard W., Loper G., and H. Shimanuki. 1994. Genetic diversity of feral honey bee (Hymenoptera: Apidae) populations in the southern United States. Ann. Entomol. Soc. Amer. **87**: 842-848.

Schiff N., and W. Sheppard. 1995. Genetic analysis of commercial honey bees (Hymenoptera: Apidae) from the Southeastern United States. J. Econ. Entomol. **88**: 1216-1220.

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

Sheppard W. 1989. A history of the introduction of honey bee races into the United States: Part I and II. Amer. Bee Journal **129**: 617-619, 664-667.

Sheppard W., Rinderer T., Garnery L., and H. Shimanuki. 1999. Analysis of Africanized honey bee mitochondrial DNA reveals further diversity of origin. Gen Mol Biol **22**:73-75.

Solignac M., Vautrin D., Loiseau A., Mougel F., Baudry E., Estoup A., Garnery L., Haberl M., and J. Cornuet. 2003. Five hundred and fifty microsatellite markers for the study of the honeybee (*Apis mellifera* L.) genome. Mol. Ecol. Notes **3**: 307-311.

Spencer C.C., Neigel J.E. and P.L. Leberg. 2000. Experimental evaluation of the usefulness of microsatellite DNA for detecting demographic bottlenecks. Molecular Ecology **9**: 1517-1528.

SPSS v12.1. 2003. SPSS Inc. Chicago, IL

Sugden, E.A., and K.R. Williams. 1990. October 15: the day the bee arrived. Glean. Bee Cult. **119**:18-21.

Walsh P.S., Metzger D.A., and R. Higuchi. 1991. Chelex ® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques **10**:507.

Whitfield C.W., Behura S.K., Berlocher S.H., Clark A.G., Johnston J.S., Sheppard W.S., Smith D.R., Suarez A.V., Weaver D. and N.D. Tsutsui. 2006. Thrice out of Africa: ancient and recent expansions of the honey bee, *Apis mellifera.* Science **314**: 642-646.

Table 1. The approximate arrival year of importation, and geographic origin of each subspecies of *Apis mellifera* introduced into the Americas. (Sheppard, 1989)

Subspecies	**Origin**	**Importation Date**
A. m. mellifera	Europe north and west of Alps, Central Russia	1622
A. m. ligustica	Italy	1859
A. m. lamarkii	Egypt, Nile river valley	1866
A. m. carnica	Southern Austrian Alps and northern Balkans	1877
A. m. caucasica	Central Caucus Mountains	1880-1882
A. m. cypria	Cyprus	1880
A. m. syriaca	Syria, Eastern Mediterranean	1880
A. m. intermissa	Africa, north of Sahara	1891
A. m. scutellata	Central and eastern equatorial Africa to South Africa	1956

Table 2. Mitochondrial mtDNA haplotype frequencies for the (a) WCBP(b) SCBP in 1993-1994 and 2004-2005.n is the number of colonies sampled.

(a) WCBP

MtDNA haplotype	WCBP 1994 (n=176)	WCBP 2004 (n=212)
C1	**0.386**	**0.425**
C2	**0.517**	**0.574**
C3	**0.063**	**0**
M3	**0.028**	**0.011**

(b) SCBP

MtDNA haplotype	SCBP 1993 (n=155)	SCBP 2005 (n=132)
C1	0.445	0.568
C2	0.471	0.393
C3	0.026	0.00
M3	0.00	0.009
M4'	0.013	0.00
M7	0.019	0.009
M7'	0.026	0.009
A	0.00	0.015

Table 3. Number of alleles per locus (Allelic richness) and number of private alleles per locus for the WCBP and SCBP. Estimates based on rarefaction.

	WCBP 1994	WCBP 2004	SCBP 1993	SCBP 2005
Locus A24				
Allelic richness	7.53	7.40	5.36	6.41
Private alleles	1.21	1.09	0.86	1.91
Locus A28				
Allelic richness	6.67	5.48	3.42	4.72
Private alleles	3.29	2.10	0.34	1.65
Locus A88				
Allelic richness	7.30	5.55	5.94	8.04
Private alleles	1.91	0.16	0.41	2.52
Locus AP66				
Allelic richness	4.99	11.00	5.67	8.14
Private alleles	0.00	6.01	1.00	3.47
Locus B124				
Allelic richness	12.73	12.68	11.14	11.45
Private alleles	1.51	1.47	1.17	1.48
Locus A7				
Allelic richness	25.44	18.50	15.19	13.57
Private alleles	8.70	1.77	4.45	2.87
Locus AP113				
Allelic richness	11.32	10.30	10.70	10.38
Private alleles	2.74	1.71	1.49	1.18
Locus AP43				
Allelic richness	17.62	13.15	11.64	9.00
Private alleles	8.87	4.39	5.70	3.06
Locus AP55				
Allelic richness	13.94	12.18	10.08	11.09
Private alleles	1.79	0.04	0.95	1.97
Locus A81				
Allelic richness	15.06	9.04	5.42	5.98
Private alleles	7.82	1.79	2.11	2.66

Table 4. Pairwise F_{st} values for the commercial breeding populations of the U.S. based on microsatellite data

	WCBP 1994	WCBP 2004	SCBP 1993	SCBP 2005
WCBP 1994	//			
WCBP 2004	0.00535*	//		
SCBP 1993	0.012*	0.01795*	//	
SCBP 2005	0.01468*	0.01697*	0.0020	//

* F_{st} value is greater than 0 with a *P* value less than 0.05

Table 5. Allele frequency data or 10 microsatellite loci used for analyzing the WCBP and SCBP from 1993-1994 and 2004-2005.

A24	WCBP 94	WCBP 04	SCBP 93	SCBP 05
92	0.016	0.009	0.014	0.021
94	0.244	0.314	0.371	0.373
96		0.004	0.005	
98	0.002	0.004	0.002	0.003
100	0.011	0.002	0.003	0.029
102	0.221	0.177	0.204	0.246
104	0.471	0.454	0.397	0.311
106	0.028	0.031		0.014
135	0.002			
total	8	8	7	7

A28	WCBP 94	WCBP 04	SCBP 93	SCBP 05
100		0.002		
101	0.034			
104	0.048	0.004		
123			0.005	0.003
126	0.003	0.002		
128	0.116	0.088	0.215	0.161
130	0.005			0.007
132			0.008	0.018
135	0.786	0.897	0.77	0.805
137		0.002		
139		0.002		0.003
141	0.005			
total	7	7	4	6

A88	WCBP 94	WCBP 04	SCBP 93	SCBP 05
130				0.007
136	0.062	0.046	0.083	0.094
139	0.288	0.266	0.339	0.279
141			0.006	0.003
143	0.003			
145	0.003		0.006	0.014
146	0.019	0.009	0.038	0.051
148	0.604	0.67	0.52	0.521
150	0.003	0.002	0.006	0.014

Table 5. Continued

141			0.006	0.003
143	0.003			
145	0.003		0.006	0.014
146	0.019	0.009	0.038	0.051
148	0.604	0.67	0.52	0.521
150	0.003	0.002	0.006	0.014
152	0.016	0.005		0.014
total	8	6	7	9

Ap66	**WCBP 94**	**WCBP 04**	**SCBP 93**	**SCBP 05**
87			0.003	
93	0.159	0.239	0.077	0.103
95	0.539	0.525	0.654	0.676
97		0.007		0.004
100			0.232	0.163
101	0.235	0.117	0.009	0.01
102		0.007		
103	0.061	0.022	0.019	0.012
104		0.044		
106		0.011		
107				0.004
113		0.003		0.012
115	0.003	0.003		
117			0.003	
129				0.004
135		0.018		0.004
total	5	11	7	10

B124	**WCBP 94**	**WCBP 04**	**SCBP 93**	**SCBP 05**
210		0.002		
212	0.014	0.002	0.009	0.011
214	0.355	0.219	0.268	0.42
216	0.187	0.288	0.222	0.061
218	0.043	0.099	0.03	0.007
220	0.133	0.159	0.147	0.192
222	0.092	0.036	0.114	0.112
224	0.086	0.089	0.114	0.091
226	0.041	0.031	0.033	0.036
228	0.008	0.005	0.006	0.003
230	0.014	0.055	0.036	0.025
232	0.014	0.005	0.003	
234	0.003		0.012	0.025
236			0.003	0.011

Table 5. Continued

242	0.005	0.002		0.003
244		0.002		
total	13	14	13	13

A7	**WCBP 94**	**WCBP 04**	**SCBP 93**	**SCBP 05**
80	0.006	0.004		
94	0.009			
96	0.003		0.003	0.004
98	0.003			0.008
102	0.012	0.031	0.023	0.047
104	0.102	0.097	0.142	0.119
106	0.006			0.008
108	0.25	0.282	0.322	0.286
110	0.009	0.005	0.011	0.029
112	0.18	0.128	0.159	0.119
114	0.105	0.094	0.121	0.158
116	0.129	0.165	0.103	0.149
118	0.003	0.002	0.009	
120	0.003		0.005	
124	0.003	0.005	0.006	
126			0.003	
128	0.012		0.003	
130	0.033	0.019	0.021	0.017
132	0.021	0.022	0.003	
134	0.009	0.002		
136	0.006	0.002		
138	0.048	0.082	0.017	0.021
140	0.0151	0.024	0.023	0.017
142				0.004
151	0.003	0.002		
155	0.009	0.005		
157	0.012	0.019	0.014	0.008
159			0.003	
168	0.003		0.003	
170		0.002		
171	0.003			
214		0.002		
total	27	21	20	15

Table 5. Continued

Ap113	WCBP 94	WCBP 04	SCBP 93	SCBP 05
145		0.003		
198	0.003			
202	0.061	0.043	0.127	0.058
214	0.301	0.251	0.154	0.227
216	0.031	0.074	0.006	0.022
218	0.039		0.006	
220	0.432	0.505	0.563	0.508
222	0.054	0.059	0.027	0.044
224	0.003	0.01	0.006	0.013
226	0.039	0.028	0.048	0.035
228	0.027	0.017	0.03	0.031
230	0.003	0.005	0.006	0.004
233	0.003		0.009	0.008
235		0.003	0.012	0.044
236			0.003	
total	12	11	13	11

Ap43	WCBP 94	WCBP 04	SCBP 93	SCBP 05
110		0.003		
114	0.007			
116	0.007			
121	0.007			
127	0.04			
131				0.015
133				0.007
135	0.06	0.04	0.11	0.073
137	0.14	0.086	0.113	0.029
139	0.013		0.026	0.029
141	0.033	0.075	0.004	
143	0.08	0.183	0.136	
145	0.426	0.433	0.428	0.397
147	0.117	0.111	0.09	0.4
149	0.016	0.019	0.015	0.029
151	0.003			
155		0.005		
165	0.033	0.013	0.045	
167		0.013		
169	0.003		0.007	
171		0.003	0.011	
173	0.003			
175		0.005		
177	0.007	0.003		0.015
187			0.008	
189	0.003		0.008	
total	18	14	13	9

Table 5. Contiued

Ap55	WCBP 94	WCBP 04	SCBP 93	SCBP 05
150	0.006			
171	0.019	0.011	0.025	0.005
173	0.191	0.231	0.148	0.081
175	0.426	0.419	0.412	0.518
177	0.092	0.091	0.058	0.057
179	0.105	0.088	0.093	0.042
181	0.082	0.113	0.167	0.133
183	0.015	0.005	0.032	0.033
185	0.012	0.011	0.019	0.019
187	0.009	0.008	0.032	0.009
189	0.006	0.003		0.019
193	0.009	0.003	0.006	0.014
199	0.015	0.002	0.003	
202	0.006	0.011		0.005
total	14	13	11	12

A81	WCBP 94	WCBP 04	SCBP 93	SCBP 05
106			0.003	
108	0.021	0.03	0.021	
114	0.006			
116		0.005		
121	0.003			
125	0.011			0.004
127	0.261	0.309	0.274	0.29
129	0.008			0.004
133				0.008
135	0.594	0.615	0.671	0.66
137	0.017	0.017	0.021	0.013
139	0.032	0.007	0.003	0.017
141	0.014	0.005	0.006	
143	0.003	0.005		
145	0.008			
147	0.005			
150	0.003			
152	0.003			
155	0.006	0.002		
163		0.002		
total	16	10	7	7

Figure 1. A map of the two distinct commercial queen breeding regions of the United States showing collection sites from 2004-2005.

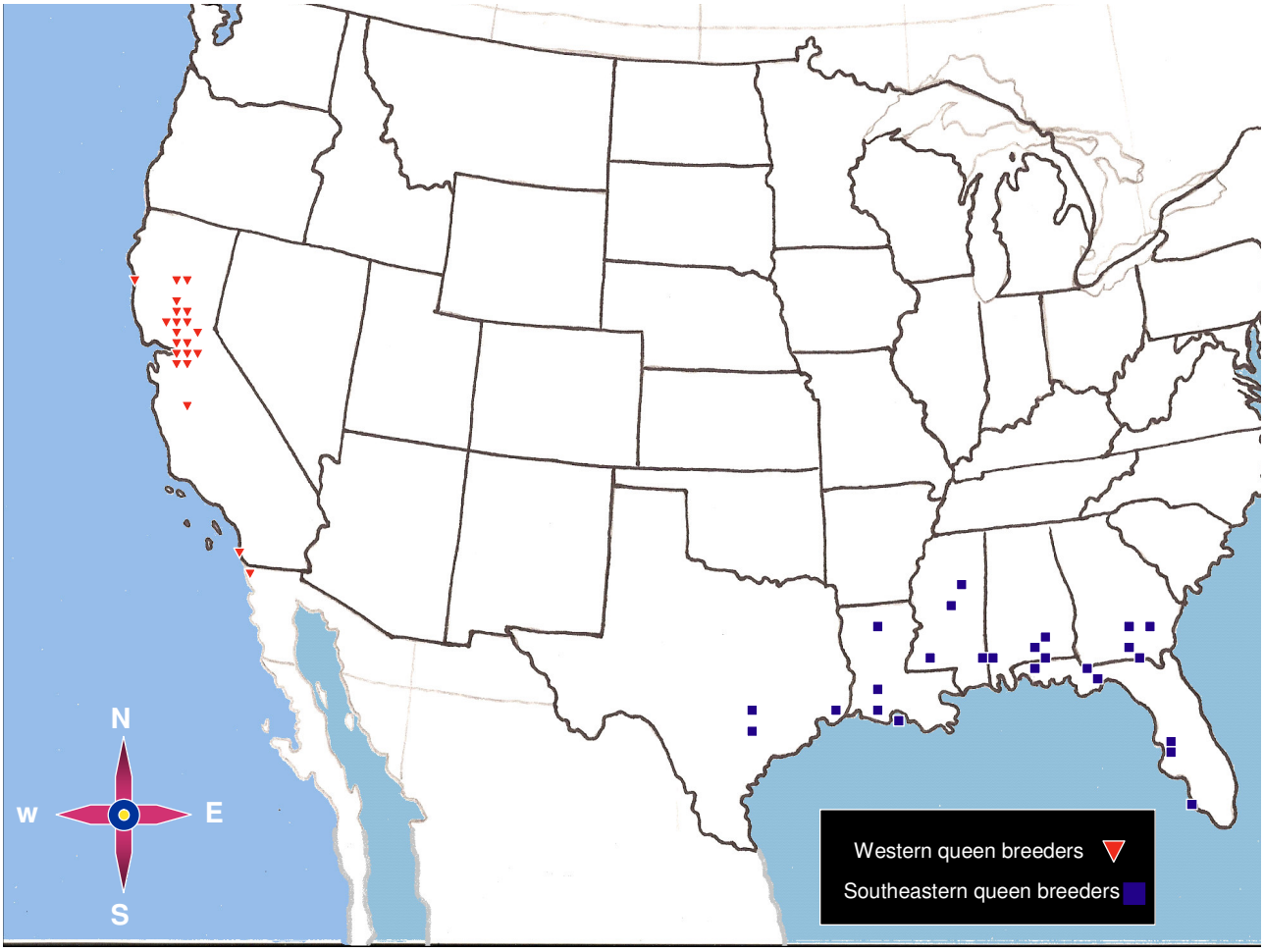
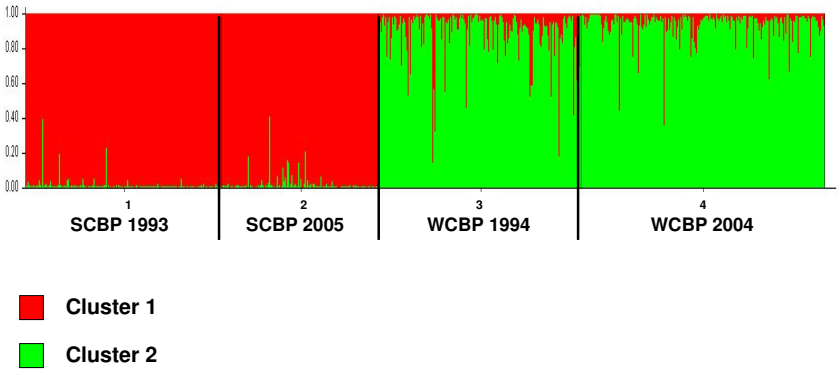


Figure 2. a). Bar plot produced by STRUCTURE assuming 2 populations showing the SCBP and WCBP samples clustering into different populations



b). Proportion of memebership of each pre-defined population in 2 clusters.

Given population	Inferred cluster 1	Inferred cluster 2	Known population
1	0.983	0.017	SCBP1993
2	0.976	0.024	SCBP2005
3	0.107	0.893	WCBP1994
4	0.058	0.942	WCBP2004

LIST OF TABLES

LIST OF FIGURES

GENETIC CHARACERIZATION OF U.S. FERAL HONEY BEE POPULATIONS USING MITOCHONDRIAL AND MICROSATELLITE MARKERS

Abstract

By Deborah A. Delaney, Ph.D.
Washington State University
August 2008

Chair: Walter S. Sheppard

The genetic composition of 692 feral honey bee nests collected in the United States between 1980 and 1992 was analyzed by assessing variation in the COI-COII intergenic spacer region of the mitochondrial DNA and 10 microsatellite loci. The genetic diversity and population structure of these pooled populations were compared to that of managed breeding populations in the southeastern and western U.S.

The mtDNA analysis revealed that 28% of the feral samples had haplotype patterns representative of the subspecies *Apis mellifera mellifera* or *Apis mellifera iberiensis.* Of the samples collected from Arizona, over 50% showed a haplotype pattern characteristic of honey bees found in Spain and/or Italy (M7). A subset of samples from the feral population with the M7 haplotype were sequenced and shown to be previously unknown variants of M7. These variants differed from the M7a and M7b variants known from Spain and Italy, respectively, leaving their origin unknown. Haplotype patterns characteristic of the European races most commonly used by commercial queen breeders were found in 66% of the feral samples.

Based on the microsatellite analysis, significant allelic and genotypic differences were found between most of the states, and genetic differences were observed between the feral samples and the managed breeding populations based on F_{st} values and Bayesian analysis of allele frequencies. The allelic and mtDNA diversity found in feral populations suggests that they may have been an important source of genetic diversity for U.S. managed honey bee populations.

Introduction

Most populations exhibit some degree of substructuring; i.e.differences in genetic variation within the population (Hedrick 2000). Genetic structuring can derive from different evolutionary processes including drift, founder events, non-random mating and selection (Hartl and Clark 1997). Many studies have shown that geographic and climatic ecological barriers, historical events and biological attributes can affect the genetic structure of populations (Donnelly and Townson 2000, Gerlach and Musolf 2000, Balloux and Lugon-Moulin 2002).

The "western" honey bee, *Apis mellifera* L, has an endemic range that includes Africa, Europe and western and central Asia. Across these divergent climatic and ecological regions, substantial intraspecific variation is evident, with around 26 described subspecies grouped into four evolutionary lineages: M (western and northern Europe), C (southeastern Europe), A (Africa) and O (Eastern Mediterranean) (Ruttner, 1988; Sheppard et al. 1997; Sheppard and Meixner, 2003). Mitochondrial analyses have yielded largely congruent classification of subspecies into "mitochondrial lineages", although subspecies of the C and O evolutionary lineages are not typically differentiated with most mitochondrial approaches (Garnery et al. 1992, Franck et al. 1998, Franck et al. 2001).

Although microsatellite and mitochondrial DNA (mtDNA) have been used to assess the genetic structure of honey bee populations throughout its native range and in Africanized regions of the New World (Hall and Muralidharan 1989, Smith et al. 1989, Rinderer et al. 1991, Sheppard et al. 1991, Franck et al. 1998, Clarke et al. 2001, Clarke et al. 2002,

Franck et al. 2001, Arias et al. 2006, Pinto et al. 2005, Collet et al. 2006, Kraus et al. 2007), studies addressing the genetic diversity of honey bee populations in the U.S. were performed prior to the decimation of the feral population due to *Varroa destructor* in areas free from Africanized honey bees (Sheppard 1988, Schiff and Sheppard 1993, Schiff et al. 1994, Schiff and Sheppard 1995, 1996, Kraus and Page, 1995, Wenner and Bushing, 1996, Sanford, 2001, Seeley, 2007). These studies of U.S. honey bee populations provide baseline data useful for comparison with data sets from these same populations collected after the arrival of *Varroa destructor* and Africanized bees, such a time series approach can measure changes in the genetic structure of these populations in response to parasite establishment and invasive movement of Africanized honey bees.

U.S. honey bee populations were established over 400 years ago during the early settlement of North America (Sheppard, 1989; Horn, 2005). Between 1621 and 1922 eight subspecies were successfully imported into North America (Sheppard, 1989; Table 1). In 1990, Africanized honey bees, descendents of an additional subspecies, *Apis mellifera scutellata,* moved into North America and became established in the U.S. However, it was during the early settlement of North America when feral populations, largely composed of descendents of the subspecies *Apis mellifera mellifera*, became established (Sheppard, 1989). These feral or unmanaged populations flourished well into the early 1990s until the arrival and dispersal of a parasitic brood mite, *Varroa destructor* (Anderson and Trueman) drastically reduced these populations (Wenner and Bushing, 1996; and Sanford 2001). The current status and genetic composition of feral populations 20 years after the arrival of *Varroa destructor* remains largely unknown.

The U.S. honey bee population can be divided into two sub-populations. The managed population is composed of populations managed and maintained by beekeepers and bee breeders in moveable frame hives (Schiff et al. 1994). The second population is the unmanaged feral population that consists of bees living in a variety of nest sites, both artificial and natural. Schiff and colleages (Schiff et al. 1994; Schiff and Sheppard 1995, 1996) analyzed these two honey bee populations in the U.S. using allozyme and mitochondrial DNA variation. They reported significant heterogeneity among U.S. feral honey bees and found that feral populations contained genetic remnants from early honey bee importation events. Subsequent studies showed that the western and southeastern managed breeding populations contained little within population variation, although *Mdh* allele frequencies were significantly different between the two managed breeding populations (Schiff and Sheppard 1995, 1996). Differences were also found between the western managed population and U.S. feral populations (Schiff and Sheppard 1996). These findings suggested reduced gene flow between the three populations.

Recent (1992-present), drastic reductions of feral populations and the annual 50% loss of managed honey bee populations in the U.S. (National Research Council, 2007) have raised concern over the genetic integrity of U.S. honey bee stock. Because significant heterogeneity was found in U.S. feral populations sampled between 1980 and 1992, it was thought that feral populations once served as a possible source of genetic diversity for the less diverse managed honey bee populations. Therefore, the loss of the feral populations may have resulted in the loss of a possible source of genetic diversity.

The purpose of this study was to characterize microsatellite and mitochondrial DNA variation in previously collected feral honey bee populations. These new markers provide a greater a better estimate of diversity for these populations. The genetic differences measured for feral populations from 11 southern states were compared to similar data from two previously analyzed managed breeding populations. A further goal of the study was to determine whether the apparent loss of feral populations had a measurable effect on the genetic composition of managed breeding populations. Mitochondrial haplotype and microsatellite allele frequencies were established for feral populations. Pairwise f_{st} values and heterozygosity measures were calculated to determine the genetic structure of feral populations. Data gathered from the feral populations and the managed breeding populations from 1993-1994 and 2004-2005 were compared.

Materials and methods

Collection protocol

Feral population

From 1980 to 1992, samples of bees from 692 feral nests were collected (Schiff et al. 1994). Of these 692 samples, 546 were used for the current mtDNA analysis and 625 were used for microsatellite analysis. This collection represented feral populations that existed in the southern United States. Samples were collected from 11 states: Alabama, Arizona, Georgia, Louisiana, Mississippi, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee and Texas. For analysis, the Oklahoma samples (n=3) were

combined with the Texas samples, and the Tennessee samples (n=2) were combined with the Georgia samples due to small sample size and geographic proximity. Adult workers were frozen in liquid nitrogen for transportation and stored at -80°C until analysis.

Molecular analysis

DNA extraction

One adult worker leg or thorax from each colony was used to extract both nuclear and mitochondrial DNA (mtDNA) following the chelex method (Walsh et al. 1991). The chelex /honey bee preparation was placed in the thermocycler for 1 h at 55°C, 15 min at 99°C, 1 min at 37°C and 15 min at 99°C. Ninety-six well PCR plates of extracted honey bee DNA were stored in the freezer at -80°C for future analysis and amplification.

Mitochondrial DNA analysis

The COI-COII intergenic region of the mtDNA was polymerase chain reaction (PCR) amplified from 1.25 µl of DNA extraction product from each sample with primers E2 and H2 in 25 µl reactions (Garnery et al. 1992). PCR conditions for the amplified region were as follows: 30 cycles at 92°C for 30 s, 1.5 min at 47°C and 2 min at 63°C. Aliquots of 6 µl of the PCR product were electrophoresed through 1.4% agarose gel, and fragments were visualized with ethidium bromide. The remainder of the amplified PCR product (19µl) was digested with 4 units of the restriction enzyme *Dra*I (Garnery et al. 1993). The digested DNA products were separated on a 10% acrylamide gel and visualized with ethidium bromide. Haplotype names were assigned to each sample based on fragment patterns described in the literature (Garnery et al. 1993, Garnery et al. 1995,

Garnery et al. 1998, Franck et al. 1998, Franck et al. 2000a, Franck et al. 2000b and Franck et al. 2001). For problematic samples, previously stored extractions based on a total nucleic acid protocol (Sheppard and McPheron 1986) were used. The software program ARLEQUIN 3.1 (Excoffier, L.G. Laval, and S. Schneider, 2005) was used to calculate pairwise F_{st} values for each population pair.

The COI-COII intergenic region of two to three samples with the M7 haplotype from each state was sequenced. The PCR conditions for the cycle sequence protocol was 25 cycles of 96°C for 10 sec, 50°C for 5 sec, 60°C for 4 min. Products were sequenced on a ABI 3730 automated sequencer (Applied Biosystems). Alignment and sequence comparisons were made using Multalign (Corpet 1988).

Microsatellite analysis

Individual honey bee samples from feral colonies (N = 625) collected in 12 southern states were analyzed for variation at 10 microsatellite loci: A7, A24, A28, A88, A113, B124 (Estoupe et al. 1995), Ap43 (Garnery et al. 1998), Ap55, Ap66 and Ap81 (Solignac et al. 2003). These were compared to data taken from managed breeding populations from the western and southeastern U.S. that had been sampled both in 1993-1994 and 2004-05 (Delaney et al. in review).

Two multiplex PCRs were used to amplify 10 µl reactions of extracted DNA containing 1.0 - 4.0 mM of flouresecent dye-labeled primer (Applied Biosystems), 1x reaction buffer, 1.5 units *Taq* polymerase, 3mM dNTP mixture, 0.001 mg bovine serum albumin,

1µl of DNA extraction product and 1.2 to 1.5 mM of $MgCl_2$ (Delaney et al. in review). The reaction products were analyzed using an ABI Biosystems 3730 automatic sequencer (Applied Biosystems). The allele sizes at each microsatellite locus were determined with the aid of GeneMapper™ software (Applied Biosystems).

Statistical analysis

The population genetics software GENETIX™ was used to calculate allele frequencies and to convert files into formats appropriate for other software programs. Hardy-Weinberg equilibrium, heterozygote deficiency, heterozygosity (observed and expected), genic and genotypic differentiation and population pairwise F_{st} values were estimated using the GENEPOP package (web version 3.4; Raymond and Rousset 1995). The average number of alleles per locus (n_a), allelic richness (a_g) and expected heterozygosity (H_e) were calculated for each feral population (Leberg 2002) using the software HP-RARE which corrects these diversity measurements for sample size variation (Kalinowski, 2005). Reductions in effective population size were determined for each sample grouping (most by state) by calculating excesses in heterozygosity using three kinds of statistical tests: sign test, standardized differences test and the Wilcoxon's signed rank test found in the software program BOTTLENECK (Piry et al. 1999).

Results

Mitochondrial DNA

DraI RFLP analysis of the COI-COII intergenic region revealed nine different haplotypes in the feral population (Figure 1): three from mitochondrial lineage C (C1, C2, and C3),

five from lineage M (M3, M4, M4', M7, M7') and one from A lineage. The haplotypes for four percent of the samples could not be identified due to degraded DNA.

The number of haplotypes found in each sample population from each state can be seen in Table 2. Honey bee populations sampled from North Carolina displayed the fewest number of haplotypes (3), whereas populations from the rest of the states each exhibited between five and seven of the nine reported haplotypes. The M7 haplotype, reported from *Apis mellifera iberiensis* and *Apis mellifera ligustica* (Franck et al. 2000), was found in 57% of the samples from Arizona. The haplotype M3, characteristic of subspecies of western European origin and also found in Italy (Kraus et al. 2007, Franck et al. 2000), was found in low frequency (0.028 to 0.122) in populations sampled from five states. Haplotypes M4 and M4' occur in honey bees from France and Belgium and more recently variants of M4 were found in Italy (Garnery et al. 1998, Franck et al. 1998, Franck et al. 2000). These two M haplotypes were found in 7 % of the feral samples and were more common east of the Mississippi River, reaching a maximum of 19% in samples from Georgia. Two percent of the feral population had a haplotype pattern characteristic of the African subspecies, *Apis mellifera lamarckii.* This subspecies is endemic to the Nile River Valley and was introduced into North America in the 1860s (Sheppard 1989, Schiff et al. 1994). The most commonly occurring haplotypes were from the C lineage: C1 and C2. These haplotypes are common in *Apis mellifera ligustica* and *Apis mellifera carnica,* respectively,whose descendents form the basis for the commercial strains currently favored by U.S. beekeepers (Franck et al. 2000, Schiff and Sheppard 1995). Together these two haplotypes constituted 63 % of the feral population sampled

(Figure 1.). The C2 haplotype occurred in 75% of the colonies sampled from North Carolina.

Sequence comparisons of the COI-COII intergenic region for 13 samples (sequences submitted to genbank, acession numbers EU831095-EU831107) with the M7 haplotype uncovered two substitutions characteristic of the M7a variant found in honey bee samples from Spain (Franck et al. 2000) (Figure 2). However, a deletion that is characteristic of the published M7a variant was not found. Other substitutions and deletions were also found in the feral sample sequences, further differentiating these sequences from the published M7 variants (Franck et al. 2000).

Microsatellite DNA

The average number of alleles varied from 6.2 ± 2.6 (Louisiana) and 9.9 ± 3.7 (Arizona) (Table 3). Allelic richness measures after rarefaction ranged from 4.7 ± 1.7 in Mississippi to 5.5 ± 1.8 in Alabama (Table 3). The average gene diversity measured as expected heterozygosity was similar across all the states sampled and ranged from 0.60 ± 0.2 for New Mexico to 0.66 ± 0.08 for Alabama (Table 3).

The proportion of alleles of feral "origin" that were found in managed populations decreased between 1993-1994 and 2004-2005. The feral samples contained 138 different microsatellite alleles across the 10 loci. Of these 138 alleles, 111 alleles (80%) were found in the managed populations sampled from 1993 and 1994 and 102 alleles (74%) were found in managed populations sampled in 2004 and 2005.

The proportion of alleles from the managed populations that may have originated from feral populations increased in the western breeding population between 1994 and 2004. Ninety-nine out of 128 alleles (77%) found in the western managed breeding population sampled in 1994 were also found in feral populations, whereas 89 out of 114 alleles (87%) found in the western managed breeding population sampled in 2004 were also found in feral populations. The proportion of alleles from the southeastern managed population that could have originated from feral populations decreased between 1993 and 2005. Ninety-two out of 102 alleles (90%) found in the southeastern managed breeding population sampled in 1993 were also found in feral populations, and 85 out of 99 alleles (85%) found in the southeastern managed breeding population sampled in 2005 were also found in feral populations.

No significant linkage disequilibrium was detected in any of the 45 locus pairs for each state (at the 5% level). The exact test for Hardy-Weinberg equilibrium (HWE) found only population samples from Mississippi and New Mexico to be in HWE ($0.451 \leq P \leq 0.580$). The other six population samples (Texas, North Carolina, Arizona, Alabama, Georgia and South Carolina) significantly deviated from HWE ($0.000 \leq P \leq 0.001$). *U*-tests revealed significant heterozygote deficiency ($0.000 \leq P \leq 0.001$) for each of the six sample populations not in HWE likely due to combining multiple subpopulations by sampling design.

Population pairwise F_{st} values are reported in Table 4. Population samples from Georgia and Louisiana showed little genetic differentiation ($F_{st} = 0.003$) as did other intra-region comparisons. In contrast, inter-regional comparisons such as between population samples from Alabama and Arizona ($F_{st} = 0.065$) and Alabama and New Mexico ($F_{st} = 0.068$) showed moderate genetic differentiation. The F_{st} values between most of the feral honey bee populations sampled were one order of magnitude higher than between the two managed breeding populations. The pooled feral honey bee populations sampled from all the states were highly diverged from all the commercial populations, based on F_{st} values that ranged from 0.258 to 0.313.

The genic differentiation test (Raymond and Rousset, 1995) performed on population pairs across all loci (Table 5) showed significant allelic differentiation between 8 population samples ($0.000 \leq P \leq 0.011$). Population samples from Louisiana and Georgia did not show significant genic differentiation ($P = 0.305$.). This supports the insignificant F_{st} values found for this population pair. Genotypic differentiation for each population pair across all loci (Table 5) was significant for 75 out of 78 population pairs. Three population pairs did not exhibit significant genotypic differences: Louisiana and Texas ($P = 0.215$), Georgia and Louisiana ($P = 0.601$) and Alabama and Georgia ($P = 0.131$). The number of alleles/locus and allele frequencies are reported in Table 6.

Sample populations from North Carolina, South Carolina and Mississippi exhibited significant bottlenecks when analyzed with the BOTTLENECK software. The values for population samples from North Carolina were statistically significant for the standardized

differences test ($P = 0.036$) and the Wilcoxon's signed rank test ($P = 0.003$). Population samples from Mississippi also had significant values for the standardized differences test ($P = 0.049$) and the Wilcoxon's signed rank test ($P = 0.012$). Population samples from South Carolina exhibited significant excess heterozygosity using the sign test ($P = 0.037$) and the Wilcoxon's signed rank test ($P = 0.004$). The results were unchanged when loci that were not in HWE were removed from the analysis.

Discussion

Because of the many beneficial services honey bees provide, humans have expanded the original range of honey bees globally. The history of the importation of honey bee subspecies is based on ship logs, journals and advertisements. It has been proposed that the first honey bees to arrive in North America entered into the U.S. from Mexico in the 1500s from early Spanish settlements (Brand, 1988), although there was no direct evidence presented to support this theory. The feral collection used in this study represents a sampling of the U.S. feral honey bee population between 1980 and 1992, almost 400 years after the first recorded arrival of honey bees into the New World. The ancestry found within these sample populations reflect the history of subspecific importations into the New World from the Old World during the settlement of North America by European settlers.

The mitochondrial DNA analysis of the feral populations clearly demonstrates a more diverse maternal ancestry than the managed breeding populations we sampled in 1993-1994 and 2004-2005 (Figure 3). The feral populations had a higher proportion of

mitochondria from the M lineage than the managed breeding populations sampled in 1993 and 1994. Twenty-eight percent of the feral colonies had mtDNA haplotypes characteristic of the M mitochondrial lineage, whereas 4 % of the managed breeding colonies sampled from 1993 and 1994 had mtDNA haplotypes attributed to the M lineage (Delaney et al. in review). This lineage includes the subspecies *Apis mellifera mellifera* and *A. m. iberiensis* (Garnery et al. 1998). These findings support the theory that *A. m. mellifera* established a feral population and flourished in North America during its settlement.

The high occurrence of mitochondria from the M lineage is especially apparent in populations sampled from Arizona. The M7 haplotype was found in 57% of the sample populations from Arizona. Founder events exacerbated by genetic drift and geographic isolation could explain the particular high frequency of this haplotype in these populations. The M7 haplotype was thought to be unique to Spain and therefore representative of *A. m. iberiensis* (Garnery et al. 1998), but has since been discovered in Italy (Franck et al. 2000, Clarke et al. 2001). Franck et al. 2000 reported that two variants of M7 (*a* and *b*) could be differentiated based on sequence differences in the COI-COII intergenic region. The variant M7*a* was observed in Spain, whereas the M7*b* variant was found in Italy. However, the M7 pattern we found in U.S. feral honey bee samples exhibited sequence differences from both variants reported by Franck et al. (2000).

The origin of this haplotype is of particular interest because of the implications it has to determine the source of the early North American honey bee importations. Additional sequencing studies on honey bees from the Old World with M7 haplotypes will be necessary to identify the origin of these new variants and to fully resolve the chronology of importation events into North America (Watkins, 1968; Brand, 1988; Sheppard, 1989). It is possible that these new variants arose in the U.S. feral populations through mutations.

MtDNA and microsatellite data verified that feral populations were an important source of diversity for the managed populations in the U.S. However, due to recent bottlenecks imposed by breeding practices and impediments to gene flow, such as geographic barriers, the two managed sub-populations have retained significant genetic differences. The mtDNA data help to explain the effect that humans have had on U.S. honey bee populations. Over a century of selecting for Italian and Carniolan queen mothers led to the dominance of these subspecies in terms of maternal ancestry. Of the managed breeder population samples from 1993 and 1994, 96% have mtDNA characteristic of the C mitochondrial lineage which is found in subspecies *Apis mellifera ligustica* and *A. m. carnica.*

The microsatellite analysis of the feral honey bee samples we observed supports previous findings that found the genetic composition of the feral population was more heterogeneous than the managed breeding populations (Schiff et al., 1994). The feral populations were highly differentiated from the commercial breeding populations based

on F_{st} values. In addition, significant differences were found in the genotypic and allelic composition between the 1980-1992 feral populations sampled and the managed breeding populations. Together, these findings suggest that barriers to gene flow existed between feral populations and both managed breeding populations and may continue to this day.

Reports of established feral populations east of the Mississippi early in the 18th century are well documented. The establishment of these feral populations in this region supports the large proportion of shared alleles (90%) between the feral populations sampled and the southeastern managed population. The decreased amount of alleles (77%) shared between the feral population and the western managed population is not surprising due to the geographical distance and topography that separates these two regions. What is surprising was the increase in shared alleles (87%) between the feral population and western managed population in 2004. New alleles were found in both the managed populations in 2004 and 2005 and these alleles were known to occur in the feral population samples collected between 1980-1992 (Delaney et al., in review). The decrease in shared alleles (85%) between the feral population and the southeastern managed population in 2005 could be a result of the loss of allelic input once provided by the allelically abundant feral population. Therefore, transfer of alleles from the feral population to the western managed population in 2004 and the decrease in shared alleles between the feral population and southeastern managed population sampled in 2005 implies that the feral population was likely a source of allelic variation for both managed populations.

Analyses of genetic variation using molecular tools have been used to explain species dispersal (Ibrahim et al. 1996; Eldridge et al. 2001), relationships between population genetic structure and environmental variables (Gerlach and Musolf, 2000), historical information and phylogenies of domesticated animals and plants (Fan et al. 2002; Geffen et al. 2004). Examining genetic structure and assessing diversity in natural populations has been critical in understanding the affects of reductions in effective breeding size and in developing appropriate conservation management strategies (Hoelzel et al., 2002; Harley et al., 2005; and Brown and Gladden et al., 1999).

Conclusions

This study set out to characterize the genetic composition and allelic variation of representative feral honey bee populations collected in the U.S. between 1980 and 1992 to determine if they could have provided genetic variation to potentially inbred, managed honey bee populations. The variation found in mtDNA and microsatellites in feral populations was compared to that of the managed honey bee populations collected at two different time periods: 1993-1994 and 2004-2005 from two regions, southeastern U.S. and western U.S. The collection times of these managed populations are significant because the most recent collection occurred after reductions in feral populations, and the earlier collection occurred while the feral population still flourished. These comparisons enable us to use the perturbation of feral populations (their dramatic reduction) to test if there is evidence that the feral population served as a source of genetic variation for the managed populations.

The results showed that feral populations had more allelic diversity than either of the managed populations collected in the early 1990s. Also, the mtDNA haplotypes diversity was greater in pooled feral populations than managed breeding populations. The higher diversity found in feral populations may reflect accumulated diversity from the history of multiple subspecific importations from the Old World during the settlement of North America by European settlers. The increase in the proportion of alleles shared between pooled feral populations and the western managed population collected in 2004, the decrease in the proportion of shared alleles between the feral population and the southeastern managed population collected in 2005 and the identification of known feral alleles newly found in both of the managed populations collected in 2004-2005 strongly indicate that the feral population was a source of allelic variation for both managed populations.

Bibliography

Arias MC, Rinderer TE, Sheppard WS (2006) Further characterization of honey bees from the Iberian Peninsula by allozyme, morphometric and mtDNA haplotype analysis. *Journal of Apicultural Research,* **45**, 188-196.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology,* **11**, 155-165.

Bottin L, Vaillant A, Sire P, Cardi C, Bouvet JM (2005) Isolation and characterization of microsatellite loci in *Santalum austrocaledonicum,* Santalaceae*. Molecular Ecology Notes*, **5**, 800-802.

Brand D (1988) The honey bee in New Spain and Mexico. *Journal of Cultural Geography,* **9**, 71-82.

Brown Gladden JG, Ferguson MM, Friesen MK, Clayton JW (1999) Population structure of North American beluga whales (*Delphinapterus leucas*) based on nuclear DNA microsatellite variation and contrasted with the population structure revealed by mitochondrial DNA variation. *Molecular Ecology,* **8**, 347-363.

Clarke KE, Oldroyd BP, Javier J, Quezada-Euán G, Rinderer TE (2001) Origin of honeybees (*Apis mellifera* L.) from the Yucatan peninsula inferred from mitochondrial DNA analysis. *Molecular Ecology,* **10**, 1347-1355.

Clarke KE, Rinderer TE, Franck P, Quezada-Euán JG, Oldroyd BP (2002) The Africanization of honeybees (*Apis mellifera* L.) of the Yucatan: a study of a massive hybridization event across time. *Evolution,* **56**, 1462-1474.

Corpet F (1988) "Multiple sequence with hierarchal clustering" *Nucleic Acids Res.*, **16**, (22), 10881-10890.

Collet T, Ferreira KM, Arias MC, Soares AEE, Del Lama MA (2006) Genetic structure of Africanized honeybee populations (*Apis mellifera* L.) from Brazil and Uruguay viewed through mitochondrial DNA COI-COII patterns. *Heredity,* **97**, 329-335.

Donnelly MJ,Townson H (2000) Evidence of extensive genetic differentiation among populations of the malaria vector *Anopheles arabiensis.* Eastern Africa. *Insect Molecular Biology,* **9**, 357-367.

Eldridge MDB, Kinnear JE, Onus ML (2001) Source population of dispersing rock-wallabies (*petrogale lateralis*) identified by assignment tests on multilocus genotypic data. *Molecular Ecology,* **10**, 2867-2876.

Excoffier LG, Laval, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1**,47-50.

Franck P, Garnery L, Solignac M, Cornuet JM (1998) The origin of west European subspecies of honeybees (*Apis mellifera)*: new insights from microsatellite and mitochondrial data. *Evolution,* **52**, 1119-1134.

Franck P, Garnery L, Celebrano G, Solignac M, Cornuet J-M (2000a) Hybrid origins of honeybees from Italy (*Apis mellifera ligustica)* and Sicily (*A. m. sicula*). *Molecular Ecology,* **9**, 907-921.

Franck P, Koeniger N, Lahner G, Crewe RM, Solignac M (2000b) Evolution of extreme polyandry: an estimate of mating frequency in two African honeybee species, *Apis mellifera monticola* and *A.m. scutellata. Insect. Soc.,* **47**, 364-370.

Franck P, Garnery L, Loiseau A, Oldroyd BP, Hepburn HR, Solignac M (2001) Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. *Heredity,* **86**, 420-430.

Freeman and Herron (2004) Evolutionary Analysis 3rd Ed. Pearson Education Inc. Upper Saddle River, New Jersey. 802 pp.

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

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

Garnery L, Mosshine EH, Oldroyd BP, Cornuet J-M (1995) Mitochondrial DNA variation in Moroccan and Spanish honey bee populations. *Molecular Ecology,* **4**, 465-471.

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

Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology*, **13**, 2481-2490.

Gerlach G, Mosolf KF (2000) Fragmentation of landscape as a cause for genetic subdivision in bank voles. *Conservation Biology,* **14**, 1066-1074.

Griffiths SW, Armstrong JD (2001) The benefits of genetic diversity outweigh those of kin association in a territorial animal. *Proc. Royal Soc. Lond. B,* **268**, 1293-1296.

Hall HG, Muralidharan K (1989) Evidence from the mitochondrial-DNA that African honey bees spread as continuous maternal lineages. *Nature,* **339**, 211-213.

Harley EH, Baumgarten I, Cunningham J, O'ryan C (2005) Genetic variation and population structure in remnant populations of black rhinoceros, *Diceros bicornus*, in Africa. *Molecular Ecology,* **14**, 2981-2990.

Hartl D.L. and A.G. Clark (1997) Principles of Population Genetics 3rd ed. Sinauer Associates Inc. Sunderland, MA, 542 pp.

Hedrick P.W. (2000) Genetics of Populations 2nd ed. Jones and Bartlett Publishers, Sudbury, MA, 553 pp.

Hoelzel AR, Fleischer RC, Campagna C, Le Boeuf BJ, Alvord G (2002) Impact of a population bottleneck on symmetry and genetic diversity in the northern elephant seal. *Journal of Evolutionary Biology,* **15**, 567-575.

Hollingsworth PM, Dawson IK, Goodall-Copestake WP, Richardson JE, Weber JC, Sotelo Montes C, Pennington RT (2005)

Hughes WOH, Boomsa JJ (2004) Genetic diversity and disease resistance in leaf-cutting ant societies. *Evolution*, **58**, 1251-1260.

Horn T. (2005) Bees in America: How the Honey Bee Shaped a Nation.The University Press of Kentucky: Lexington. pp. 352.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282-291.

Kalinowski ST (2005). HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes,* **5**, 187-189.

Kraus B, Page RE (1995) Effects of *Varroa jacobsoni* (Mesostigmata: Varroidae) on feral *Apis mellifera* (Hymenoptera: Apidae) in California. *Environ. Entomol*., **24**, 1473-1480.

Kraus B, Franck P, Vandame R (2007) Asymmetric introgression of African genes in honeybee populations (*Apis mellifera* L.) in Central Mexico. *Heredity,* **99**, 233-240.

Lande R (1988) Genetics and demography in biological conservation. *Science,* **241**, 1455-1459.

Leberg PL (2002) Estimating allelic richness: effects of sample size and bottlenecks. *Molecular Ecology,* **11**, 2445-2449.

Luikart G, Cornuet J-M (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology,* **12**, 228-237.

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology,***7**, 963-974.

Mattila HR Seeley, TD (2007) Genetic Diversity in Honey Bee Colonies Enhances Productivity and Fitness. *Science,* **317**, 362-364.

National Research Council (2007) Status of pollinators in North America. The National Academies Press, 500 Fifth Street, NW Washington DC 20001. pp 307.

Pinto MA, Rubink WL, Patton JC, Coulson RN, Spencer Johnston J (2005) Africanization in the United States: Replacement of Feral European Honeybees (*Apis mellifera* L.) by an African Hybrid Swarm. *Genetics,* **170**, 1653-1665.

Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity,* **90**, 502-503.

Pritchard JK, Stephens M, Donnelly P (2000) Genetics 155: 945

Raymond M, Rousset F (1995) Population genetics software for exact test and ecumenism. *J. Hered.,* **86**, 248-250.

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

Rinderer TE, Stelzer JA, Oldroyd BP, Buco SM, Rubink WL (1991) Hybridization Between European and Africanized Honey-Bees in the Neotropical Yucatan Peninsula. *Science,* **253,** 309-311.

Sanford MT (2001) Introduction, spread, and economic impact of *Varroa* mites in North America, in: Webster TC, Delaplane KS (Eds.), Mites of the honey bee, Dadant and Sons, Hamilton, Illinois, pp. 149-162.

Schiff N, Sheppard WS (1993) Mitochondrial DNA evidence for the 19th century introduction of African honey bees into the United States. *Experientia,* (Basel) **49**, 350-352.

Schiff N, Sheppard W, Loper G, Shimanuki H (1994) Genetic diversity of feral honey bee (Hymenoptera: Apidae) populations in the southern United States. *Ann. Entomol. Soc. Amer*., **87**, 842-848.

Schiff N, Sheppard W (1995) Genetic analysis of commercial honey bees (Hymenoptera: Apidae) from the Southeastern United States. *J. Econ. Entomol*., **88**, 1216-1220.

Schiff N, Sheppard W (1996) Genetic differentiation in the queen breeding population of the western United States. *Apidologie,* **27**, 77-86.

Seeley TD (2007) Honey bees of the Arnot Forest: a population of feral colonies persisting with *Varroa destructor* in the northeastern United States. *Apidologie*, **38**, 19-29.

Seeley TD, Tarpy DR (2007) Queen promiscuity lowers disease within honeybee colonies. *Proc. R. Soc. B*, **274,** 67-72.

Sheppard WS (1988) Comparative study of enzyme polymorphism in the United States and European Honey bee (Hymenopetra: Apidae) populations. *Annals of the Entomological Society of America*, **81**, 886-889.

Sheppard WS, McPheron BA (1986) Genetic variation in honey bees from an area of racial hybridization in western Czechoslovakia. *Apidologie,* **17**, 21-32.

Sheppard W (1989) A history of the introduction of honey bee races into the United States: Part I and II. *Amer. Bee Journal,* **129**, 617-619, 664-667.

Sheppard WS, Rinderer TE, Mazzoli JA, Stelzer JA Shimanuki H (1991) Gene Flow Between African-Derived and European-Derived Honey-Bee Populations in Argentina. *Nature* **349**, 782-784.

Sheppard W, Rinderer T, Garnery L, Shimanuki H (1999) Analysis of Africanized honey bee mitochondrial DNA reveals further diversity of origin. *Gen Mol Biol,* **22**:73-75.

Smith DR, Taylor OR Brown WM (1989) Neotropical Africanized Honey Bees Have African Mitochondrial DNA. *Nature* **339**, 213-215.

Spencer CC, Neigel JE, Leberg PL (2000) Experimental evaluation of the usefulness of microsatellite DNA for detecting demographic bottlenecks. *Molecular Ecology* **9**, 1517-1528.

Tarpy DR, Seeley TD (2006) Lower disease infections in honeybee (*Apis mellifera*) colonies headed by polyandrous vs monandrous queens. *Naturwissenschaften* **93**, 195-199.

Walker CW, Vila C, Landa A, Linden M, Ellegren H (2001) Genetic variation and population structure in Scandinavian wolverine (*Gulo gulo*) populations. *Molecular Ecology*, **10**, 53-63.

Walsh PS, Metzger DA, Higuchi R (1991) Chelex ® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, **10**:507
SPSS v12.1. 2003. SPSS Inc. Chicago, IL

Watkins LH (1968) California's first honey bees. *American Bee Journal,* **108,** 191-192.

Wenner AM, Bushing WW (1996) *Varroa* mite spread in the United States, *Bee culture,* **124**, 342-343.

Yasui Y (1998) The genetic benefits of female multiple mating reconsidered. *Trends in Ecology and Evolution,* **13**, 246-250.

Table 1.
The approximate arrival date, and geographic origin of each introduced subspecies of *Apis mellifera* into the Americas.

Subspecies	**Origin**	**Importation Date**
A. m. mellifera	Europe, North and west of Alps, central Russia	1622 (known to be present)
A. m. ligustica	Italy	1859
A. m. lamarckii	Egypt, Nile River valley	1866
A. m. carnica	Southern Austrian Alps and northern Balkans	1877
A. m. caucasica	Central Caucus Mountains	1880-1882
A. m. cypria	Cyprus	1880
A. m. syriaca	Syria, eastern Mediterranean region	1880
A. m. intermissa	Africa, north of Sahara	1891
A. m. scutellata	Central and eastern equatorial Africa to South Africa	1956

Table 2.
Number of haplotypes in each state of the feral population

	C haplotypes			M haplotypes					A haplotype
State	**C1**	**C2**	**C3**	**M3**	**M4**	**M4'**	**M7**	**M7'**	***Apis mellifera lamarckii***
TX (n=73)	21	34	2	2	1		10		3
NC (n=32)	6	24		2					
AZ (n=150)	20	35				2	86	4	3
LA (n=44)	15	16		2		3	7	1	
GA (n=52)	10	29			1	9	1		2
AL (n=33)	2	17	3		3	3	4		1
MS (n=36)	10	14	3	4	1		2		2
NM (n=57)	11	33	7	2			1	2	1
SC (n=69)	17	36	2		8	6			
Total	112	238	17	12	14	23	111	7	12

Table 3.

The average number of alleles (n_a), allelic richness (a_g) and expected heterozygosity (H_e) for each state

Population	n_a	a_g	H_e
TX	8.8±3.6	5.3±2	0.64±0.1
NC	5.8±2.5	5.1±2.1	0.65±0.1
AZ	9.9±3.7	5.4±1.7	0.62±0.1
LA	6.2±2.6	5.1±2	0.63±0.1
GA	6.6±2.3	5.4±1.8	0.65±0.1
AL	7.1±2.7	5.5±1.8	0.66±0.08
MS	6.2±2.9	4.7±1.7	0.63±0.09
NM	7.6±2.9	5.3±1.8	0.60±0.2
SC	7±2.4	5.1±1.6	0.64±0.1

Table 4.
Pairwise multilocus Fst estimates of 13 *Apis mellifera* populations based on microsatellite frequencies

	TX	NC	AZ	LA	GA	AL	MS	NM	SC	WCBP 94	WCBP 04	SCBP 93
NC	0.019											
AZ	0.022	0.021										
LA	0.013	0.018	0.032									
GA	0.015	0.021	0.034	-0.003								
AL	0.021	0.039	0.065	0.040	0.018							
MS	0.014	0.038	0.052	0.022	0.009	0.006						
NM	0.025	0.049	0.028	0.042	0.047	0.068	0.050					
SC	0.018	0.037	0.044	0.026	0.029	0.039	0.035	0.040				
WCBP 94	0.266	0.268	0.280	0.263	0.258	0.259	0.277	0.289	0.278			
WCBP 04	0.289	0.292	0.301	0.287	0.285	0.281	0.298	0.313	0.301	0.007		
SCBP 93	0.273	0.268	0.284	0.273	0.269	0.268	0.284	0.300	0.291	0.016	0.018	
SCBP 05	0.284	0.283	0.293	0.287	0.281	0.281	0.298	0.310	0.301	0.023	0.016	0.002

Table 5.
P-values for genotypic (above diagonal) and genic (below diagonal) differentiation for 13 *Apis mellifera* populations based on microsatellite data

	NC	**TX**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**	**WCBP94**	**WCBP04**	**SCBP93**	**SCBP05**
NC		0.000	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TX	0.000		0.000	0.215	0.011	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000
AZ	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
LA	0.000	0.011	0.000		0.601	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
GA	0.000	0.001	0.000	0.305		0.130	0.002	0.000	0.000	0.000	0.000	0.000	0.000
AL	0.000	0.000	0.000	0.000	0.002		0.007	0.000	0.000	0.000	0.000	0.000	0.000
MS	0.000	0.000	0.000	0.000	0.000	0.001		0.000	0.000	0.000	0.000	0.000	0.000
NM	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
SC	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000
WCBP 94	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000
WCBP 04	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000
SCBP 93	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
SCBP 05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	

Table 6.
Allele frequncies of honey bee samples collected from nine southern states where n = the number of samples from each state at each locus.

A24	**TX (n=80)**	**NC (n=34)**	**AZ (n=198)**	**LA (n=51)**	**GA (n=36)**	**AL (n=39)**	**MS (n=50)**	**NM (n=63)**	**SC (n=72)**
alleles									
90								0.086	
92			0.019		0.015				0.054
94	0.512	0.516	0.541	0.512	0.468	0.513	0.592	0.232	0.479
96	0.006		0.003					0.008	0.013
100	0.006		0.003						0.041
102	0.141	0.083	0.086	0.122	0.156	0.175	0.118	0.267	0.116
104	0.294	0.4	0.333	0.329	0.343	0.27	0.289	0.465	0.28
106	0.019		0.012	0.024	0.015	0.04		0.017	0.013
133	0.019								
140				0.012					

A28	**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
alleles									
101				0.026					
104				0.026					
112			0.003						
126				0.026	0.015				0.021
128	0.324	0.362	0.245	0.21	0.328	0.486	0.385	0.129	0.292
130	0.006								0.007
132	0.013	0.034	0.01	0.013	0.015	0.013		0.017	
135	0.642	0.603	0.711	0.684	0.609	0.486	0.614	0.853	0.678
137	0.006		0.016		0.015	0.013			
139	0.006		0.013	0.013	0.015				

A88	**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
alleles									
116			0.003						
130			0.006						

133		0.032							
136	0.037	0.112	0.045	0.013	0.015	0.027	0.052	0.094	0.077
139	0.407	0.338	0.254	0.381	0.453	0.486	0.473	0.206	0.338
141			0.029			0.013			0.014
146	0.037	0.032	0.025	0.039		0.04	0.026	0.025	0.028
148	0.487	0.451	0.606	0.526	0.531	0.378	0.434	0.655	0.514
150	0.012	0.032	0.019			0.054	0.013	0.008	0.028
152	0.018		0.009	0.039				0.008	

Ap66	**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
alleles									
87			0.025					0.02	
93	0.08	0.08	0.112	0.088	0.104	0.258	0.053	0.081	0.344
95	0.58	0.4	0.653	0.588	0.479	0.344	0.464	0.653	0.442
100	0.198		0.005	0.176	0.27	0.344	0.446	0.183	0.18
101	0.11	0.52	0.193	0.147	0.083	0.017	0.035	0.03	0.016
103	0.022		0.005		0.02			0.02	0.016
107								0.01	
113	0.007								
115					0.02				
117			0.005		0.02	0.017			
129						0.017			

B124	**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
alleles									
212		0.031	0.013		0.032	0.013		0.009	0.007
214	0.37	0.375	0.469	0.153	0.209	0.175	0.303	0.509	0.45
216	0.148	0.781	0.118	0.192	0.29	0.121	0.03	0.027	0.064
218	0.012	0.031	0.016	0.096		0.054		0.009	

220	0.228	0.156	0.125	0.269	0.225	0.283	0.378	0.209	0.164
222	0.074	0.062	0.047	0.057	0.08	0.04	0.09	0.118	0.078
224	0.098	0.14	0.054	0.115	0.096	0.229	0.045	0.036	0.157
226	0.012	0.031	0.01	0.038	0.048	0.013		0.036	0.035
228	0.018		0.003	0.019		0.013			0.014
230	0.012		0.013	0.019	0.016	0.027	0.06		0.014
232							0.015		
234	0.024	0.046	0.057	0.038		0.027	0.06	0.045	0.007
236		0.046	0.067				0.015		0.007
238			0.003						

A7	**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
alleles									
96								0.009	
102	0.047		0.041	0.016		0.025		0.009	
104	0.123	0.214	0.207	0.306	0.2	0.089	0.097	0.113	0.112
106	0.006							0.037	
108	0.349	0.428	0.293	0.306	0.4	0.487	0.413	0.264	0.411
110	0.02	0.017		0.016		0.038			
112	0.102	0.053	0.092	0.129	0.066	0.089	0.217	0.113	0.233
114	0.116	0.089	0.197	0.032	0.15	0.141	0.141	0.198	0.072
116	0.143	0.142	0.092	0.193	0.083	0.089	0.097	0.16	0.121
118	0.006		0.003		0.033	0.012			0.008
120			0.012					0.009	0.008
124			0.003						
126	0.013		0.003				0.01		
128			0.015			0.012		0.009	
130	0.006		0.015						0.024

132	0.013		0.006		0.016		0.01		
134		0.017	0.003						
136								0.009	
138	0.006	0.017			0.033			0.009	
140	0.034	0.017	0.003				0.01	0.037	0.008
142	0.006		0.003						
153			0.003						
157			0.003		0.016			0.018	
171							0.012		

Ap113 *alleles*	**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
198				0.013					
202	0.157	0.129	0.071	0.138	0.112	0.184	0.218	0.071	0.093
204	0.006								
206				0.013			0.01		
210						0.026			
212			0.005						
214	0.198	0.145	0.1	0.111	0.096	0.078	0.125	0.196	0.195
216			0.02					0.008	
218			0.005		0.032				
220	0.513	0.58	0.688	0.611	0.645	0.526	0.531	0.616	0.671
222	0.027	0.064	0.02	0.027	0.032	0.105	0.041	0.026	0.007
224	0.02		0.04	0.055			0.02		0.015
226	0.006	0.032	0.028	0.013	0.032	0.026	0.01	0.035	
228	0.027	0.016	0.017		0.016	0.013	0.031	0.035	0.015
233	0.013			0.013	0.016	0.013		0.008	
235	0.006	0.032					0.01		
236	0.006								
237	0.013		0.002		0.016	0.026			

Ap43		**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
alleles										
	135	0.042	0.117	0.127		0.047	0.08	0.039	0.031	0.096
	137	0.074	0.205	0.158	0.4	0.381	0.24	0.368	0.212	0.129
	139	0.01	0.029							0.048
	143					0.071				
	145	0.297	0.147	0.147	0.366	0.214	0.06	0.105	0.127	0.435
	147	0.521	0.352	0.519	0.233	0.238	0.54	0.394	0.5	0.209
	149	0.01	0.147	0.031		0.023	0.04	0.026	0.063	0.032
	151					0.023		0.013	0.021	0.016
	167	0.01		0.003				0.013		
	171			0.003					0.021	
	175									0.016
	177							0.013		
	181							0.013		
	183								0.01	
	187	0.01		0.003				0.013		0.016
	189	0.021		0.003					0.01	
	192						0.02			

Ap55		**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
alleles										
	163			0.003						
	165			0.003						
	171	0.007								0.122
	173	0.088	0.023	0.061	0.023	0.038	0.042	0.058	0.072	0.112
	175	0.411	0.5	0.391	0.547	0.384	0.257	0.337	0.406	0.306
	177	0.066	0.095	0.057	0.047	0.038	0.114	0.069	0.114	0.081
	179	0.066	0.071	0.258	0.047	0.096	0.028	0.093	0.104	0.03
	181	0.257	0.19	0.078	0.214	0.346	0.428	0.383	0.187	0.244
	183	0.022	0.071	0.091	0.047	0.038	0.028	0.011	0.031	0.02
	185	0.029		0.01	0.023	0.019	0.014	0.011	0.031	0.071
	187	0.029	0.047	0.03	0.023		0.057	0.034	0.041	0.01
	189			0.01						

	191				0.023				
	193							0.01	
	195	0.014		0.003			0.014		
	199					0.019			
	202	0.007				0.019	0.014		

A81		**TX**	**NC**	**AZ**	**LA**	**GA**	**AL**	**MS**	**NM**	**SC**
alleles										
	108	0.007		0.054	0.023					
	114			0.003			0.027	0.032		
	127	0.388	0.44	0.446	0.452	0.553	0.319	0.391	0.278	0.293
	135	0.597	0.52	0.42	0.523	0.41	0.583	0.554	0.634	0.672
	137		0.02	0.05		0.035	0.069		0.019	0.025
	139		0.02	0.018				0.021	0.057	0.008
	141	0.007		0.007					0.009	

Figure 1.
Percentage of 9 haplotypes found in the feral population collected between 1980 and 1992.

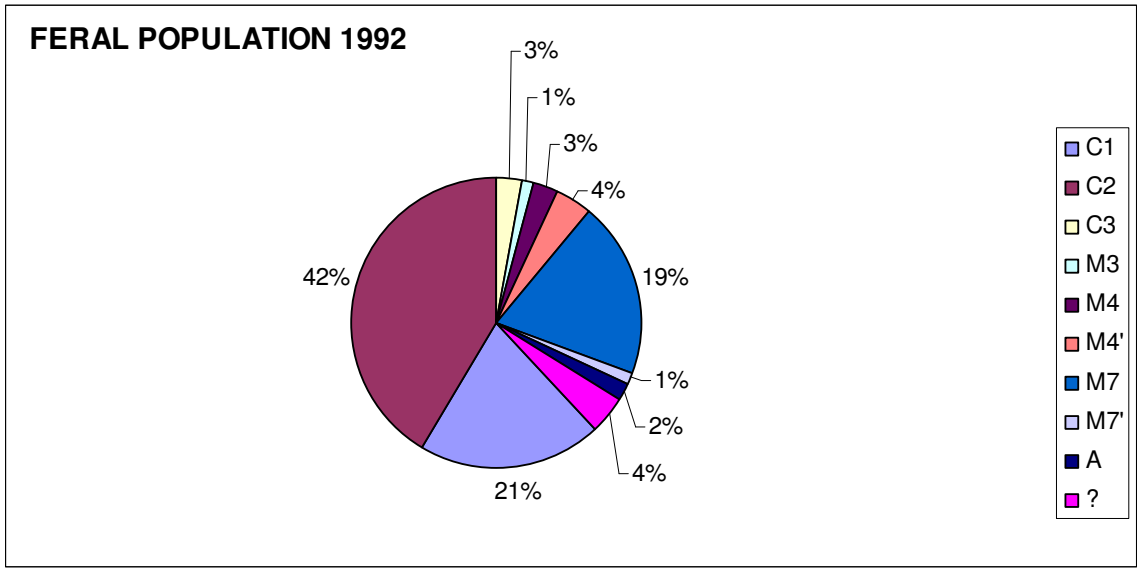


Figure 2. Sequneces of 13 feral samples with the M7 haplotype and the two M7 variants (*a* and *b*) reported in Franck et al. (2000).

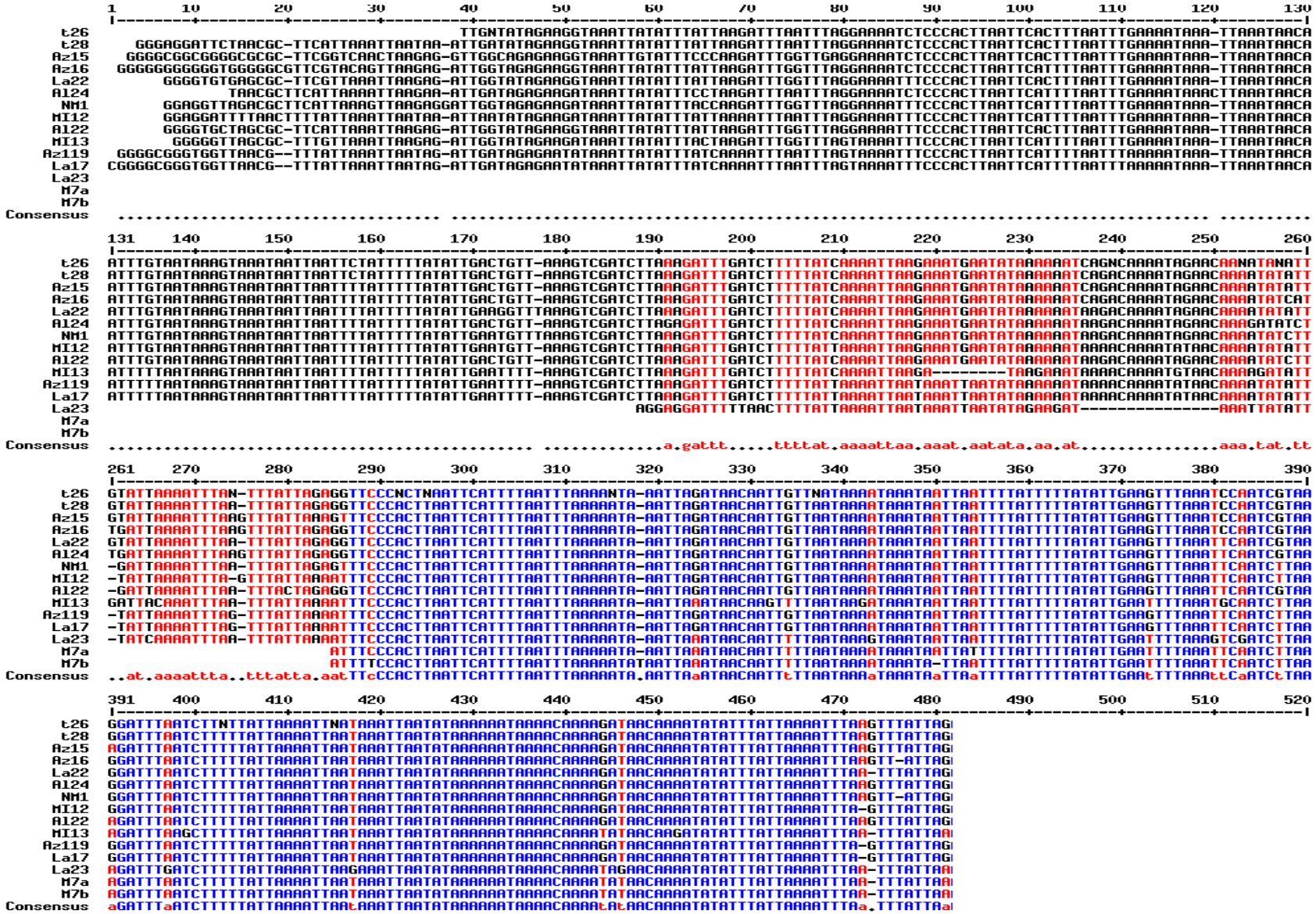


Figure 3. Distributions of the COI-COII mitochondrial haplotypes in honey bee collected from feral nests from 9 southern states. Pie charts represent the frequencies of nine haplotype patterns (C1, C2, C3, M3, M4, M4', M7, M7' and A).