Mitochondrial DNA Diversity, Parasite and Pathogen Occurrence, and a Potential Disease Vector in Managed and Unmanaged Honey Bee, Apis mellifera L. Populations

Dylan Cleary
University of Arkansas, Fayetteville

Follow this and additional works at: https://scholarworks.uark.edu/etd

Citation

This Dissertation is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact scholar@uark.edu.
Abstract

The western honey bee, *Apis mellifera* L., is a globally important pollinator plagued by several harmful stressors impacting colony health and survival. At least eight *A. mellifera* subspecies were imported and continue to be the genetic ancestors of U.S. honey bee populations today. Successive genetic bottle-neck events have led to reduced genetic diversity in U.S. honey bees. First, the subset of subspecies imported into the U.S. represents only a third of *A. mellifera* subspecies. Next, the parasitic *varroa* mite reduced managed and feral populations. Third, ongoing breeding practices have selected for traits from a single genetic lineage and bred from a limited stock of queens. Due to these genetic bottle-neck events, and limited access to outside germplasm, interest has arisen concerning increasing genetic diversity in U.S. honey bee populations.

Analysis of the mitochondrial DNA (mtDNA) in honey bees has allowed for genetic studies to characterize the origin of honey bee populations. Research has focused on commercial colonies, finding a limited pool of haplotypes from the Eastern European honey bee lineage. Research concerning the genetic diversity in non-commercial honey bee populations has been largely neglected. We studied the genetic variation of managed and unmanaged honey bee populations in the U.S., sequencing a portion of the mtDNA cytochrome oxidase I and II intergenic region. Molecular diagnostics were utilized to detect for seven pathogen species. Additionally, molecular and taxonomic techniques were used to explore a common colony pest’s identity and vector potential.

Haplotype diversity occurred among regions and between management types: a total of 43 haplotypes within four genetic lineages, ‘C’ (67.4%), ‘A’ (16.3%), ‘M’ (12.9%), and ‘O’ (3.4%); were detected throughout our study. We detected 20 haplotypes in unmanaged Utah
honey bee populations, the majority of which (48%) were from the A lineage, suggesting African
descent honey bees are common and diverse in Utah. In Arkansas, 25 haplotypes were detected
in hobbyist-kept and unmanaged honey bee populations. Six commonly detected C lineage
haplotypes accounted for 88% of the samples; however, 17 haplotypes from four lineages were
detected in the remaining samples. Together, these findings suggest that hundred-year-old
genetic remnants of historical importations have survived in feral honey bee populations despite
the arrival of harmful stressors.

This study is among the first to comprehensively explore genetic origin, management,
and regionality as factors of pathogen infection in U.S. honey bee populations. Honey bee
pathogens *Nosema* sp., *Lotmaria passim*, and *Varroa destructor* were detected in all lineages,
regions, and management types. All three pathogens were least prevalent in the A lineage,
potentially due to African honey bees’ hygienic behaviors. Additionally, managed colonies
exhibited higher mite loads compared to unmanaged colonies.

This study identified cockroaches sampled from Arkansas honey bee colonies as
*Parcoblatta* sp. Additionally, Hymenoptera DNA was detected within the guts of the samples,
indicating they likely fed on deceased bees. *Nosema ceranae* and *L. passim* were not detected in
our samples, signaling further studies are necessary to understand the wood roach’s potential role
in spreading honey bee pathogens.
Acknowledgements

I would like to first acknowledge my major advisor, Allen Szalanski, and thank him for mentoring me for the last seven years. I am overwhelmed by the amount of experience and growth I have been able to acquire while working in the University of Arkansas Insect Genetics lab. Dr. Szalanski wrote countless letters of recommendation, edited hundreds of pages of my drafts, and offered invaluable advice and resources to me to assist in making the graduate student experience rewarding and the degree achievable. I must thank him for inviting me into the Insect Genetics lab, and letting it be a second home over the last seven years. I do not know how I would have fared without his guidance and mentorship.

I would also like to thank the members of my committee, Don Steinkraus, Jackie Lee, and Neel Joshi for their support and guidance. My understanding of insect morphology, pathology, and honey bee biology and history would be severely lacking without Dr. Steinkraus. I appreciate his enthusiasm for the world of entomology as well as his appreciation for the world outside entomology; his love of music, art, and history is inspiring. Dr. Lee has given me the opportunity to work outside the laboratory, whether through fieldwork or presentations, I am grateful for the experiences she has facilitated. Dr. Joshi taught me more in one semester about insect toxicology than I had grasped in all my years prior. Thank you for offering career advice and talking gardening. My committee has provided valuable feedback and asked important questions throughout this process.

Next, I must thank my lab mates, Clint Trammel and Mary-Kate Williams. I am convinced that Clint could fix any lab issue that could possibly arise. He taught me the inner workings of graduate school and how to troubleshoot in the lab. I am sure I asked Clint hundreds of dumb questions and he answered every single one without hesitation and without judgement.
Mary-Kate Williams has an inspiring work ethic and was always willing to help and bounce ideas off. She has provided me with constant support, guidance, and friendship.

My entomology cohort has supplied me with endless amounts of knowledge, advice, and entertainment during my time at the University of Arkansas. I must express my gratitude to Hillary Fischer for serving as a gold standard of a graduate student, researcher, and human being. I would also like to thank Austin Jones for always offering the most real advice and showing me that there is always a light at the end of the tunnel. There are numerous other students that have helped me during my time at the University of Arkansas and I owe them so much more than a line in a dissertation. Thank you.

Last, I must thank my family. Paul and Julie Cleary have supported me in every aspect of life. Their love and guidance are unconditional, and they have always set the standard to achieve while prioritizing personal well-being. My grandparents, Howard and Virginia DelCour, while I wish you could be here to see this degree to its completion, I cannot thank you both enough for the strength you have provided me with. My siblings, Conor and Caitlin Cleary, are my winning lottery tickets in life. They have always provided me with a place to readjust, re-evaluate, and reset when I needed it. Lastly, I would like to thank Brian Threlkeld. You have loved and supported me through seven states, three degrees, and 11 years. You have fed, watered, and cared for me when I was unable to do so for myself. You have let hissing cockroaches stay in our home and you have assured me that vacations are in our future.
Dedication

This dissertation is dedicated to my family: Paul, Julie, Conor, Caitlin, Brian, Virginia, and Howard.
# Table of Contents

Chapter 1: Introduction ........................................................................................................... 1

Background .............................................................................................................................. 1
Honey Bees ............................................................................................................................... 1
History of Honey Bees in the United States ........................................................................ 3
Honey Bee Pollination and Economics.................................................................................. 4
Honey Bee Health, Colony Losses, and Stressors ................................................................. 6
Honey Bee Management ......................................................................................................... 8
Managed Honey Bees ............................................................................................................. 8
Swarms ..................................................................................................................................... 10
Feral Honey Bees .................................................................................................................... 11
Genetic Diversity and Bee Breeding....................................................................................... 12
Parasites, Pathogens, and Pests ............................................................................................. 14
*Varroa destructor* .................................................................................................................... 15
Microsporidian Pathogens ....................................................................................................... 16
Protists ...................................................................................................................................... 17
Bacterial Pathogens ................................................................................................................ 18
Molecular Techniques ............................................................................................................. 19
Objectives ............................................................................................................................... 20
References ............................................................................................................................... 22

Chapter 2: Genetic variation of feral honey bees (*Apis mellifera* L.) from Utah (USA) ...... 39

Abstract ................................................................................................................................... 39
Introduction .............................................................................................................................. 40
Materials and Methods .......................................................................................................... 42
Results ...................................................................................................................................... 43
Discussion ............................................................................................................................... 44
Acknowledgements ............................................................................................................... 47
References ............................................................................................................................... 48

Chapter 3: Mitochondrial DNA genetic variation in Arkansas honey bee, *Apis mellifera* L., colonies .................................................................................................................................. 56

Abstract ................................................................................................................................... 56
Introduction .............................................................................................................................. 57
Materials and Methods .......................................................................................................... 62
Results ...................................................................................................................................... 63
Discussion ............................................................................................................................... 64
Acknowledgements ............................................................................................................... 68
References ............................................................................................................................... 69
List of Published Papers

Chapter 2:

Chapter 1: Introduction

Background

The Western honey bee, *Apis mellifera* L. 1758, is the most widely utilized commercial pollinator worldwide, contributing over 15 billion USD to the U.S. economy in pollinator services (Calderone, 2012; Klein et al., 2007). Fluctuations in honey bee populations in the U.S. are longstanding but reached dramatic lows in 2008 at approximately 2.3 million managed colonies, compared to peak colony numbers in the 1940s at approximately 6 million, a 61% decline (Ellis, 2012; Ellis et al., 2010; USDA-NASS, 2022). Large-scale colony losses pose a sizable threat to global food production and have prompted concern amongst the public (Elobeid & Hart, 2007). Multiple stressors contribute to reduced honey bee health (Oldroyd, 2007; Runckel et al., 2011; Simone-Finstrom et al., 2016; vanEngelsdorp et al., 2009). Given the role of infectious microorganisms in reduced honey bee health, there is a need to understand factors that influence susceptibility and transmission.

Honey Bees

Honey bees are a subset of eusocial bees that compose the genus *Apis* (Order: Hymenoptera). As eusocial insects, honey bees exhibit cooperative brood care, overlapping generations, and reproductive division of labor (Wilson, 1975). The polyandrous reproductive queen utilizes a haplodiploid sex-determination system in which fertilized eggs produce diploid female worker bees, and unfertilized eggs produce haploid male drones (Beukeboom & Perrin, 2014; Wilson, 1975). Honey bees are characterized by their production and storing of honey, unique communication tactics, and elaborate comb nests constructed by workers using wax secreted from their abdominal glands (Ruttner, 1988; Seeley, 1985; Winston, 1991). There are 7-11 recognized species of honey bees within three subgenera, *Megapis* (giant honey bees),
Micrapis (dwarf honey bees), and Apis (cavity-nesting honey bees) (previously Sigmatapis) (Bustamante et al., 2021; Crane, 2009; Ruttner, 1988; Yadav et al., 2017). The subgenera differ in morphology, behavior, genetics, and distribution (Oldroyd & Wongsiri, 2009; Otis, 1996; Ruttner, 1988). While humans gather honey from all species of honey bees, only the cavity-nesting bees have been successfully managed (Crane, 2009). Worldwide, A. mellifera is the most widely kept pollinator due to its productivity as a pollinator and as a producer of wax and honey that humans can harvest without destroying the colony (Bakker, 1999; Crane, 1999; Delaplane & Mayer, 2000; Oldroyd & Nanork, 2009).

Apis mellifera is the only honey bee species with a native range beyond Asia. The proposed native range of A. mellifera includes parts of Africa, Asia, and Europe (Ruttner, 1988; Smith, 1991). Today, A. mellifera has a broad geographic distribution spanning every continent except Antarctica, primarily due to human-mediated migration and natural swarming (Crane, 1999; Ruttner, 1988; Seeley, 1985; Sheppard & Meixner, 2003).

Approximately 30 A. mellifera subspecies within six evolutionary lineages have been identified. These lineages include the A (African group), M (Northern and Western Europe), C (Southeastern Europe), O (Near East and Middle East) (Ruttner et al., 1978; Ruttner, 1988; Franck et al., 2001; Kandemir et al., 2006; Ferreira et al., 2008; Shaibi et al., 2009), Y (Ethiopia) (Franck et al., 2001), and Z (Syria and Lebanon) (Alburaki et al., 2013). Considerable variation exists in genetics, morphology, physiology, and behavior, even between closely related subspecies. This variation is associated with adaptations to different climates and ecological environments (Ruttner, 1988). For example, subspecies originating from northern climates, such as A. m. mellifera L. (M), are characterized by their large, broad, and densely “haired” abdomens. Subspecies originating from Africa, such as A. m. lamarckii Cockerell (A), are
characterized by their short wings, tongue, and legs, and small slender body size. Behaviorally, 
*A. m. ligustica* Spinola (C) is considered a docile and productive subspecies, while *A. m. syriaca* Skorikov (O) is also known to yield high volumes of honey; however, it is known for its highly

defensive nature and propensity to swarm (Ruttner, 1988). *Apis mellifera* subspecies were

traditionally differentiated using more than 24 morphological characteristics (Ruttner, 1988). Measurements of body parts and analysis of wing shape are among the morphological characteristics used to distinguish subspecies; however, hybridization may complicate morphometric analysis if variations occur in characteristics (DuPraw, 1965; Ruttner, 1988).

**History of Honey Bees in the United States**

*Apis mellifera* is not native to the U.S. and was first recorded in the U.S. in 1622, when European colonizers transported *A. mellifera* from England to Jamestown, Virginia (Oertel, 1976). Only *A. mellifera mellifera* L. 1758 (M), the European dark bee, was imported into the U.S. for over 200 years (Crane, 1999; Sheppard, 1989a, 1989b). The *A. m. mellifera* colonies quickly adapted to the eastern U.S. forests and began swarming, moving south and eastward, establishing feral colonies (Crane, 1999; Horn, 2005; Kritsky, 1991). By 1792 the first honey bee swarm reportedly crossed the Mississippi River in St. Louis (Crane, 1999; Kritsky, 1991).

Following the successful establishment of *A. m. mellifera* in the U.S., importation of other honey bee subspecies from different lineages began from 1859 to 1922. These subspecies included *A. mellifera ligustica* Spinola (C), *A. mellifera lamarckii* Cockerell (A), *A. mellifera carnica* Pollmann (C), *A. mellifera cypria* Pollmann (O), *A. mellifera syriaca* Skorikov (O), *A. mellifera caucasia* Pollman (C), and *A. mellifera intermissa* Buttel-Reepen (A) (Crane, 1999; Horn, 2005; Pellett, 1938; Sheppard, 1989a, 1989b). Additionally, while there is no official documentation of its importation, *Apis mellifera iberiensis* Engel (M & A) was likely imported
by Spanish settlers to Mexico in the 1500s (Brand, 1988; Seeley, 2019; Sheppard, 1989b). *Apis mellifera iberiensis* was also among the subspecies hybridized with *A. m. scutellata* in Brazil (Crane, 1999). Movement and importation of honey bee colonies increased from 1859-1914 due to: advances in transportation, i.e.: transcontinental railroads; migration of human settlers across the country; and honey bees’ drive to reproduce via swarming (Cobey et al., 2012; Crane, 1999).

Before the 1922 Honey Bee Act halted the importation of honey bees due to the spread of *Acarapis woodi* Rennie 1921, eight subspecies from four lineages (A, C, M, and O) were introduced to the U.S. (Sheppard, 1989a, 1989b). In 1990, *A. m. scutellata* hybrids were detected in the U.S. after entering from Brazil (Sugden & Williams, 1990; Winston, 2014). Thus, 9-10 subspecies have been documented as introduced into the U.S. (Sheppard, 1989a, 1989b).

*Apis m. ligustica*, *A. m. carnica*, and *A. m. caucasica* were maintained and widely favored by beekeepers due to their docile behavior and increased honey production leading beekeepers to dismiss the less productive and more aggressive subspecies, specifically those from the M, A, and O lineage (Carpenter & Harpur, 2021; Watkins, 1968). However, feral populations maintained these genetics as they dispersed across the U.S. (Cobey et al., 2012; Ruttner, 1988; Schiff et al., 1994). The Eastern European lineage honey bees continue to be favored by beekeepers commercially due to their gentleness, productivity, and less propensity to swarm or abscond (Sheppard, 1989a, 1989b). The honey bee subspecies commercially available in the U.S. are *A. m. carnica* and *A. m. ligustica* (Sheppard, 2012).

**Honey Bee Pollination and Economics**

Estimated to contribute greater than $215 billion (USD) worth of pollination services to crops globally and over $15 billion (USD) annually in the U.S., *A. mellifera* is the most economically important pollinator worldwide. Furthermore, *A. mellifera* pollination service
values are speculated to be far greater than previously estimated (Calderone, 2012; Jordan et al., 2021; Morse & Calderone, 2000). Additionally, honey bees are also important for honey production, valued at approximately $330 million (USD) annually in the U.S. (USDA-NASS, 2021a).

*Apis mellifera* is a successful commercial pollinator due to its versatility and efficiency (Alger et al., 2018; Gallai et al., 2009; Klein et al., 2007). Also, *A. mellifera* is polylectic and exhibits floral fidelity, a behavior in which bees only collect pollen and nectar from a single plant species on a foraging trip (Amaya-Márquez, 2009; Brosi & Briggs, 2013). This mutualistic short-term pollination specialization increases pollination and yields of hundreds of crops while increasing pollen concentration and efficient pollen collection (Aizen et al., 2008; Fünfhaus et al., 2018).

Honey bees have several advantages as reliable commercial pollinators beyond their diverse dietary preferences and foraging behaviors. Honey bees are eusocial, with 10,000-40,000 worker bees within a single colony, providing a substantial workforce of foraging worker adults (Seeley, 1985). The perennial nature of honey bees and their ability to be easily maintained in human-made boxes with removable frames, ensure pollinator numbers and easy transport to a particular area (Seeley, 1978). Additionally, beekeepers supplementing with artificial diets can accelerate colony numbers in preparation for pollination (vanEngelsdorp & Meixner, 2010).

The largest single pollination event in the U.S. is the almond pollination in California, where approximately 80% of the world’s almonds are produced (USDA-FAS, 2020). A total of 1.33 million almond-bearing acres were reported in California as of 2021, a 95.6% increase since 2008 (Bond et al., 2021; USDA-NASS, 2021c). In 2020, an estimated 2.4 million honey bee colonies were used to pollinate California almonds, of which 1.9 million colonies were shipped
to California (Goodrich & Goodhue, 2020; USDA-NASS, 2021b). Almond pollination, and other large-scale pollination events, rely on honey bee colonies to sustain a multibillion-dollar industry; however, honey bee colony health and populations are unstable. Notably, elevated winter colony losses (up to 40%) are a major concern for commercial beekeepers considering the almond bloom follows the winter months in February (Cavigli et al., 2016; Döke et al., 2015; Potts et al., 2010; vanEngelsheldorp et al., 2012).

**Honey Bee Health, Colony Losses, and Stressors**

Between 1987 and 2008, U.S. managed colonies declined by 33% (3.4 million to 2.3) (Daberkow et al., 2009; Ellis et al., 2010; vanEngelsheldorp & Meixner, 2010). As of 2021, the number of managed, honey-producing colonies is estimated at 2.7 million (Ellis, 2012; USDA-NASS, 2022). Declines in honey bee populations are not novel. Honey bees have an extensive history of fluctuating populations, documented as far back as the 1800s (Critchlow, 1904; Oldroyd, 2007; vanEngelsheldorp & Meixner, 2010). For example, in 1903, a phenomenon known as ‘disappearing disease’ was observed in Cache Valley, Utah, when 2,000 colonies were lost following a harsh winter and cool spring (Critchlow, 1904; Kulincevic et al., 1984; Root, 1990; Wilson & Menapace, 1979). Disappearing disease was described as a phenomenon in which unexplained losses of adult worker bees occurred suddenly, typically with no visible cause for the bees to abandon the colony (i.e., ample food and brood) (Kulincevic et al., 1984; Wilson & Menapace, 1979). Other similar phenomena have been named since (Finley et al., 1996; Oldroyd, 2007; Root, 1990).

Decades later, in 2006, a similar phenomenon known as Colony Collapse Disorder (CCD) was named and described as the sudden disappearance of worker honey bees from a colony. Colonies described as having CCD lacked dead bees in or around the colony and is
marked by the queen and perhaps a few recently developed workers remaining within the colony, typically with food reserves and capped brood (Cox-Foster et al., 2007; Oldroyd, 2007; Stokstad, 2007; vanEngelsdorp et al., 2009). Colony Collapse Disorder continues to be researched; however, it was heavily studied between 2010-2017, when roughly 276 CCD-related research articles were published in an eight-year period. To date, no definitive single scientific cause has been identified for the dramatic and rapid loss of adult worker honey bees described as CCD (Martin, 2015; Watson & Stallins, 2016). In more recent years, the number of honey bee colonies in the U.S. has increased compared to the losses of 2008 (USDA-NASS, 2022). However, the quantity of colonies is still disproportionate compared to the growth in pollinator-dependent crop acreage, which has increased by greater than 300% globally in recent years (Aizen & Harder, 2009; Calderone, 2012).

Honey bee colony losses have not been attributed to a single factor. Instead, several biotic and abiotic factors have been identified and implicated as harmfully impacting honey bee health. Furthermore, a growing body of research indicates the potential and realized negative consequences of these stressors, both singly and combined (Jack et al., 2016; Klein et al., 2007; Potts et al., 2010). Several scientific reviews have detailed the factors related to honey bee decline (Goulson et al., 2015; Oldroyd, 2007; Potts et al., 2010; vanEngelsdorp & Meixner, 2010).

Among, but not limited to, the factors implicated are improper pesticide usage and pesticide accumulation (Mullin et al., 2010; Pettis et al., 2013; Straub et al., 2016; Tison et al., 2016); habitat loss, fragmentation, and degradation (Foley, 2005; Vanbergen et al., 2013); dietary stress (Brodschneider & Crailsheim, 2010; Taric et al., 2019); climate change (Cornelissen et al., 2019; Willmer, 2012); genetic diversity (López-Uribe et al., 2017; Mattila &
Seeley, 2007; Tarpy, 2003); pests, parasites, and pathogens (Evans & Schwarz, 2011; Smith et al., 2013); migratory and commercial beekeeping (Alger et al., 2018; Simone-Finstrom et al., 2016); and colony management and treatment (Oldroyd, 2007; Raymann et al., 2017; Taric et al., 2019).

**Honey Bee Management**

As of March 2021, an estimated 2.7 million honey-producing managed honey bee colonies were reported in the U.S., with an estimated 32.3% winter loss and 45.5% total annual loss of managed colonies (Steinhauer et al., 2021; USDA-NASS, 2022). While declines are concerning, annual colony losses are best interpreted as a turnover rate rather than a change in populations because beekeepers can replace lost colonies throughout the year via splits, requeening, or swarm catching. Therefore, the high levels of losses do not necessarily result in a decrease in the total number of colonies managed in the U.S. (Steinhauer et al., 2021). Management practices and colony movement are implicated as major stressors to honey bee health (Alger et al., 2018). Honey bee colonies can occur in maintained colonies, temporary swarms, or feral, unmanaged colonies.

**Managed Honey Bees**

Managed colonies typically occur in man-made bee boxes with movable frames (which maintain the bee space of 9.5 mm) and are maintained by one of three types of beekeepers: Backyard (hobbyist), sideliners, or commercial (Lee et al., 2015; Royce & Rossignol, 1990; Schiff et al., 1994). Beekeeping operations can differ in the number of colonies managed, the intensity and type of management tactics used, and whether colonies remain stationary or are transported (Pilati & Prestamburgo, 2016). Commercial operations typically manage 500 or more colonies, sideliners manage 51-500 colonies, and hobbyists tend 50 or fewer colonies (Lee et al.,
Preliminary results from the 2020-2021 Bee Informed Partnership’s annual U.S. colony loss survey indicate higher winter losses in hobbyist colonies (42%) than commercial colonies (32.9%) but slightly higher summer losses in commercial (30.9%) compared to hobbyist (27%) (Steinhauer et al., 2021). Also, studies have detected elevated infection rates of *N. ceranae* and *L. passim* in commercial honey bee colonies compared to stationary colonies (Williams et al., 2019; Zhu et al., 2014).

Beekeepers can be further divided into migratory beekeepers, bee breeders, and those maintaining colonies for honey or wax production (Underwood et al., 2019). Migratory beekeeping involves transporting honey bee colonies to different locations for pollination services. Colonies are often transported thousands of kilometers during warm months in hives with poor ventilation, which can cause stress (Simone-Finstrom et al., 2016). Upon arrival at their destination, colonies are intermingled with other colonies from across the U.S., thus, exposing each other to new viruses, parasites, and pathogens (Klee et al., 2007; Simone-Finstrom et al., 2016). Additionally, honey bees may receive poor nutrition due to the lack of floral diversity in large-scale monocultures (Bacandritsos et al., 2010; Hendriksma & Shafir, 2016; Smith et al., 2013). The combination of stressed honey bees with poor nutrition exposed to harmful agents may result in compromised immune systems unable to fight off infectious agents adequately (Bacandritsos et al., 2010; Hendriksma & Shafir, 2016; Smith, 2012).

Commercial bee breeders are large-scale producers of honey bee queens and packaged bees sold to beekeepers regionally as well as across the country (Cobey et al., 2012). In terms of honey bee health, studies have detected *Nosema* sp. (Strange et al., 2008) and *L. passim* (Williams, 2018) in packaged bees and commercial queen breeder colonies, indicating these operations may facilitate the spread of harmful pathogens across the country. Last, Stationary
commercial beekeepers receive most of their profits from honey bee products, i.e., honey and wax (Underwood et al., 2019).

Maintained honey bee colonies benefit from beekeeper management in many ways. Combining weak colonies and supplementing colonies with pollen patties to increase stores may enhance overwintering success (Brodschneider & Crailsheim, 2010; Haydak, 1970). Insulation and ventilation provided by beekeepers, and in some cases transporting colonies to warmer climates, can aid in winter survival, particularly in harsh environments (Caron & Connor, 2013; Döke et al., 2015). Miticides are administered to control *Acarapis woodi* and *V. destructor*, two mite species known to have decimated feral and untreated populations (Rosenkranz et al., 2010). Furthermore, antibiotics are frequently used to fight off and prevent harmful microorganisms. However, long-term antibiotic usage in honey bee colonies has been shown to impact honey bee behavior negatively and is linked to reducing beneficial gut biota associated with pathogen resistance (Li et al., 2017; Ortiz-Alvarado et al., 2020; Raymann et al., 2017; Tian et al., 2012). Also, the close proximity of colonies in apiaries can increase transmission of pathogens between colonies due to shared floral resources, robbing events, or other interactions (Renz & Rosenkranz, 2001; Youngsteadt et al., 2015). Thus, while management of colonies can be beneficial, management tactics also have harmful implications, from reducing pest resistance, encouraging pathogen transmission, maintaining weak honey bees, to overexposure to antibiotics.

**Swarms**

Honey bee swarming is a process in which reproduction at the colony level naturally occurs, i.e., the original established colony (parental colony) splits into multiple. Colonies begin rearing new queens in response to 90% or more of the brood comb utilized (Winston et al.,
Roughly two-thirds of worker honey bees and the old mother queen depart from the parental colony and temporarily cluster as a bivouac; comb is not built during a swarming phase (Grozinger et al., 2014; Rangel & Seeley, 2012). Typically occurring in spring and early summer when resources are most plentiful, swarming is a response to prevent overcrowding in a colony. However, swarming may occur in colonies overcrowded or not (Fell et al., 1977; Winston, 1991).

The clustering typically occurs on a nearby bush, tree branches, or fence post near the parent colony (~90 m). Scout bees will then search for a new permanent location, reporting back to the swarm and recruiting other workers to evaluate the site until a suitable location is determined via a consensus. Honey bee swarms can fly several kilometers from their parent colony to establish, although, on average fly 300-600 meters from the parent colony (Lindauer, 1955; Seeley & Buhrman, 1999; Villa, 2004). The swarm then migrates and colonizes the new location (Seeley et al., 2006; Seeley & Visscher, 2003, 2004).

Swarm catching is a common practice in beekeeping in which beekeepers will capture swarms, providing them with a bee box and frames to start new colonies (Crane, 1999). While swarms are often splits from beekeepers’ existing colonies, there is potential for captured swarms to be splits from nearby feral colonies. Additionally, subspecies from the A, M, and O lineage have a stronger propensity to swarm than the C lineage subspecies (Ruttner, 1988). Little is known about pathogen occurrence in swarms.

**Feral Honey Bees**

Feral honey bee colonies are those which receive no active management by humans; when established in isolated areas, these colonies are left unmanaged for multiple generations (Schiff et al., 1994). Feral colonies typically occur in natural crevices and tree cavities; however,
they can also occur in manmade structures and materials such as buildings and trash (Szalanski et al., 2016). Following the numerous introductions of honey bees from 1622-1922, a large population of feral colonies became established across the U.S. Genetic studies demonstrated that feral populations displayed unique genetic markers consistent with historical importation stock and different from commercial queen breeder colonies, suggesting that feral populations could act as a genetic reserve for honey bee breeders (Schiff et al., 1994; Schiff & Sheppard, 1995, 1996).

Unfortunately, feral honey bee colonies were largely eliminated due to the introduction of *Varroa destructor* Anderson & Trueman (2000) in 1987 (Kraus & Page Jr, 1995; Loper, 1995; Seeley et al., 2015). The small surviving populations of feral colonies have been dubbed ‘survivor stock’ due to their ability to persist and adapt to various stressors, namely parasites and pathogens (Loper et al., 2006; McNeil, 2009, Seeley, 2007). Additionally, while feral honey bee populations crashed upon the arrival of the *varroa* mite, instances of persistence and resurgence of feral populations have been documented in the U.S. (Seeley, 2007; Villa et al., 2008), Europe (Le Conte et al., 2007), and South America (Rosenkranz, 1999). Feral colonies have also shown a lower disease burden and stronger immune response than managed colonies; however, the mechanism for this immunity is not well understood (López-Urbe et al., 2017; Youngsteadt et al., 2015). These feral honey bee populations may have survived without beekeeper manipulation for a long time and may have adapted to the different stressors that threaten their survivability (Loper et al., 2006; Seeley, 2007; Villa et al., 2008).

**Genetic Diversity and Bee Breeding**

Reduced genetic diversity has been linked with increased susceptibility to disease-causing agents in honey bees (Oxley & Oldroyd, 2010). Currently, research indicates limited
genetic variability in U.S. commercial honey bee breeding stock (Delaney et al., 2009; Magnus et al., 2011; Sheppard, 1989b). According to mitochondrial DNA (mtDNA) genetic data, seven of the nine known honey bee subspecies introduced into the U.S. have detected haplotypes (Carpenter & Harpur, 2021). It has been speculated that the diversity in climate and geography in the U.S. allowed for success amongst the different subspecies and their feral populations (Ruttner, 1988).

Managed honey bee colonies are primarily purchased as packages, nucs, or queens from honey bee breeders in the U.S. Studies on the genetic variation of the mtDNA COI-COII region of queen breeder colonies have revealed that the vast majority of queens have only the C lineage (Delaney et al., 2009; Magnus et al., 2011). As of 2009, 500 breeder queens were used to produce an estimated 900,000 daughter queens for commercial sale in the U.S., suggesting a limited pool of genetic stock (Delaney et al., 2009). A 2010 multistate study on commercial queen breeder colonies detected seven haplotypes exclusively within the C lineage (Magnus et al., 2011).

Honey bee populations were substantially affected by the accidental introduction of the varroa mite in 1987 (Wenner & Bushing, 1996). Notably, over 85% of feral colonies were decimated in parts of California (Kraus & Page Jr, 1995) and over 96% in parts of Arizona (Loper et al., 2006). Before their decimation, studies showed that feral honey bee populations in the U.S. differed distinctly from managed populations (Schiff et al., 1994; Schiff & Sheppard, 1993). Selective breeding and treatment of commercial honey bees, and large-scale loss of feral populations, may have hindered natural selection, reduced disease tolerance, and hampered local and seasonal adaptations to their environments resulting in reduced colony survivorship (Hatjina et al., 2014; Seeley et al., 2015).
Studies on mtDNA diversity of feral honey bee colonies and swarms from the continental U.S. (Magnus et al., 2014; Magnus & Szalanski, 2009) and Hawaii (Szalanski et al., 2016) have revealed several M and O lineage haplotypes not observed in commercial queen breeder colonies (Delaney et al., 2009; Magnus, 2011). Feral honey bee genetics are of interest to diversify commercial honey bee stock and improve colony health; however, there is concern over introducing aggressive Africanized traits (Mattila & Seeley, 2007). Additionally, Africanized honey bees can outcompete native pollinator species in southern Utah, resulting in the local extinction of the Andrenid bee *Perdita meconis* Griswold (Portman et al., 2017).

**Parasites, Pathogens, and Pests**

Harmful microorganisms are among the direct factors affecting honey bee health; the increasing threat they pose to honey bee pollination services has caused alarm globally (Budge et al., 2015; Cavigli et al., 2016; Youngsteadt et al., 2015). Pathogens include bacteria, fungi, protozoa, and viruses; harmful parasites and pests also influence bee health (Bailey & Ball, 1991; Oldroyd, 2007). Harmful microorganisms may cause physical, physiological, and behavioral changes to honey bees (Gómez-Moracho et al., 2017). Known parasites and pathogens such as the parasitic mite *V. destructor* and microsporidian pathogens, *Nosema apis* Zander (1909) and *N. ceranae* Fries (1996), are among the identified threats to honey bee health (Botías et al., 2013; Oldroyd, 2007; Rosenkranz et al., 2010; Uroš et al., 2014). However, newly identified and lesser-studied parasites and pathogens, such as trypanosomes and spiroplasmas are emerging as underlying threats to bee health (Fünfhaus et al., 2018; Williams et al., 2019).

Furthermore, there are pests within honey bee colonies whose presence has been largely unexplored, as we see with cockroaches. Numerous harmful pests and pathogens negatively impact honey bee health (Bailey & Ball, 1991; Evans & Schwarz, 2011). Here, we provide a
brief overview of a few harmful species currently known to substantially affect honey bee health and emerging species of interest related to honey bee health.

**Varroa destructor**

*Varroa destructor* is the single most harmful threat to honey bees worldwide (Locke, 2015; Rosenkranz et al., 2010; Uroš et al., 2014). Originally a parasite of *A. cerana*, the mite shifted hosts to *A. mellifera*; thus, there is a lack of co-evolution between *A. mellifera* and the mite (Anderson & Trueman, 2000). Often considered ubiquitous in honey bee colonies, *V. destructor* is an obligate parasite that feeds on honey bee adults, pupae, and larvae (Rosenkranz et al., 2010).

*Varroa destructor* are particularly damaging as they cause harm in multiple ways to honey bees. *Varroa destructor* was previously thought to feed on honey bee hemolymph exclusively. However, scientists recently found the mites feed on the honey bee’s fat body tissue, which is responsible for an array of crucial functions, including winter food and nutrient storage, detoxification of pesticides, and assisting in the management of the honey bee’s immune system among others. With a compromised fat body, winter mortality is among the potential risks (Arrese & Soulages, 2010; Ramsey et al., 2019). *Varroa destructor* has been linked to improper development of adult bees, reduced immune system function, and increased susceptibility to infection. In addition, *V. destructor* can vector multiple viruses (Bailey & Ball, 1991; Chen et al., 2004; Di Prisco et al., 2011; Posada-Florez et al., 2019; Sammataro et al., 2000) and pathogens to honey bees (Kanbar & Engels, 2003; Meixner et al., 2014; Rosenkranz et al., 2010). *Nosema ceranae* has been identified in both mite and *V. destructor* hemolymph samples (Uroš et al., 2014). Also, DNA from the trypanosome pathogen *Lotmaria passim* Schwarz was recently
detected in *V. destructor* found on honey bees positive for *L. passim* in Argentina, suggesting passive transport of the pathogen (Quintana et al., 2021).

**Microsporidian Pathogens**

Two species of unicellular microsporidian pathogens threaten *A. mellifera*: *N. apis* and *N. ceranae*. The pathogens develop and germinate in the honey bee’s gut and spread via spores during trophallaxis and grooming (Smith, 2012). Once in the gut, *Nosema* sp. spores rapidly multiply, often resulting in digestive disorders, but are also linked to decreased colony productivity, increased winter mortality, and reduced overall colony health (Botías et al., 2013; Rangel et al., 2016). *Nosema apis* was discovered in 1909, while *N. ceranae* was not identified until 1996. Furthermore, while *N. apis* evolved with the European honey bee as its host, *N. ceranae*’s original host was *A. cerana* until its introduction to *A. mellifera* (Botías et al., 2013; Fries, 2010). It is speculated that due to the lack of co-evolution between *N. ceranae* and *A. mellifera*, *N. ceranae* is more pathogenic and detrimental than *N. apis* to *A. mellifera* (Fries, 2010).

Several studies have examined *Nosema* sp. in managed honey bee colonies (Rangel et al., 2020; Szalanski et al., 2013; Zhu et al., 2014). Szalanski et al. (2013) found *Nosema* sp. in 44% of the apiaries sampled in New York and 29% of the apiaries sampled in South Dakota, with 96.8% (NY) and 100% (SD) being *N. ceranae*.

A 2016 study performed by Rangel et al. (2016) examined feral honey bee colonies in southern Texas and found that *N. apis* was only found in samples collected from 1991 to 1995. In contrast, *N. ceranae* was detected in feral colonies every year from 1991 to 2013, with the highest level of infection occurring in 2013 (85.7%). Still, fewer than 6% of the total feral samples were infected with *Nosema* sp. (Rangel et al., 2016). Similarly, Szalanski et al. (2014)
explored *Nosema* sp. in feral Africanized colonies from six U.S. states, finding relatively low levels of *Nosema* sp. (8.3%), 82.1% of which were *N. ceranae*. Overall, the above studies suggest that feral colonies have relatively low occurrences of *Nosema* compared to managed colonies. Furthermore, there is evidence that *N. ceranae* is the predominant species in U.S. managed and feral colonies (Chen et al., 2008; Fries, 2010; Smith, 2012).

**Protists**

Trypanosomatids are protozoan parasites known to infect numerous organisms and have been widely detected in the U.S. (Cox-Foster et al., 2007; Runckel et al., 2011; Schwarz et al., 2015; vanEngelsdorp et al., 2009). *Crithidia mellificae* Langridge (1967) and McGhee and *Lotmaria passim* Schwarz (2014) are two obligate pathogenic trypanosome species found in honey bees. While the impact of the trypanosome species is not well known, *Crithidia bombi* Lipa & Triggiani is a known and harmful pathogen of bumble bees. *Crithidia bombi* is known to cause loss of mass in overwintering bumble bee queens and reduced overall queen fitness and productivity (Brown et al., 2003; Yourth et al., 2008). *Apis mellifera* has been speculated to play a role in the transmission of *C. bombi* to bumble bees (Ruiz-González & Brown, 2006).

*Crithidia mellificae* was first described in 1967 in Australian colonies (Langridge & McGhee, 1967). The parasite is known to infect the rectum of the honey bee; however, little is known about its pathogenicity. *Lotmaria passim* was described more recently in 2015 following a molecular study in which the pathogen was differentiated from *C. mellificae* (Ravoet et al., 2015; Schwarz et al., 2015). *Lotmaria passim* is currently considered the more dominant species in the U.S., Belgium, Switzerland, Japan, Uruguay, Chile, and Argentina (Castelli et al., 2019; Morimoto et al., 2013; Ravoet et al., 2015). Castelli et al. (2019) found that honey bee colonies in Uruguay, Chile, and Argentina infected with *L. passim* displayed high *V. destructor*
parasitization levels, suggesting *L. passim* infection may increase bees’ susceptibility to *V. destructor* (Castelli et al., 2019). Additionally, Williams et al. (2017) found managed honey bee colonies had a higher occurrence (16%) of *L. passim* than feral colonies (4%).

**Bacterial Pathogens**

*Spiroplasmas* are small, helical, wall-less bacteria belonging to the Mollicutes class (Bailey & Ball, 1991). Intracellular organisms with associations with eukaryote hosts, consisting of plants, crustaceans, ticks, and insects, several species of *spiroplasma* have detrimental impacts on their hosts (Regassa & Gasparich, 2006; Tozkar et al., 2015). Two *spiroplasma* species are known to occur in honey bees, *Spiroplasma apis* Mouches (1984) and *S. melliferum* Clark (1985) (Clark, 1977; Clark et al., 1985; Mouches et al., 1983).

*Spiroplasma* sp. has been implicated in causing a seasonally occurring neurological disease known as “spiroplasmosis” or “May disease.” Symptoms of spiroplasmosis in honey bees include the inability to digest pollen, difficulty flying, and colony abandonment (Evans & Schwarz, 2011; Mouches et al., 1982; Schwarz et al., 2014). Known to invade the gut lumen, *S. apis* and *S. melliferum* also invade the hemolymph of honey bees, where parthenogenesis may occur (Bailey & Ball, 1991; Mouches et al., 1982). Further, higher mortality rates have been observed in honey bees, which carry spiroplasma (Clark, 1977, 1982; Clark et al., 1985). Mouches et al. (1982) infected honey bees with *Spiroplasma* via injection into the hemolymph. These bees died within five days unless given tetracycline.

*Spiroplasma melliferum* was first reported in Beltsville, MD in 1977 (Clark 1977, 1982, Clark et al. 1985). By 1980, colonies displaying symptoms of spiroplasmosis were observed in France, where *S. apis* was detected in large quantities (Mouches et al. 1982, 1983). *Spiroplasma* sp. infection is seasonally dependent, occurring most commonly in the spring months when
flowers bloom. *Spiroplasma* sp. has been isolated from flowers and honey bees; it is widely accepted that the pathogen uses flower surfaces as transmission sites (Clark, 1982; Schwarz et al., 2014).

**Wood Roaches**

Several insects occur within honey bee hives that act as occasional pests, one of which is the wood roach pests (Bailey & Ball, 1991). Ectobiidae is an insect family in the order Blattodea composed of several cockroach species, approximately 40 of which occur in the U.S. (Beccaloni, 2019). Within Ectobiidae are the wood cockroaches, mostly found within the genus *Parcoblatta*. Twelve species of *Parcoblatta* have been identified in North America, with most species occurring in eastern locations. These species include *Parcoblatta Americana* Scudder, *P. bolliana* Saussure & Zehnter, *P. caudelli* Hebard, *P. desertae* Rehn & Hebard, *P. divsa* Saussure & Zehntner, *P. fulvescens* Saussure & Zehntner, *P. lata* Brunner, *P. notha* Rehn & Hebard, *P. pensylvanica* DeGeer, *P. uhleriana* Sassure, *P. virginica* Brunner, and *P. zebra* Hebard (Atkinson et al., 1990; Beccaloni, 2019; Hebard, 1917; Pratt, 1988). As scavengers that have been found to feed on human feces, sewage, and garbage, cockroaches are known indirect mechanical vectors and disease reservoirs of numerous pathogens, including bacteria, protozoa, and viruses (Baumholtz et al., 1997). While roaches are known to occur in honey bee colonies, the species are currently unknown.

**Molecular Techniques**

Due to polymorphic variability, taxonomic and morphological identification and detection of understudied internal pathogens can be tedious and unreliable. Molecular diagnostics, including PCR, multiplex PCR, PCR-RFLP, and Q-PCR, allow for reliable and
robust identification of pathogens at the species level (Arismendi et al., 2016; Klee et al., 2007; Meeus et al., 2012; Szalanski et al., 2016).

DNA sequence analysis of a portion of the mtDNA cytochrome oxidase I and II (COI-II) intergenic region can provide information on the mitochondrial lineage of the ancestral queen of a colony (Franck et al., 2001; Meixner et al., 2013). The COI-COII intergenic region of *A. mellifera* exhibits a high level of genetic variability within and among *A. mellifera* lineages, making it useful for distinguishing lineages and identifying unique haplotypes (Cornuet et al., 1991; Garnery et al., 1992). Furthermore, this region’s sequencing data is consistent with previous morphological and biogeographical lineage distinctions (Arias & Sheppard, 1996; Cornuet et al., 1991; Garnery et al., 1992; Ruttner, 1988). For example, Africanized honey bees in the U.S. are virtually indistinguishable in the field from European honey bees and require morphometric analysis for morphological identification; however, molecular analysis using mtDNA can reliably characterize a colonies matriline (Rinderer et al., 1993; Sheppard and Smith, 2000). Introggression of Africanized genes using a mtDNA marker is, however, not detectable if a European queen has mated with Africanized drones.

Unlike nuclear genotypes, which can be altered during segregation or recombination during reproduction, mtDNA markers are maternally inherited, thus allowing mtDNA analysis to focus on the genealogies of individual lineages (Lansman et al., 1981). This also allows the use of a single individual to genetically characterize a honey bee colony (Sheppard and Smith, 2000).

**Objectives**

The objectives of this dissertation are:

1. Characterize the genetic diversity of honey bees from feral colonies and swarms in Utah using DNA sequence data of the COI-COII mtDNA region of *A. mellifera*
2. Characterize the mitochondrial genetic variation in Arkansas managed and unmanaged honey bee colonies and swarms

3. Identify mtDNA genetic haplotypes and explore pathogen occurrence (*N. apis, N. ceranae, C. mellificae, L. passim, S. apis, S. melliferum, and V. destructor*) associations between genetic lineages, management sources, and regional locations using novel and previously collected data from the Insect Genetics Laboratory

4. Identify cockroaches found in Arkansas honey bee colonies to the species level and determine if the cockroaches are feeding on dead honey bees or carrying honey bee pathogens (*N. ceranae* and *L. passim*)
References


Beccaloni, G. (2019). *Cockroach species file (8.13)* [Computer software]. Species File Group. [https://doi.org/10.15468/A7UXKL](https://doi.org/10.15468/A7UXKL)


Döke, M. A., Frazier, M., & Grozinger, C. M. (2015). Overwintering honey bees: Biology and management. *Current Opinion in Insect Science, 10*, 185–193. [https://doi.org/10.1016/j.cois.2015.05.014](https://doi.org/10.1016/j.cois.2015.05.014)


Le Conte, Y., de Vaublanc, G., Crauser, D., Jeanne, F., Rousselle, J.-C., & Bécard, J.-M. (2007). Honey bee colonies that have survived Varroa destructor. *Apidologie, 38*(6), 566–572. [https://doi.org/10.1051/apido:2007040](https://doi.org/10.1051/apido:2007040)


Occurrence of parasites and pathogens in honey bee colonies used in a European genotype-environment interactions experiment. *Journal of Apicultural Research, 53*(2), 215–229. [https://doi.org/10.3896/IBRA.1.53.2.04](https://doi.org/10.3896/IBRA.1.53.2.04)


Oldroyd, B. P. (2007). What’s killing American honey bees? *PLOS Biology, 5*(6), e168. [https://doi.org/10.1371/journal.pbio.0050168](https://doi.org/10.1371/journal.pbio.0050168)


Posada-Florez, F., Childers, A. K., Heerman, M. C., Egekwu, N. I., Cook, S. C., Chen, Y., Evans, J. D., & Ryabov, E. V. (2019). Deformed wing virus type A, a major honey bee pathogen, is vectored by the mite *Varroa destructor* in a non-propagative manner. *Scientific Reports*, 9(1), 12445. [https://doi.org/10.1038/s41598-019-47447-3](https://doi.org/10.1038/s41598-019-47447-3)


Renz, M., & Rosenkranz, P. (2001). Infestation dynamics and reinvasion of Varroa destructor mites in honey bee colonies kept isolated and in groups. *Apidologie (France)*.


Schwarz, R. S., Teixeira, É. W., Tauber, J. P., Birke, J. M., Martins, M. F., Fonseca, I., & Evans, J. D. (2014). Honey bee colonies act as reservoirs for two *Spiroplasma* facultative


Simone-Finstrom, M., Li-Byarlay, H., Huang, M. H., Strand, M. K., Rueppell, O., & Tarpy, D. R. (2016). Migratory management and environmental conditions affect lifespan and oxidative stress in honey bees. *Scientific Reports, 6*(1). https://doi.org/10.1038/srep32023


USDA-NASS. (2021a). *Honey* (Honey). National Agricultural Statistics Service (NASS), Agricultural Statistics Board (ASB), United States Department of Agriculture (USDA). [https://downloads.usda.library.cornell.edu/usda-esmis/files/hd76s004z/7h14bh90x/w9505v43v/hony0321.pdf](https://downloads.usda.library.cornell.edu/usda-esmis/files/hd76s004z/7h14bh90x/w9505v43v/hony0321.pdf)


vanEngelsdorp, D., & Meixner, M. D. (2010). A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. *Journal of Invertebrate Pathology, 103*, S80–S95. [https://doi.org/10.1016/j.jip.2009.06.011](https://doi.org/10.1016/j.jip.2009.06.011)


Chapter 2: Genetic variation of feral honey bees (*Apis mellifera* L.) from Utah (USA)

Dylan Cleary, Allen L. Szalanski, Clinton Trammel, Mary-Kate Williams, Amber Tripodi and Danielle Downey

**Abstract**

A study was conducted on the mitochondrial DNA genetic diversity of feral colonies and swarms of *Apis mellifera* from ten counties in Utah by sequencing the intergenic region of the cytochrome oxidase (COI-COII) gene region. A total of 20 haplotypes were found from 174 honey bee colony samples collected from 2008 to 2017. Samples belonged to the A (African) (48%); C (Eastern Europe) (43%); M (Western Europe) (4%); and O (Oriental) lineages (5%). Ten African A lineage haplotypes were observed with two unique to Utah among A lineage haplotypes recorded in the US. Haplotypes belonging to the A lineage were observed from six Utah counties located in the southern portion of the State, from elevations as high as 1582 m. All five C lineage haplotypes that were found have been observed from queen breeders in the US. Three haplotypes of the M lineage (n=7) and three of the O lineage (n=9) were also observed. This study provides evidence that honey bees of African descent are both common and diverse in wild populations of honey bees in southern Utah. The high levels of genetic diversity of A lineage honey bee colonies in Utah provide evidence that the lineage may have been established in Utah before the introduction of A lineage honey bees from Brazil to Texas in 1990.
Introduction

Colonies of honey bees, *Apis mellifera* L., were first brought to Utah, USA in covered wagons in 1848, and by 1872 there were approximately 2000 honey bee colonies in Utah (Nye, 1976). The Africanized honey bee (AHB) was first detected in Texas, USA in 1990 (Sugden & Williams, 1990), and by 2008 it was discovered in southern Utah (Hodgeson et al., 2010). By 2010 AHB had spread to three counties in Utah (Szalanski & Magnus, 2010). The Africanized honey bee in the United States is virtually indistinguishable in the field from the European honey bee (EHB) and requires a morphometric analysis for morphological identification (Rinderer et al., 1993). Mitochondrial DNA (mtDNA) can be used as a genetic marker for identifying colonies that have an Africanized queen since a single worker honey bee can represent the entire honey bee colony (Sheppard & Smith, 2000). Introgression of AHB genes using a mitochondrial DNA marker is, however, not detectable if an EHB queen has mated with AHB drones. Besides the concerns about the aggressive nature of AHB to humans, it has been recently shown that AHB can outcompete native pollinator species in southern Utah, resulting in the local extinction of the Andrenid bee *Perdita meconis* Griswold (Portman et al., 2017).

Honey bees have more than 26 subspecies which have been placed into six evolutionary lineages based primarily on morphometrics and their historical geographic distribution (Ruttner, 1988; Sheppard et al., 1997; Franck et al., 2001; Sheppard & Meixner, 2003; Ferreira et al., 2008; Alburaki et al., 2013). These lineages include the A (African group), M (North and Western Europe), C (Southeastern Europe), O (Near East and Middle East) (Ruttner et al., 1978; Ruttner, 1988; Franck et al., 2001; Kandemir et al., 2006; Ferreira et al., 2008; Shaibi et al., 2009), Y (Ethiopia) (Franck et al., 2001), and Z (Syria and Lebanon) (Alburaki et al., 2011). Eight subspecies from four lineages (A, C, M and O) were introduced to the United States
(Pellett 1938; Sheppard, 1989a, 1989b) before the enactment of the 1922 Honey Bee Act which ended all importation of adult honey bees due to the mite, *Acarapis woodi* Rennie, which was responsible for the Isle of Wight disease (Phillips, 1923).

DNA sequence analysis of a portion of the mitochondrial DNA (mtDNA) COI-COII genome can provide information on the mitochondrial lineage of the queen honey bee in a colony. Unlike nuclear genotypes which can be altered during segregation or recombination during reproduction, mtDNA markers are maternally inherited, thus allowing mtDNA analysis to focus on the genealogies of individual lineages (Lansman et al., 1981). This also allows the use of a single individual to genetically characterize a honey bee colony (Sheppard and Smith, 2000). Studies on the genetic variation of the mtDNA COI-COII region of queen breeder honey bee populations in the United States have revealed that the vast majority of queens have only the C lineage (Delaney et al., 2009; Magnus et al., 2011). However, several studies on mtDNA diversity of feral honey bee colonies and swarms from the continental United States and Hawaii (Magnus & Szalanski, 2010; Magnus, 2015; Szalanski et al., 2016) have revealed a number of M and O lineage haplotypes that have not been observed in previous studies of queen breeders (Delaney et al., 2009; Magnus, 2011).

It is believed that feral colonies of European honey bees in the United States are rare in natural areas due to parasites and pathogens, especially *Varroa destructor* Anderson and Trueman, which greatly reduced feral bee populations after its introduction (Seeley, 2015). Several studies have shown that remaining feral colonies of European honey bees have persisted for at least 10 years with infections of *V. destructor* in New York, USA (Seeley, 2007) and in Europe (Fries et al., 2006, Le Conte et al., 2007). Also, beekeepers in the United States have become interested in unmanaged feral colonies for their breeding programs, dubbed ‘survivor
stock’ due to their ability to persevere despite the presence of pathogens and parasites (McNeil, 2009a, 2009b; Jacobson, 2010). These feral colonies may not have been manipulated by beekeepers for a long period of time and may have adapted to the various stresses that threaten their survival (Loper and Sammataro, 2006; Seeley, 2007; Villa et al., 2008).

The objective of this study was to characterize the genetic diversity of honey bees from feral colonies and swarms in Utah, USA using DNA sequence data of the COI-COII mtDNA region of A. mellifera.

Materials and Methods

Adult worker honey bees were collected from feral colonies and swarms, from 10 Utah counties, into 70% ethanol from 2008 to 2017 (Figure 2.1, Table 2.1). Following Schiff et al. (1994), feral colonies are defined as established colonies occurring in unmanaged homesites (e.g., trees, caves, buildings). Samples collected from masses of bees discovered in the open, without comb, are defined as swarms. Swarms are commonly derived from unmanaged colonies, but their origin can be difficult to determine with certainty. DNA was extracted from individual honey bees using a salting-out protocol with in-house reagents (Sambrook & Russell, 2001). PCR Primers E2 and H2 (Garnery et al., 1993) were used to amplify extracted DNA via PCR. These primers will amplify an approximately 530 bp to 1230 bp portion of the mtDNA COI-COII genes. A total of 2 μL of extracted DNA was used for PCR. The PCR reaction, following Taylor et al. (1996), consisted of holding the samples for 5 min at 94°C, then 40 cycles of 94°C for 45s, 46°C for 1 min, and 72°C for 1 min, followed by a final extension of 72°C for 5 min. Amplicon verification was conducted by gel electrophoresis using 1% agarose gels and visualizing PCR products using a BioDoc-it™ Imaging System (UVP, Inc., Upland, CA). Samples were purified and concentrated with VWR centrifugal devices (VWR, Radnor, PA) and
sent to Eurofins Genomics (Huntsville, AL) for direct sequencing in both directions. Voucher specimens are deposited at the University of Arkansas Insect Genetics Laboratory in Fayetteville, Arkansas, USA. Consensus sequences with the primer ends removed were obtained using Geneious v6.1.6 (Biomatters Ltd., Auckland, New Zealand). An unrooted maximum likelihood tree using the observed haplotypes was constructed using Geneious v6.1 using the PHYLIP plugin (Felsenstein 1989) with 1000 resamplings using the HYY 85 substitution model and NNI topology search.

**Results**

A total of 174 samples of feral honey bee colonies and swarms were collected from 10 Utah counties (Figure 2.1) and successfully sequenced (Table 2.1 & 2.2). Genetic similarity of shared nucleotide sites among the observed haplotypes ranged from 92.1 to 99.8% (Table 2.3). The majority of the samples were from Washington (n=94), Iron (n=34), and San Juan (n=26) counties located in southern Utah. A total of 20 COI-COII haplotypes were observed (Table 2.1), with 48% belonging to the A lineage, 43% to the C lineage, 4% to the M lineage, and 5% to the O lineage (Figure 2.2). Samples with A lineage haplotypes (n=83) were observed from six Utah counties (Emery, Iron, Kane, San Juan, Washington, and Garfield). Ten haplotypes of the African A lineage were observed with A1e (n=29), A26a (n=10), and A1b (n=9) being the most common. A total of five C lineage haplotypes were observed (n=75), with C1 (n=40) being most common. For the M lineage, three haplotypes were observed, M3 (n=3), M7 (n=1), and M3a (n=3). Two haplotypes of the O lineage were also observed (n=9) (O2 and O5), with O2 (n=7) being the most common. The maximum likelihood cladogram (Figure 2.2) revealed the C, M, and O lineages forming single clades among their haplotypes, while the A lineage haplotypes formed three distinct
clades. This consisted of haplotype A4a' forming a sister group with haplotypes A26, A26a, A26c, A4 and A4'', while the A1 haplotypes (A1, A1b, A1d, and A1e) formed another distinct clade.

**Discussion**

All of the M lineage haplotypes (M3, M7, and M3a) observed in this study have been previously observed in other western states and in the southern United States (Delaney et al., 2009, Magnus et al., 2014). Haplotype O2 has been observed from California (Kono & Kohn, 2015), and O5 from Oklahoma, New Mexico, and Texas (Magnus et al., 2014). All C lineage haplotypes observed in this study have been observed from queen breeders throughout the United States (Delaney et al., 2009, Magnus et al., 2011).

Among the 10 A lineage haplotypes, only two (A4a’ & A4’’) are unique to Utah when compared to the rest of the United States. Five of the A haplotypes observed in Utah (A1, A1e, A26, A26c, and A4) have also been found in the bordering state, Arizona (Darger 2013). Although Darger (2013) observed a total of 22 A lineage samples in Arizona, only these five A haplotypes were found in Utah in the present study. Also, three of the Utah A lineage haplotypes observed (A1, A1d, and A26a) have been documented in New Mexico, which also borders Utah, although only A1 was observed in all three states (Szalanski & Magnus, 2010). As New Mexico and Arizona are both south of Utah, the currently accepted AHB expansion scenario would suggest that AHB spread northward into Utah from these states. The diversity of haplotypes found throughout the southwest, and in Utah particularly, is surprising, given this expansion scenario. A recent study of genetic diversity of feral and managed honey bee colonies from San Diego County, California found 60% of the 48 worker honey bees sequenced for the COI-COII mtDNA region belonged to the A lineage with 10%, 17%, and 13% belonging to the C, O, and M lineages, respectively (Kono & Kohn, 2015). Only three (A2b, A4a, and A26) of the 10 A
lineage haplotypes observed in Utah occurred among the 10 A lineage haplotypes from San Diego County, California, none of these were also observed in Arizona or New Mexico. This level of genetic diversity is interesting given the close regional proximity between the San Diego samples and those from this study and the proposed patterns of AHB expansion throughout the region.

Most A lineage bees in Utah belong to the A1 haplotype group (55 of 83) as reported in a smaller survey in the past (Szalanski & Magnus, 2010). However, the overall diversity of A lineage haplotypes in Utah is quite high given the limited geographic range sampled. The high number of haplotypes of A lineage honey bees, 10, observed in this study from southern Utah is on par with a previous study of AHB genetic diversity from seven southern U.S. states, which found a total of 12 A lineage haplotypes (Szalanski & Magnus, 2010). In addition, A lineage haplotypes were far more common in this sample (48%) than haplotypes from the other lineages, and two haplotypes have only been observed in Utah thus far. Also, the high proportion of AHB from the feral Utah honey bee samples is of importance given the ability of AHB to outcompete native pollinator species in southern Utah, resulting in the local extinction of the Andrenid bee *Perdita meconis* (Portman et al., 2017). This brings a new component to the importance of documenting the occurrence of AHB in Utah and throughout the southern United States.

In addition, given the high diversity of mtDNA haplotypes observed among feral honey bees in Utah, this gives evidence that these feral honey bees were not exposed to *varroa* mites. Populations of feral honey bees found in a forest in New York, USA had a dramatic loss of mtDNA genetic diversity after exposure to *varroa* mites (Mikheyev et al. 2015). Also, from our samples for our study, usually consisting of 20-40 worker honey bees, we have not observed any
varroa mites. This sample size was not ideal for varroa mite surveillance, but it does provide evidence of a lack of exposure to Varroa mites in these populations.

The commonness of the A lineage in Utah, Arizona (Darger, 2013), New Mexico (Szalanski & Magnus, 2010) and San Diego County, California (Kono & Kohn, 2015) provides evidence that A lineage honey bees may be well adapted to the arid climate in the southwestern United States. The adaptiveness to this climate is also evident by A lineage samples in our study being recovered from elevations as high as 1357 m. The highest elevation that samples were found in our study was 1868 m for a colony of O lineage honey bees. A recent study by Wallberg et al. (2017) on honey bees from Eastern Africa found that A. m. scutellata honey bees were only recovered from locations where elevations remained below 1100 m. The presence of A lineage honey bees at elevations greater than 1300 m. in Utah could be due to hybridization between Africanized honey bees and European honey bees allowing them to adapt to higher elevations. Another possible reason is that the A lineage samples collected in our study from Utah, which were found at elevations as high as 1357 m., are not A. m. scutellata, but a different A lineage subspecies. This could provide evidence that some of the A lineage honey bees from Utah, are not A. m. scutellata that entered the United States in 1990 from Brazil (Sugden & Williams, 1990), and were introduced to Utah before 1990.
Acknowledgements

We especially want to thank Jerry Shue for collecting honey bee samples, as well as the beekeepers in the state of Utah and the Utah Department of Agriculture and Food for assisting with sample collection. This research was supported in part by the University of Arkansas, Arkansas Agricultural Experiment Station.
References


Kono, Y., & Kohn, KR. (2015). Range and frequency of Africanized honey bees in California (USA). *PLOS ONE*, DOI: 10.1371/journal.pone.0137407


Figure 2.1. Counties sampled in Utah for *Apis mellifera* feral colonies and swarms. Numbers indicate sample sizes. Counties with Africanized honey bees (AHB) are shown in yellow; counties that were sampled, but in which no AHB were observed are shown in green.
| County     | A1 | A1b | A1d | A1e | A26 | A26a | A26c | A4  | A4a | A4" | C1  | C2  | C11 | C12 | C31 | M3  | M7  | M13a | O1 | O2 | O5 | Total |
|------------|----|-----|-----|-----|-----|------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|----|----|----|-------|
| Carbon     | 1  |     |     |     |     |      |      |     |     |     |     |     |     |     |     |     |      |    |    |    |       |
| Emery      | 1  |     |     |     |     |      |      |     |     |     |     |     |     |     |     |     |      |    |    |    |       |
| Garfield   | 1  | 1   |     |     |     |      |      |     |     |     |     |     |     |     |     |     |      |    |    |    |       |
| Grand      |     |     |     |     |     |      |      |     |     |     |     |     |     |     |     |     |      |    |    |    |       |
| Iron       | 3  | 3   | 7   | 3   | 4   | 1    | 3    | 1   | 2   | 1   | 1   | 3   | 34  |     |     |     |      |    |    |    |       |
| Kane       |     |     |     |     |     |      |      |     |     |     |     |     |     |     |     |     | 6    |    |    |    |       |
| Salt Lake  |     |     |     |     |     |      |      |     |     |     |     |     |     |     |     |     |      |    |    |    |       |
| San Juan   | 4  | 6   |     |     |     |      | 9    | 1   | 1   | 1   | 1   | 2   | 26  |     |     |     |      |    |    |    |       |
| Sevier     |     |     |     |     |     |      |      |     |     |     |     |     |     |     |     |     |      |    |    |    |       |
| Washington | 5  | 6   | 2   | 15  | 1   | 6   | 2    | 1   | 5   | 22  | 2   | 18  | 1   | 2   |     |     |      |    |    |    |       |
| Total      | 12 | 9   | 5   | 29  | 4   | 10  | 3    | 1   | 7   | 40  | 7   | 23  | 1   | 4   | 3   | 1   | 1   | 2   | 7   | 2   | 174 |     |
Figure 2.2. Maximum likelihood phylogenetic tree showing the relationship among observed haplotypes.
Table 2.2. Observed haplotypes, percent match to GenBank Accessions, and matching GenBank Accession number.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Haplotype</th>
<th>Percent Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>EF033649</td>
<td>100</td>
</tr>
<tr>
<td>A1b</td>
<td>FJ477985</td>
<td>100</td>
</tr>
<tr>
<td>A1d</td>
<td>FJ743639</td>
<td>100</td>
</tr>
<tr>
<td>A1e</td>
<td>GU326335</td>
<td>100</td>
</tr>
<tr>
<td>A26</td>
<td>KJ661742</td>
<td>100</td>
</tr>
<tr>
<td>A26a</td>
<td>FJ743640</td>
<td>99.9</td>
</tr>
<tr>
<td>A26c</td>
<td>FJ890929</td>
<td>100</td>
</tr>
<tr>
<td>A4</td>
<td>EF033650</td>
<td>100</td>
</tr>
<tr>
<td>A4a’</td>
<td>KX463808</td>
<td>100</td>
</tr>
<tr>
<td>A4”</td>
<td>FJ478009</td>
<td>99.7</td>
</tr>
<tr>
<td>C1</td>
<td>EF033655</td>
<td>100</td>
</tr>
<tr>
<td>C2</td>
<td>JF934704</td>
<td>100</td>
</tr>
<tr>
<td>C11</td>
<td>FJ037776</td>
<td>100</td>
</tr>
<tr>
<td>C12</td>
<td>FJ037777</td>
<td>100</td>
</tr>
<tr>
<td>C31</td>
<td>HQ287900</td>
<td>100</td>
</tr>
<tr>
<td>M3</td>
<td>FJ743636</td>
<td>100</td>
</tr>
<tr>
<td>M7</td>
<td>KX463911</td>
<td>99.7</td>
</tr>
<tr>
<td>M3a</td>
<td>KX463884</td>
<td>100</td>
</tr>
<tr>
<td>O2</td>
<td>FJ477996</td>
<td>100</td>
</tr>
<tr>
<td>O5</td>
<td>FJ743633</td>
<td>100</td>
</tr>
</tbody>
</table>
Chapter 3: Mitochondrial DNA genetic variation in Arkansas honey bee, *Apis mellifera* L., colonies

Abstract

This study characterized the mitochondrial DNA (mtDNA) genetic variation in Arkansas honey bee, *Apis mellifera* L., populations by sequencing a portion of the mitochondrial cytochrome oxidase (COI-COII) intergenic region. The samples were primarily of hobbyist-managed origin (n=179), as well as 32 feral colonies and two swarms. Of the 213 honey bee colonies and swarms sampled, 25 haplotypes were observed. The haplotypes were from the: A (African) (1.88%); C (Eastern European) (92.49%); M (Northern and Western European) (3.29%); and O (Near East and Middle East) lineages (2.35%). Six C lineage haplotypes were predominantly detected (n=188, 88.26%), all of which are common in U.S. commercial queen breeder colonies. The remaining 24 honey bee samples represented 17 haplotypes, all of which are absent from commercial queen breeder colonies but have been observed in feral honey bee populations collected in other States. These haplotypes, particularly those from the M and O lineages, are likely hundred-year-old remnants of historical importations, surviving for generations despite the arrival of threats, such as varroa mites. Understanding honey bee genetics and population structure are valuable for maintaining genetic diversity. Results from this study provide evidence that Arkansas honey bee populations differ from U.S. commercial queen breeder colonies. The 17 haplotypes detected in our Arkansas study, which were absent from commercial queen breeder colonies, could be important sources of genetic diversity in future honey bee breeding programs, highlighting the importance of State-level genetic surveys.
Introduction

The western honey bee, _Apis mellifera_ L. (Hymenoptera: Apidae), has a geographic distribution spanning every continent except Antarctica (Crane, 1999; Ruttner, 1988). This expansive distribution, beyond its proposed native range of Africa, Asia, and Europe, is due to human-mediated migration and natural swarming (Crane, 1999; Ruttner, 1988; Seeley, 1985; Sheppard & Meixner, 2003). Humans have kept honey bees for over 10,000 years, primarily managing for honey, wax, and, more recently, pollination services (Batra, 1995; Crane, 1999). Today, _A. mellifera_ is the single most economically important pollinator worldwide, valued at greater than 200 billion USD annually for its pollination services (Gallai et al., 2009). Over the past several decades, increased winter mortality and reduced colony health have resulted in elevated concern and an increasing interest in identifying contributing stressors (Cox-Foster et al., 2007; Oldroyd, 2007; USDA-NASS, 2022; vanEngelsdorp et al., 2012). Reduced genetic variation is among the suggested contributing factors to honey bee health (De la Rúa et al., 2009; vanEngelsdorp & Meixner, 2010; Zayed, 2009). Adequate genetic diversity is associated with increased disease resistance, immunity, productivity, and fitness in honey bees (Mattila & Seeley, 2007; Tarpy, 2003).

Based on morphometric analysis, molecular data, and biogeographic origin, _A. mellifera_ has been divided into approximately 30 subspecies, within six evolutionary lineages (Alburaki et al., 2013; Ferreira et al., 2009; Franck et al., 2001; Ilyasov et al., 2020; Meixner et al., 2013; Ruttner, 1988; Sheppard et al., 1997; Sheppard & Meixner, 2003). These lineages include the A (African group), M (North and Western Europe), C (Southeastern Europe), O (Near East and Middle East) (Ferreira et al., 2009; Franck et al., 2001; Kandemir et al., 2006; Ruttner, 1988; Ruttner et al., 1978; Shaibi et al., 2009), Y (Ethiopia) (Franck et al., 2001), and Z (Syria and
Lebanon) (Alburaki et al., 2013). Within lineages, unique genetic haplotypes occur; there is an average nucleotide divergence of 2.5% between the A, M, and C lineages (Garnery et al., 1992).

Considerable morphological and behavioral variation is exhibited between A. mellifera lineages and subspecies adapted from their geographic origin’s climate and environmental conditions (Ruttner, 1988). This variation primarily manifests in the form of varied body, wing, and tongue size as well as overwintering, foraging, defense, and swarming behavior (Cobey et al., 2012; Ruttner, 1988; Ruttner et al., 1978). For example, Ruttner (1988) characterizes subspecies originating from northern latitudes, such as A. m. mellifera L. (M), by their large, broad, and densely “haired” abdomens. Subspecies originating from Africa, such as A. m. lamarckii Cockerell, are characterized by their short wings, tongue, and legs and small slender body. Behaviorally, A. m. ligustica Spinola (C) is regarded as a docile and productive subspecies, while A. m. syriaca Skorikov (O) is also known to yield large volumes of honey; however, it is notorious for its highly defensive nature and propensity to swarm.

The European dark bee, A. m. mellifera (M), was the only subspecies imported into the U.S. for nearly 200 years. The A. m. mellifera colonies adapted to the eastern U.S. forests, eventually swarming and establishing feral colonies (Crane, 1999; Horn, 2005; Kritsky, 1991). Reportedly, honey bees first crossed the Mississippi River by swarm in 1792, and by 1818 honey bees had swarmed to Arkansas (Crane, 1999; Kritsky, 1991). Following the successful establishment of A. m. mellifera in the U.S., importation of other honey bee subspecies began, including A. m. ligustica (C), A. m. lamarckii (A), A. m. carnica Pollmann (C), A. m. cypria Pollmann (O), A. mellifera syriaca (O), A. mellifera caucasica Pollman (C), and A. mellifera intermissa Buttel-Reepen (A) (Crane, 1999; Horn, 2005; Pellett, 1938; Sheppard, 1989a, 1989b). Between 1859 and 1914 movement and importation of honey bee colonies accelerated due to
advances in transportation and migration of humans to the west coast as well as via natural swarming (Cobey et al., 2012; Crane, 1999).

*Apis mellifera ligustica* (C), *A. m. carnica* (C), and *A. m. caucasica* (C) were maintained and widely favored by beekeepers due to their docile behavior and increased honey production (Watkins, 1968). Conversely, M, O, and A lineage subspecies were quickly dismissed by beekeepers due to undesirable behaviors such as lower honey production, aggressiveness, and a propensity for swarming; however, there is evidence of their persistence in feral populations in the U.S. (Cleary et al., 2018; Cobey et al., 2012; Magnus & Szalanski, 2009; Schiff et al., 1994; Sheppard, 1989a; Szalanski et al., 2016).

The limited subset of *A. mellifera* imported into the U.S. is the first of three successive genetic bottleneck events which contributed to reduced genetic variation in U.S. honey bee populations (Cobey et al., 2012; Horn, 2005; Sanford, 2001; Sheppard, 1989a, 1989b; vanEngelsdorp & Meixner, 2010). Importation of honey bees, and later their germplasm, was halted in 1922 in response to the Isle of Wright disease and later to avoid the introduction of *A. m. scutellata* Lepeletier (A) (Cobey et al., 2012; Fracker et al., 1923). Nevertheless, *A. m. scutellata* hybrids swarmed into southern Texas in 1990 after an unintentional release in 1957 from an experimental breeding program in Brazil (Winston, 1992). Thus, a total of nine *A. mellifera* subspecies, compared to the estimated 30 subspecies globally, have reportedly arrived in the U.S. since 1622.

The second genetic bottleneck relates to the commercial breeding industry and apicultural practices. U.S. beekeepers have long prioritized C lineage honey bees due to their desirable colony size, productivity, docility, and low proneness to swarm (Watkins, 1968). This selectivity resulted in commercial honey bee stock being almost exclusively C lineage origin (Delaney et
al., 2009; Magnus et al., 2011; Schiff & Sheppard, 1995, 1996). Reportedly, an estimated 500 breeder queens have been used to produce approximately 900,000 daughter queens for commercial sale in the U.S., suggesting commercial queen production in the U.S. has a very limited genetic pool (Delaney et al., 2009; Schiff & Sheppard, 1995, 1996). Also, selective breeding, supplemental feeding, and frequent administering of medication to control harmful agents have the potential to sustain and breed weak colonies unlikely to survive under unmanaged conditions (Brodschneider & Crailsheim, 2010; DeGrandi-Hoffman & Chen, 2015; Raymann et al., 2017). These breeding and apicultural practices may have hindered natural selection, selecting for docility and productivity rather than disease resistance and local adaptiveness (Hatjina et al., 2014; Seeley et al., 2015). However, studies have found feral populations of C, M, O, and A lineage honey bees different than the common C haplotypes in commercial breeding operations, suggesting remnants of past importations have persisted for several generations (Cleary et al., 2018; Magnus et al., 2014; Schiff et al., 1994). There is increasing interest in enhancing the genetic variability of U.S. honey bee breeding stock. More specifically, incorporating feral survivor stock into breeding operations has the potential to diversify breeder stock and improve colony health; however, there is concern over introducing aggressive Africanized traits (Mattila & Seeley, 2007; McNeil, 2009; Rangel et al., 2020).

The third genetic bottleneck was the arrival of the parasitic *V. destructor* mite in the 1980s, dramatically reducing honey bee populations in the U.S., particularly feral populations (Kraus & Page Jr, 1995; Sanford, 2001; Wenner & Bushing, 1996). In parts of California, over 85% of feral colonies were killed between 1990 and 1994 (Kraus & Page Jr, 1995), and 96-99% were lost in parts of Arizona between 1996 and 1998 (Loper et al., 2006) due to mite infestation. The small percentage of feral colonies persisting after *V. destructor*’s arrival have since been
dubbed “survivor stock” (McNeil, 2009). Prior to their decimation, genetic studies showed that feral honey bee populations in the U.S. were distinctly different from managed populations (Schiff et al., 1994; Schiff & Sheppard, 1993). Furthermore, studies have shown feral honey bee colonies have higher immunocompetency in response to not only *V. destructor* but also *Nosema* sp. and exhibit a lower occurrence of pests and parasites compared to managed colonies (Gilliam & Taber, 1991; López-Uribe et al., 2017; Szalanski et al., 2014; Youngsteadt et al., 2015).

Morphometric analysis to distinguish between honey bee subspecies is tedious and requires measuring several characteristics (>24 characteristics) from multiple individual bees within a single colony and averaging this data to perform probability analyses (Alattal et al., 2014; Ruttner, 1988). Furthermore, morphometric analysis is unsuitable for inferring phylogenetic relationships (Garnery et al., 1992). Molecular techniques are sensitive and reliable tools for identification purposes, i.e., distinguishing between Africanized and European or identifying genetic haplotypes (Cleary et al., 2018; Szalanski et al., 2014; Szalanski & McKern, 2007; Szalanski & Tripodi, 2014). Molecular techniques, specifically DNA sequencing, also allow for phylogenetic analysis to better understand the dispersion of subspecies and their genetic relationships (Arias & Sheppard, 1996; Ilyasov et al., 2021). Mitochondrial markers are maternally inherited in animals and do not undergo recombination, allowing for the ancestral lineage genealogy to be maintained (Brown, 1985; Lansman et al., 1981). Because all individuals within a colony are progeny of the queen or the queen’s daughter, a single individual may be used to characterize an entire colony’s maternal ancestor (Garnery et al., 1992; Sheppard & Smith, 2000).

In this study, we characterized the mtDNA genetic diversity in Arkansas honey bee populations, analyzing samples from hobbyist-managed colonies, feral colonies, and swarms.
The mtDNA COI-COII intergenic region of *A. mellifera* exhibits a high level of genetic variability within and among *A. mellifera* lineages, making it useful for distinguishing lineages as well as identifying unique haplotypes (Franck et al., 2001; Meixner et al., 2013). Exploring the relationship and influence of management and geographic location on genetic diversity may improve future honey bee breeding programs.

**Materials and Methods**

Adult worker honey bees were collected from managed and unmanaged (feral) colonies, as well as from swarms in Arkansas. Samples were acquired by beekeepers, the Arkansas Plant Board, and our own collection efforts from 2005 to 2022 and preserved in 70% ethanol. Voucher specimens are stored at the University of Arkansas Insect Genetics Laboratory in Fayetteville, Arkansas, USA.

DNA was extracted from individual worker honey bees using a salting-out protocol with in-house reagents and stored at -20°C (Sambrook & Russell, 2001). PCR was performed as described in Szalanski (2000), using paired COI-COII PCR primers E2 (5’-GGCAGAATAA GTGCATTG-3’) and H2 (5’-CAATATC ATTGATGACC-3’) (Garnery et al., 1998) and the following thermocycler conditions: denatured initially for 5 minutes at 94°C then 40 cycles at 94°C for 45 seconds, 46°C for 1 minute, 72°C for 1 minute and a final extension of 72°C for 5 minutes (Taylor et al., 1997). These primers amplify the polymorphic intergenic region between the COI and COII genes of the mtDNA. PCR products were run on a 2% agarose check gel and visualized using the BioDoc-it™ Imaging System (UVP, Inc., Upland, CA). A 600-1200 bp amplicon indicated successful DNA extraction; the size variation is due to an intergenic spacer region which varies in size among honey bee lineages.

62
Successfully amplified PCR product was purified and concentrated using VWR centrifugal devices (VWR, Radnor, PA). A 1% agarose check gel with 2 μL of filtered product was run and visualized to confirm DNA amplification. Purified DNA was combined with the E2 and H2 primers and sent to Eurofins Genomics (Diatherix, Huntsville, AL) for direct sequencing in both directions. Consensus sequences with the primer ends removed were obtained using Geneious v6.1.8 (Biomatters Ltd., Auckland, New Zealand), and unique haplotypes were identified by eye. The assignment and comparison of haplotypes was conducted using an NIH BLAST (National Center for Biotechnology Information) search of DNA sequences available on GenBank and the University of Arkansas Insect Genetics lab’s database (ALS unpublished data).

Results

A total of 213 honey bee colony and swarm samples from 47 Arkansas counties were successfully characterized by mtDNA COI-COII sequencing (Figure 3.1). The samples were from hobbyist-managed colonies (84.0%) as well as feral colonies (15.0%) and swarms (1.0%).

From the 213 honey bee samples, 25 haplotypes were detected representing four genetic lineages; frequencies and percentages of detected lineages are presented in Table 3.1. Within the C lineage, 13 haplotypes were observed, six of which, C1, C2, C11, C12, C19, and C31, represented 88.26% of the samples (n=188). The M lineage represented 3.29% of the samples and was composed of seven haplotypes, while the A and O lineage were composed of three and two haplotypes, respectively (Table 3.1). The distribution of COI-COII haplotypes by Arkansas counties is provided in Figure 3.1.

Twelve haplotypes occurred in the 32 feral colonies (4 lineages), 20 haplotypes in the 179 hobbyist-managed colonies (4 lineages), and two haplotypes from the two swarms (2
lineages). Geographically, all four A lineage honey bee samples occurred in southeastern Arkansas. The remaining lineages were distributed across the state of Arkansas (Figure 3.1).

**Discussion**

Of the 213 honey bee colonies sampled, 25 haplotypes within four genetic lineages (A, C, M, and O) were observed (Table 3.1). These results are markedly different from those found in the Magnus et al. 2011 commercial queen breeder study, which detected seven haplotypes within a single lineage (C). Unsurprisingly, the C lineage represented the majority of the colonies sampled (92.49%); these results are consistent with previous genetic studies in commercial breeding operations (Delaney et al., 2009; Magnus et al., 2011), managed colonies (Rangel et al., 2020; Seeley et al., 2015), and unmanaged colonies (Magnus et al., 2014; Rangel et al., 2020; Seeley et al., 2015). Furthermore, six C lineage haplotypes (C1, C2, C11, C12, C19, and C31), all commonly detected amongst commercial queen breeders, were the most frequently found haplotypes, representing 88.26% of the colonies sampled (Magnus et al., 2011). This suggests that hobbyist-managed colonies in Arkansas are primarily receiving their genetics from commercial breeder stock. The remaining 17 haplotypes were not detected in the Magnus et al. (2014) study; however, they have been detected in feral and unmanaged honey bee populations, suggesting that genetics from initial historical importations have persisted in Arkansas honey bee populations (Cleary et al., 2018; Magnus et al., 2014; Rangel et al., 2020; Szalanski et al., 2016).

Relative to previous mtDNA genetic studies, this Arkansas study had high haplotype diversity (Magnus et al., 2014; Rangel et al., 2020; Seeley et al., 2015). A 2014 study examined unmanaged honey bee populations from 12 states, observing 23 haplotypes within three lineages (C, M, & O) (Magnus et al., 2014). Studies in New York (Seeley et al., 2015) and Pennsylvania (Rangel et al., 2020) sampled both managed and unmanaged honey bee populations detecting
three and nine haplotypes, respectively, the majority of which (100% and 91.30%) were of the C lineage. While sampling bias may contribute to differences in observed haplotypes and their frequency, the reduced haplotype diversity in the unmanaged colonies in New York and Pennsylvania compared to other feral studies in the west, southwest, and southeast U.S. is potentially due to differences in climate (Cleary et al., 2018; Magnus et al., 2014; Rangel et al., 2020; Seeley et al., 2015). The cold harsh climates of the northeast U.S., combined with parasite pressure, may have eliminated feral populations, specifically populations of tropical and subtropical descent (A and O), more fully compared to the temperate southern climates which have feral populations of A, C, M, and O descent.

Geographically, C and M lineages are common across the country in feral colonies (Magnus et al., 2014). Whereas O and A lineage haplotypes have only been detected in western, southwestern, and southeastern States (Cleary et al., 2018; Kono & Kohn, 2015; Magnus et al., 2014; Magnus & Szalanski, 2009). Furthermore, our study found that haplotypes within the C, O, and M lineage were distributed across the state, while the A lineage haplotypes were isolated to the southwestern part of the State (Figure 3.1). An Arizona study conducted by Atmowidjojo et al. (1997) found that feral honey bee colonies in Arizona were more tolerant of high temperatures than managed colonies. Arkansas is a geographically diverse state, consisting of mountains, forests, river valleys, lakes, and marshes, located in the southeastern U.S. All of Arkansas falls within the humid subtropical Köppen climate classification, characterized by long hot summers and short mild winters (Belda et al., 2014). The Arkansas climate may be more suitable than other States that are more arid or those with colder winters, for the survival and persistence of feral honey bee colonies.
Understanding honey bee genetics is essential for retaining genetic diversity and developing commercial breeding programs. Our study shows that there are populations of honey bees in Arkansas genetically distinct from commercial queen breeder colonies and northeastern U.S. honey bee populations (Magnus et al., 2011; Rangel et al., 2020; Seeley et al., 2015). The haplotype diversity observed in our study suggests that Arkansas hobbyist-managed colonies likely have been diversified by well-established feral colonies which are remnant populations of historical importations from the 17th-19th century. This provides evidence that feral colonies are not all recent swarms from managed colonies derived from commercial stock. Feral honey bees have been proposed as a source of untapped genetic diversity which could be implemented into honey bee breeding programs (Cobey et al., 2012). Our study provides evidence that hobbyist-managed colonies may serve as a source of genetic diversity, via swarms caught by beekeepers that have ancestral lineage haplotypes. These reservoirs may harbor genetic variation, with allelic combinations potentially related to desirable immunity or disease-resistant traits.

In its native range within Europe, *A. m. mellifera* has been the subject of conservation efforts due to the subspecies being heavily hybridized and displaced by introduced C lineage honey bees (Hassett et al., 2018; Pinto et al., 2014; Soland-Reckeweg et al., 2009). Surviving populations of *A. m. mellifera* throughout Europe are being explored as a source of desirable traits which could be implemented into breeding programs (De la Rúa et al., 2009; Hassett et al., 2018; Meixner et al., 2010). Based on our study, there is evidence that surviving populations of M lineage honey bees have persisted and adapted to U.S. conditions for over a hundred years, indicating a domestic source of genetics distinct from current U.S. queen breeders.

Future research should examine parasite and pathogen occurrence within the represented genetic lineages and explore susceptibility to harmful agents in commonly occurring haplotypes.
in managed populations compared to haplotypes associated with persisting feral populations. Also, it would be useful to determine if these colonies with historical lineages have any trait difference relative to colonies obtained from queen breeders (i.e., use of winter stores, spring brood development, varroa mite resistance). Daughter queens from these colonies with useful traits could then be further evaluated by U.S. queen breeders, and potentially distributed to beekeepers in the U.S.
Acknowledgements

We would like to acknowledge the Arkansas beekeepers who agreed to participate in this study. This research was supported in part by the University of Arkansas, Arkansas Agricultural Experiment Station.
References


De la Rúa, P., Jaffé, R., Dall’Olio, R., Muñoz, I., & Serrano, J. (2009). Biodiversity, conservation and current threats to European honeybees. *Apidologie, 40*(3), 263–284. [https://doi.org/10.1051/apido/2009027](https://doi.org/10.1051/apido/2009027)


honey bee colonies (*Apis mellifera*). *Conservation Genetics, 18*(3), 659–666. https://doi.org/10.1007/s10592-017-0942-x


Schiff, N. M., & Sheppard, W. S. (1996). Genetic differentiation in the queen breeding population of the western United States. *Apidologie, 27*(2), 77–86. [https://doi.org/10.1051/apido:19960202](https://doi.org/10.1051/apido:19960202)


vanEngelsdorp, D., Caron, D., Hayes, J., Underwood, R., Henson, M., Rennich, K., Spleen, A., Andree, M., Snyder, R., Lee, K., Roccasecca, K., Wilson, M., Wilkes, J., Lengerich, E.,


Figure 3.1. Distribution of COI-COII *Apis mellifera* haplotypes in Arkansas. The size of each pie chart is proportional to the number of samples analyzed for that county, and the slices indicate the proportion of samples of each haplotype found. Haplotypes previously detected in commercial queen breeding operations were combined and comprise the QB_Haps in dark grey (C1, C2, C11, C12, C19, & C31). The counties sampled from are indicated in grey with counties not sampled in white.
Table 3.1. Frequency and percent of haplotypes detected within the Africanized (A), Eastern European (C), Western European (M), and Oriental (O) lineages in total and between management source types.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Haplotype</th>
<th>Frequency</th>
<th>Unmanaged (%)</th>
<th>Managed (%)</th>
<th>Total Samples (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (1.88%)</td>
<td>A1</td>
<td>2</td>
<td>100</td>
<td>0.00</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>A1d</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>A4&quot;</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>C1</td>
<td>75</td>
<td>2.67 (1.39)</td>
<td>95.83</td>
<td>35.21</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>19</td>
<td>73.68</td>
<td>26.32</td>
<td>8.92</td>
</tr>
<tr>
<td></td>
<td>C2j</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>C2s</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>C11</td>
<td>54</td>
<td>3.70</td>
<td>90.74</td>
<td>25.35</td>
</tr>
<tr>
<td></td>
<td>C12</td>
<td>12</td>
<td>16.67</td>
<td>83.33</td>
<td>5.63</td>
</tr>
<tr>
<td></td>
<td>C19</td>
<td>5</td>
<td>0.00</td>
<td>100</td>
<td>2.35</td>
</tr>
<tr>
<td></td>
<td>C26</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>C29</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>C30</td>
<td>2</td>
<td>0.00</td>
<td>100</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>C31</td>
<td>21</td>
<td>4.76</td>
<td>95.24</td>
<td>9.86</td>
</tr>
<tr>
<td></td>
<td>C33</td>
<td>3</td>
<td>0.00</td>
<td>100</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>C35</td>
<td>2</td>
<td>0.00</td>
<td>100</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>M1</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>M4</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>M5</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>M6</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>M17a</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.47</td>
</tr>
<tr>
<td>O (2.35%)</td>
<td>O5&quot;b</td>
<td>4</td>
<td>25.00 (25.00)</td>
<td>50.00</td>
<td>1.88</td>
</tr>
<tr>
<td></td>
<td>O5d</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.47</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>25</td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

Note: Values in parentheses represent swarm source type.
Appendix 3.1. Fisher's exact test with Yate’s correction results. Significance indicates an association between genetic lineage and source type.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Feral</th>
<th>Managed</th>
<th>Swarm</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>24[29.75]</td>
<td>173[166.39]*</td>
<td>1[1.86]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>A</td>
<td>3[0.60]*</td>
<td>1[3.36]</td>
<td>0[0.04]</td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>2[0.75]*</td>
<td>2[4.20]</td>
<td>1[0.05]*</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3[0.90]*</td>
<td>3[5.04]</td>
<td>0[0.06]</td>
<td></td>
</tr>
</tbody>
</table>

Note: Values formatted as Observed[Expected]. *Significance level of α=0.05
### Appendix 3.2. Frequency of haplotypes recorded from previous U.S. mtDNA COI-COI studies compared to our study. Parentheses indicate unmanaged source type.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>AR Managed &amp; Unmanaged (n=213)</th>
<th>U.S. Queen Breeders (n=140)</th>
<th>U.S. Unmanaged (n=247)</th>
<th>PA Wild &amp; Managed (n=46)</th>
<th>NY Arnot Forest (n=28)</th>
<th>UT Feral (n=174)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>(2)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>A1b</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>A1d</td>
<td>(1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>A1e</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>A4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>A4a'</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>A4''</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>A26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>A26a</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>A26c</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C1</td>
<td>69(3)</td>
<td>46</td>
<td>(11)</td>
<td>6(17)</td>
<td>3(4)</td>
<td>40</td>
</tr>
<tr>
<td>C1a</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C1i</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C2</td>
<td>5(14)</td>
<td>24</td>
<td>94(1)</td>
<td>0</td>
<td>17(4)</td>
<td>7</td>
</tr>
<tr>
<td>C2j</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C2s</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C2c</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>(1)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C11</td>
<td>49(5)</td>
<td>45</td>
<td>(49)</td>
<td>3(8)</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>C12</td>
<td>10(2)</td>
<td>15</td>
<td>(15)</td>
<td>(2)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>C19</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C26</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C27</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C29</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C30</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C31</td>
<td>20(1)</td>
<td>8</td>
<td>(20)</td>
<td>(2)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>C33</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C34</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C35</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M1</td>
<td>(1)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M2</td>
<td>(1)</td>
<td>0</td>
<td>(2)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M3</td>
<td>1</td>
<td>0</td>
<td>(10)</td>
<td>(1)</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>M3a</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M4</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M4q</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>(2)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M5</td>
<td>(1)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>(2)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>M17a</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>O1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>O2</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>O5</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>O5'b</td>
<td>2(2)</td>
<td>0</td>
<td>(3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>O5d</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>173(34)</td>
<td>139</td>
<td>(247)</td>
<td>11(35)</td>
<td>20(8)</td>
<td>174</td>
</tr>
</tbody>
</table>

Appendix 3.3. Comparison of the proportion of C lineage samples compared to samples from the A, M, or O lineage between management source types. Significance indicates an association.

<table>
<thead>
<tr>
<th>Source</th>
<th>C Lineage (%)</th>
<th>A+M+O Lineages</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managed (n=179)</td>
<td>96.65</td>
<td>3.35</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Unmanaged (n=34)</td>
<td>73.53</td>
<td>26.47</td>
<td></td>
</tr>
<tr>
<td>Total (n=213)</td>
<td>92.96</td>
<td>7.04</td>
<td></td>
</tr>
</tbody>
</table>

*Significance level of α=0.05.

Appendix 3.4. Comparison of the proportion of C and AMO lineage groups between total Arkansas colonies samples in our study compared to previous COI-II genetic studies.

<table>
<thead>
<tr>
<th>COI-II Studies</th>
<th>C Lineage (%)</th>
<th>A+M+O (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR (n=213)</td>
<td>92.96</td>
<td>7.04</td>
</tr>
<tr>
<td>U.S. Queen Breeders¹ (n=140)*</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>U.S. Unmanaged² (n=247)*</td>
<td>83.40</td>
<td>16.60</td>
</tr>
<tr>
<td>PA &amp; NY³,⁴ (n=74)</td>
<td>93.24</td>
<td>6.76</td>
</tr>
</tbody>
</table>

*Significance level of α=0.05. ¹(Magnus et al. 2011), ²(Magnus et al. 2014), ³(Rangel et al. 2020), ⁴(Seeley et al. 2015).

Appendix 3.5. Comparison of COI-II studies’ proportion of C and AMO group lineages. Significance is relative to our Arkansas study.

<table>
<thead>
<tr>
<th>COI-II Studies</th>
<th>C Lineage (%)</th>
<th>A+M+O (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR (n=213)</td>
<td>92.96</td>
<td>7.04</td>
</tr>
<tr>
<td>AR Managed (n=179)</td>
<td>96.65</td>
<td>3.35</td>
</tr>
<tr>
<td>AR Unmanaged (n=34)</td>
<td>73.53</td>
<td>26.47</td>
</tr>
<tr>
<td>U.S. Queen Breeders¹ (n=140)*</td>
<td>100</td>
<td>0.00</td>
</tr>
<tr>
<td>U.S. Unmanaged² (n=247)*</td>
<td>83.40</td>
<td>16.60</td>
</tr>
<tr>
<td>PA &amp; NY³,⁴ (n=74)</td>
<td>93.24</td>
<td>6.76</td>
</tr>
</tbody>
</table>

*Significance level of α=0.05. ¹(Magnus et al. 2011), ²(Magnus et al. 2014), ³(Rangel et al. 2020), ⁴(Seeley et al. 2015).
**Appendix 3.6.** Comparison of COI-II studies’ proportion of C and AMO group lineages. Significance is relative to our Arkansas study.

<table>
<thead>
<tr>
<th>COI-II Studies</th>
<th>C Lineage (%)</th>
<th>A+M+O (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR Managed (n=179)</td>
<td>96.65</td>
<td>3.35</td>
</tr>
<tr>
<td>U.S. Queen Breeders¹ (n=140)*</td>
<td>0.00</td>
<td>100%</td>
</tr>
<tr>
<td>U.S. Unmanaged² (n=247)*</td>
<td>83.40</td>
<td>16.60</td>
</tr>
<tr>
<td>PA &amp; NY³,⁴ (n=74)</td>
<td>93.24</td>
<td>6.76</td>
</tr>
</tbody>
</table>

*Significance level of α=0.05. ¹(Magnus et al. 2011), ²(Magnus et al. 2014), ³(Rangel et al. 2020), ⁴(Seeley et al. 2015).

**Appendix 3.7.** Comparison of COI-II studies’ proportion of C and AMO group lineages. Significance is relative to unmanaged honey bee samples from our Arkansas study.

<table>
<thead>
<tr>
<th>COI-II Studies</th>
<th>C Lineage (%)</th>
<th>A+M+O (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR Unmanaged (n=34)</td>
<td>73.53</td>
<td>26.47</td>
</tr>
<tr>
<td>U.S. Queen Breeders¹ (n=140)*</td>
<td>0.00</td>
<td>100%</td>
</tr>
<tr>
<td>U.S. Unmanaged² (n=247)</td>
<td>83.40</td>
<td>16.60</td>
</tr>
<tr>
<td>PA &amp; NY³,⁴ (n=74)*</td>
<td>93.24</td>
<td>6.76</td>
</tr>
</tbody>
</table>

*Significance level of α=0.05. ¹(Magnus et al. 2011), ²(Magnus et al. 2014), ³(Rangel et al. 2020), ⁴(Seeley et al. 2015)
Chapter 4: Genetic variation and pathogen occurrence in managed and feral honey bee colonies from the south-central U.S., Utah, and Hawaii

Abstract

A mitochondrial DNA (mtDNA) genetic diversity and pathogen study was conducted examining swarms, managed, and feral honey bee populations from three areas of the U.S., south-central (SC) U.S., Hawaii, and Utah. Pathogens were identified using a combination of molecular diagnostic techniques. We sequenced a portion of the mtDNA cytochrome oxidase (COI-COII) intergenic region to characterize the maternal origin of 766 honey bee samples. The C lineage was the most common at 67.36%, followed by the A (16.32%), M (12.92%), and O lineage (3.39%). Within the four lineages, 43 different haplotypes were detected, several of which have not been found in commercial queen breeding operations. Unique haplotypes occurred regionally and within management source types. Honey bee pathogens and parasites Nosema sp., Lotmaria passim, and Varroa destructor were found in all lineages, all three were least prevalent in A lineage samples. Nosema sp. was detected in 17.68% of the samples, while L. passim was observed in 14.36% of the samples. A significant association was detected between Nosema sp. infection and lineage (p<0.001). The rate of Nosema sp. infection was highest in the M (29.29%) and C lineages (19.19%). Lotmaria passim was most prevalent in the M (22.92%) and C (14.03%) lineages. The proportion of samples with L. passim infection was significantly associated with management source (p<0.001). Varroa destructor was most common and abundant in the C (25.84%) lineage samples, however, no differences in mite loads were observed between lineages. Varroa loads were significantly different between management source (p<0.001) and between regions (p<0.001). This study is among the first to comprehensively explore genetic origin, management source, and regionality as factors of pathogen infection in U.S. honey bee populations.
Introduction

Managed and feral honey bees, *Apis mellifera* L., in the U.S., have suffered from several biotic and abiotic stressors over the past several decades resulting in reduced colony health and fluctuations in populations (Loper et al., 2006; Potts et al., 2010; vanEngelsdorp et al., 2009). Among the major stressors attributable to reduced health in *A. mellifera* are: pathogens; improper pesticide usage; habitat loss; poor nutritional resources; management stress; and reduced genetic diversity (Alger et al., 2018; Dolezal & Toth, 2018; Evans & Schwarz, 2011; Oldroyd, 2007).

In 1622 *Apis mellifera mellifera* L. (North and Western European lineage, “M”) was the first and, for over 200 years, the only honey bee subspecies imported into the U.S. *Apis m. mellifera* acclimated quickly and became widely distributed across the country due to natural swarming and human-mediated movement (Crane, 1999; Sheppard, 1989a, 1989b). Eventually, an additional eight subspecies from three genetic lineages, Eastern European (C); African (A); and Near and Middle East (O), were introduced to the U.S.; this includes the accidental introduction of *Apis m. scutellata* (A) (Horn, 2005; Pellett, 1938; Sheppard, 1989b, 1989a).

While there is no official documentation of its importation, *Apis mellifera iberiensis* Engel (M & A) was likely imported by Spanish settlers to Mexico in the 1500s (Brand, 1988; Seeley, 2019; Sheppard, 1989b). *Apis m. iberiensis* was also among the subspecies hybridized with *A. m. scutellata* in Brazil (Crane, 1999). Movement and importation of honey bee colonies increased from 1859-1914 due to advances in transportation, i.e.: transcontinental railroads; migration of human settlers across the country; and the honey bee’s drive to reproduce via swarming, resulting in managed and feral honey bee populations, from different origins, across the U.S. (Cobey et al., 2012; Crane, 1999).
Sequential genetic bottleneck events led to reduced genetic stock in U.S. honey bee populations. The initial bottleneck was the modest subset of honey bee subspecies imported into the U.S. Of the estimated 30 *A. mellifera* subspecies, only nine were introduced into the U.S. (Sheppard, 1989a, 1989b). Additionally, the U.S. commercial breeding industry is responsible for further bottlenecking managed honey bee genetics by prioritizing three ‘C’ lineage subspecies, *A. m. ligustica*, *A. m. carnica*, and *A. m. caucasica*, in breeding programs and largely abandoning the other introduced subspecies (Watkins, 1968). The C lineage honey bees displayed docile behaviors, produced greater honey yields, and were less prone to swarming and absconding compared to the other lineages; thus, they were considered to be more desirable to beekeepers (Delaney et al., 2009; Watkins, 1968). Thirdly, importations of honey bees, and later their germplasm, were halted by the Honey Bee Act of 1922 in an effort to reduce introductions of harmful pathogens and invasive species (Cobey et al., 2012; Crane, 1999). Subsequently, less genetics were incorporated into the U.S. honey bee populations due to the importation restrictions. Lastly, the arrival of the *varroa* mite, *Varroa destructor* Anderson & Trueman to the U.S., reduced honey bee populations across the U.S., particularly decimating feral honey bee populations (Kraus & Page Jr, 1995; Oldroyd, 2007; Wenner & Bushing, 1996).

Parasites and pathogens (henceforth, “pathogens”) are among the direct factors affecting honey bee health, often causing harmful physiological and behavioral changes (Evans & Schwarz, 2011; Gómez-Moracho et al., 2017). Known pathogens such as the parasitic mite *Varroa destructor* or the microsporidian pathogens *Nosema apis* Zander and *N. ceranae* Fries are among the identified and widespread threats to honey bee health. However, newly identified and lesser-studied pathogens such as trypanosomes and bacterial *spiroplasmas* are emerging as underlying threats to bee health (Fünfhaus et al., 2018; Williams et al., 2019).
Varroa destructor is an ectoparasitic mite that feeds on the fat body of honey bees, increasing the risk of winter mortality (Ramsey et al., 2019). Additionally, the mite is known to vector harmful honey bee viruses, such as deformed wing virus and acute bee paralysis virus (Martin, 2001; Posada-Florez et al., 2020). Nosema apis and N. ceranae are microsporidian gut pathogens associated with digestive disorders, decreased colony productivity, increased winter mortality, and reduced overall colony health (Botías et al., 2013; Rangel et al., 2016). Nosema ceranae is the more virulent species due to its switching hosts from the Asian honey bee, A. cerana F. (Fries, 2010). Nosema ceranae is currently the more commonly detected and widely distributed Nosema species in the U.S. (Chen et al., 2008; Klee et al., 2007; Traver & Fell, 2011).

Crithidia mellifica Langridge and McGhee and Lotmaria passim Schwarz are two obligate pathogenic trypanosome species commonly found in honey bee hindguts. Lotmaria passim has been reported as the more prevalent of the two species in the U.S. (Schwarz et al., 2015; Williams et al., 2019). These trypanosomes are poorly understood in terms of transmission and how they affect honey bees but have been linked to altered honey bee immunity and lifespan (Ravoet et al., 2013; Runckel et al., 2011; Strobl et al., 2019). Two bacterial Spiroplasma species, Spiroplasma apis Mouches and S. melliferum Clark are known to invade the gut lumen and hemolymph of honey bees (Clark, 1977; Clark et al., 1985; Mouches et al., 1982, 1983). Mouches et al. (1982) found that mortality occurred when S. apis was injected into adult honey bees or fed on by honey bees under lab conditions.

Researchers in South America (Moretto et al., 1991; Moretto & Mello Jr., 1999) and Mexico (Vandame et al., 2002) have shown that susceptibility to pathogens can vary between honey bee subspecies, specifically between A lineage and C lineage subspecies. Varroa
*destructor* infestation levels have remained consistently low in A lineage honey bee colonies in Brazil for several decades (De Jong, 1996). Furthermore, Africanized colonies had lower *varroa* mite infestation (5.78 mites/100 bees) compared to European hybrid colonies (7.53 mites/100 bees) (Moretto et al., 1991). The aggressive grooming behavior observed in A lineage honey bees is thought to serve as the primary mechanism for reducing *V. destructor* development and success (Moretto et al., 1993). While C lineage honey bees are generally more susceptible to *varroa* infestation, isolated feral populations have shown tolerance and resistance (Locke et al., 2012; Seeley, 2007). *Nosema* infection is inconsistent among studies, some indicating an association with specific genetic lineages and others showing no significant association (Herrera et al., 2017; Szalanski et al., 2014).

Several studies suggest that feral honey bee colonies have relatively low occurrences of *V. destructor* and *Nosema* sp. compared to managed colonies (Rangel et al., 2016; Seeley, 2007; Seeley et al., 2015; Szalanski et al., 2014). Additionally, studies have found elevated infection rates of *N. ceranae* and *L. passim* in commercial honey bee colonies compared to stationary colonies (Williams et al., 2019; Zhu et al., 2014). There is also evidence that feral honey bee colonies can persist for years despite *varroa* mite presence, indicating greater resistance to *V. destructor* (Seeley, 2007). While feral honey bee genetics are of interest to diversify commercial honey bee stock and improve colony health, there is concern over introducing aggressive Africanized traits (Herrera et al., 2017; Mattila & Seeley, 2007).

For this study, novel and previously collected data from pathogen (Cleary, 2017; Szalanski et al., 2014; Williams et al., 2019) and mtDNA genetic diversity (Cleary et al., 2018; Magnus et al., 2014; Szalanski et al., 2016) studies were combined and analyzed to explore relationships between pathogen infection and genetics, management, and location. The
objectives of this study were to compare pathogen prevalence (N. apis, N. ceranae, C. melliferae, L. passim, S. apis, S. melliferum, and V. destructor) among honey bee genetic lineages (A, C, M, O), management sources (managed, swarm, feral), and regional locations (SC U.S., HI, UT). Additionally, colony and swarm samples were haplotyped to characterize mtDNA genetic diversity from south-central U.S., Utah, and Hawaii.

**Material and Methods**

**Sample Collection and Varroa Detection**

Worker honey bees were collected from 2005-2020 from managed colonies, feral colonies, and swarms located in the south-central U.S. (SC), including Arkansas, southern Missouri, and Oklahoma; Hawaii; and Utah. Samples were stored in 70% ethanol at room temperature, and voucher specimens are maintained at the Insect Genetic Lab, University of Arkansas, Fayetteville, AR, USA.

The honey bee samples underwent a varroa mite check using a varroa mite wash adapted from Oliver (2013). Samples were agitated in the mite wash for 30 seconds, dislodging mites to the opposite side of the mesh. Varroa mite counts were databased, and all mites were labeled and stored in ethanol-filled Eppendorf tubes. Additionally, counts of the number of honey bees in each sample were performed, and bees were inspected to ensure mites had dislodged. The number of mites was divided by the number of honey bees in each colony sample and multiplied by 100 to determine the number of mites per 100 bees.

**DNA Extraction and Sequencing**

Novel and previously collected data from pathogen (Cleary, 2017; Szalanski et al., 2014; Williams et al., 2019) and mitochondrial DNA (mtDNA) genetic diversity (Cleary et al., 2018;
Magnus et al., 2014; Szalanski et al., 2016) studies were combined and analyzed to explore relationships between pathogen prevalence and genetic origin, management, and location.

For samples collected during this study, DNA was extracted from honey bees per Cleary et al., 2018. DNA extractions followed a salting-out protocol with in-house reagents per Sambrook and Russell (2001). DNA samples were stored in a -20°C freezer. PCR reactions included 2 μL of extracted DNA along with the PCR master mix: 1 μL of each primer, 5 μL 10x PCR buffer, 4 μL dNTP, 0.4 μL Taq, and 40 μL of PCR water for a final volume of 50 μL per sample. PCR was performed as described in (Szalanski, 2000) using honey bee mtDNA COI-COII PCR primers E2 and H2 (Garnery et al., 1998) and the following thermocycler conditions: denatured initially for 5 min at 94°C, then 40 cycles at 94°C for 45 seconds, 46°C for 1 min, 72°C for 1 min and a final extension of 72°C for 5 min (Table 4.1) (Taylor et al., 1997). Detection of PCR amplicons for this and all subsequent molecular diagnostic assays was done by subjecting PCR products to electrophoresis on a 2% agarose gel and visualized using a BioDoc-it™ Imaging System (UVP, Inc., Upland, CA).

Following successful DNA extraction, PCR products were purified and concentrated using VWR centrifugal devices (VWR, Radnor, PA). Forward and reverse primers E2 and H2 were combined with the purified DNA and sent to Eurofins Genomics (Diatherix, Huntsville, AL) for direct sequencing in both directions. Consensus sequences with the primer ends removed were obtained using Geneious v6.1.8 (Biomatters Ltd., Auckland, New Zealand), and unique haplotypes were identified by eye. Haplotypes were assigned and compared using an NIH BLAST search of DNA sequences available on NCBI GenBank (National Center for Biotechnology Information) and the University of Arkansas Insect Genetics lab’s database (ALS unpublished data).
Molecular Diagnostics

Nosema

Samples were tested for Nosema sp. using the DNA extraction product and primers NosemaSSU-1F and NosemaSSU-1R under conditions: 2 minutes at 94°C, then 40 cycles of 94°C for 45 seconds, 50°C for 1 minute, and 72°C for 1 minute, followed by a final extension of 72°C for 5 minutes (Table 4.1) (Szalanski et al., 2014). These PCR primers were developed by Szalanski et al. (2014) and amplify a 222 bp amplicon for N. apis and a 237 bp amplicon for N. ceranae using the small subunit gene region specific for Nosema.

Samples positive for Nosema sp. underwent PCR-RFLP to determine whether the sample was positive for N. apis, N. ceranae, or both. The RFLP digestion utilizes restriction enzymes Dra I, cutting only N. ceranae at 79 bp, and Rsa I, only cutting N. apis at 130 bp. Samples were incubated overnight at 37°C.

Trypanosomes

Samples were tested for trypanosomes using a multiplex PCR with the primer CBSSU rRNA F2/B4 (Schmid-Hempel & Tognazzo, 2010) to identify any trypanosomatid species and L. passim 18S-F, which identifies L. passim (Szalanski et al., 2016) using the thermocycler program conditions: denaturing step of 5 minutes at 95°C; followed by 40 cycles of 30 seconds at 95°C; primer annealing for 30 seconds at 57°C (Table 4.1) (Schmid-Hempel & Tognazzo, 2010). A negative control (DNA extraction replaced with water) and positive controls for C. mellificae (30254 ATCC) and L. passim (PRA-422 ATCC) were obtained from the American Type Culture Collection (ATCC, Manassas, VA). These primers result in a 608 bp product for samples positive for all trypanosomatids and a 499 bp product for those positive for only L. passim.

Spiroplasma
Multiplex PCR using primers S.apis ITS-F, S.apis ITS-R, Ms-160 F, and Ms-160-R were used to detect *S. apis* and *S. melliferum* (Table 4.1). The S.apis ITS primers yield a 190 bp amplicon from the 3’ end of 16S rRNA to the ITS-1 region, while the Ms-160 primers target a spiralin-like gene of *S. melliferum*. The PCR conditions were: 2 minutes at 94°C; then 39 cycles of 94°C for 45 seconds; 59°C for 1 minute; and 72°C for 1 minute, followed by a final extension of 72°C for 5 minutes (Schwarz et al. 2014). Positive controls were acquired from type strain from the American Type Culture Collection (ATCC, Manassas, VA) for *S. apis* (33834 ATCC) and *S. melliferum* (33219 ATCC). Samples positive for *S. apis* result in a 190 bp amplicon, while samples positive for *S. melliferum* result in a 160 bp amplicon.

**Statistical Analysis**

All statistical analysis was performed using R v.4.0.5, unless otherwise stated, with an alpha level of 0.05 (R Core Team, 2021). A one-tailed Fisher’s exact test with pairwise post hoc comparisons and Holm-Bonferroni correction in R v.4.0.5 with the package *rstatix* was used to determine if there was a significant association between lineage and pathogen infection for each pathogen species (Kassambara, 2021; R Core Team, 2021). A three-sample test for equality of proportions with pairwise post hoc comparisons and Holm-Bonferroni corrections was used to analyze the relationship between sample location compared to pathogen infection for each pathogen species (Ebbert, 2019; R Core Team, 2021). The above analysis was used to analyze the relationship between pathogen infection and management source (managed, swarm, feral) for each pathogen species, where cell size dictated the test used (Fisher’s for comparisons with ≤5 observations in a cell and proportion test for comparisons with cell sizes >5). When significant, analysis indicates an association between pathogen infection and the predictive variable; pairwise analysis indicates significance between factor levels.
The average number of *varroa* mites per 100 adult bees was calculated and compared between lineages for statistical significance using a Kruskal-Wallis rank sum test in JMP Pro 16.0.0. All descriptive statistics are given as the mean ± 1 SE. The Kruskal-Wallis test is a non-parametric alternative to the one-way ANOVA and does not share the ANOVA’s distributional assumptions (Conover & Iman, 1981; SAS Institute Inc, 2021). Kruskal-Wallis tests were also conducted to compare significant differences in average *varroa* mite loads between source types and regions. Pairwise post hoc comparisons with Dunn-Bonferroni correction were performed following significant tests to assess differences between factor levels.

**Results**

A total of 766 honey bee samples were subjected to DNA sequencing and molecular diagnostics. *Nosema* sp. was detected in 17.68% of the samples, of which only one sample was identified as *N. apis*, the remainder were *N. ceranae*. The honey bee parasitic trypanosome, *Lotmaria passim*, was found in 14.36% of the samples. *Varroa destructor* was observed in 23.32% of the samples screened (n=415). None of the samples were positive for the trypanosome *C. mellificae* nor bacterial pathogens *S. apis* and *S. melliferum*.

A total of 43 mtDNA COI-COII haplotypes were detected within four genetic lineages (A, C, M, O), with the C lineage detected most frequently (67.36%, n=516), followed by the A (16.32%, n=125), M (12.92%, n=99), and O lineage (3.39%, n=26). A total of nine haplotypes were unique to feral colonies within our study, while seven were unique to managed colonies and two to swarms (*Table 4.2*). Geographically, C1, C2, C11, C33, M3, and O1 occurred in all three regions; unique haplotypes occurred within each of the sampled regions (*Table 4.3*).

The infection rate of all three detected pathogens was lowest in A lineage samples (*Nosema* sp., *L. passim*, & *V. destructor*) (*Table 4.4 & Table 4.5*). Based on the analysis, there
was a significant association between *Nosema* sp. and lineage (p<0.001). *Nosema* sp. infection was significantly lower in the A lineage compared to the C lineage (p<0.001) as well as the M lineage samples (p<0.001); no significant differences were detected among the other lineages (Table 4.6). Within regions, *Nosema* sp. was most common in Hawaii (38.46%) and least common in Utah (5.26%); *Nosema* sp. infection had a significant association with region (p<0.001) (Table 4.4). Our analysis suggests that the proportion of *Nosema* sp. presence was significantly higher in Hawaii compared to Utah ($\chi^2=55.874$, df=1, p<0.001) and the SC U.S. ($\chi^2=69.381$, df=1, p<0.001); however, there was no significant difference in the rate of *Nosema* sp. infection observed between the SC U.S. and Utah ($\chi^2=2.514$, df=1, p=0.06). The proportion of samples with *Nosema* sp. presence was more significant in swarms compared to feral colonies (p<0.001) and significantly greater in managed colony samples compared to feral colony samples (p<0.001). No difference was observed between swarms and managed colonies (p=1).

The infection rate of *L. passim* was highest in the M lineage (22.22%) and lowest in the A lineage samples (Table 4.4); however, *L. passim* infection was not significantly associated with genetic lineage based on our analysis (p=0.067). *Lotmaria passim* did have a significant association with region ($\chi^2=28$, df=2, p<0.001). Pairwise analysis found that *L. passim* infection was higher in Hawaii compared to the SC U.S. (p<0.001) and Utah (p=0.004); no significant difference was detected between the SC U.S. and Utah (p=0.45). The proportion of samples with *L. passim* infection was significantly associated with management source ($\chi^2=18.51$, df=2, p<0.001), with managed colonies having significantly higher infection rates compared to feral colonies (p<0.001). No significant differences in *L. passim* prevalence was detected between swarms and feral colonies (p=0.10), nor managed colonies compared to swarms (p=0.57).
Although the proportion of samples positive for *V. destructor* was not significantly different between lineages (p=0.07), the C lineage samples had the most samples with *V. destructor* present (25.84%); they had the highest average *varroa* mite load (2.28 mites/100 bees) (Table 4.5). Overall, the total average mite load amongst samples was 2.13 mites per 100 bees. *Varroa* loads were not significantly different between lineages based on the Kruskal-Wallis ($\chi^2=5.02$, df=3, p=0.170), indicating the mean rank of *varroa* mite load was similar for each level of lineages. *Varroa* loads were significantly different between management source ($\chi^2=21.22$, df=2, p<0.001) as well as between regions ($\chi^2=69.51$, df=2, p<0.001) (Figure 4.1 & Figure 4.2). The results of the pairwise multiple comparisons indicated significant differences of mite loads between managed colonies and swarms (p=0.003) as well as managed and feral colonies (p<0.001) (Table 4.7). Pairwise comparison results indicated significant differences in mite loads between samples from the SC U.S. and Utah (p<0.001) as well as the SC U.S. and Hawaii (p<0.001) (Table 4.8).

**Discussion**

This study is among the first and most geographically comprehensive studies to characterize genetic origin and pathogen prevalence in managed and unmanaged honey bee populations. Beyond characterizing maternal haplotypes of the samples, our results address three main factors regarding pathogen occurrence in U.S. honey bee populations: 1. Genetic lineage; 2. Management source; and 3. Regional location. All four genetic lineages previously reported in the U.S. were represented in each region and management type sampled. Four of the seven pathogen species screened for were detected, including *V. destructor*, *L. passim*, *N. apis*, and *N. ceranae*. 
Our study demonstrates there are populations of honey bees genetically distinct from commercial queen breeder colonies, suggesting remnant populations of historical importations from the 19th century have persisted for over 100 years (Magnus et al., 2011). Three haplotypes, C1 (A. m. ligustica); C2 (A. m. carnica); and C11 (A. m. ligustica), represented 54% of the samples. *Apis mellifera ligustica* and *A. m. carnica* are the most widely distributed and commonly detected *A. mellifera* subspecies in the U.S., occurring in both managed and unmanaged honey bee populations (Delaney et al., 2009; Magnus et al., 2011, 2014; Rangel et al., 2020). Interestingly, our study identified all four lineages known to have been introduced into the U.S. in all three of the geographic areas sampled.

Based on the distribution of the detected haplotypes across the different regions sampled compared to previous studies in the northern U.S., there is likely a preference amongst haplotypes for specific climates. In the U.S., A lineage honey bees are restricted to the southern states because they have a reduced capacity for winter survival (Schneider et al., 2004). Studies in New York and Pennsylvania, sampling from both managed and unmanaged populations, detected haplotypes exclusively from the cold-hardy M and C lineages (Rangel et al., 2020; Seeley et al., 2015). In contrast, our study, as well as previous research in the southern and southwestern U.S., have found four lineages with multiple haplotypes not detected in those found in the northeastern U.S. (Kono & Kohn, 2015; Magnus & Szalanski, 2009; Szalanski et al., 2016). Additional research should explore genotyping analysis to complement our mtDNA study and investigate traits from ancestral populations, which could be incorporated into breeding stock to increase resistance to harmful disease-causing agents and diversify genetic stock.

Based on our analysis, *Nosema* sp. was the only pathogen species with a statistically significant association with genetic lineage. *Nosema* sp. infection rates were higher in the C and
M lineage samples compared to the A lineage samples. Research is limited pertaining to internal pathogen infection amongst honey bee lineages, especially concerning the M and O lineages; however, M lineage subspecies, *A. m. mellifera*, has displayed increased susceptibility to brood disease compared to C lineage honey bees (Jensen et al., 2009). Quantitative measures of *Nosema* sp. and *L. passim* infection would be beneficial to assess the level of infection between lineages and explore whether certain lineages can persist despite high levels of pathogen infection.

A study performed by Herrera et al. (2017) found similar results to our C lineage samples in Texas honey bee colonies, where *N. ceranae* was detected more frequently in C lineage colonies compared to A lineage colonies. However, because C lineage bees were most common in managed populations, the high level of pathogen infection could also be impacted by management. *Nosema ceranae* has shown to be more prevalent in managed colonies than feral colonies (Gilliam & Taber, 1991; Szalanski et al., 2014) and increase in infection intensity with urbanization (Youngsteadt et al., 2015). Management tactics often treat for harmful pathogen; however, the risk of pathogen transmission increases when contaminated equipment is used between infected and uninfected colonies (Oertel, 1967). Furthermore, studies have detected *Nosema* sp. (Strange et al., 2008) and *L. passim* (Williams, 2018) in packaged bees and commercial queen breeder colonies, indicating these facilities may be spreading pathogens across the country. Specifically, *L. passim* was detected in queen breeder colonies in Hawaii, explaining the high rate of *L. passim* infection (Williams, 2018).

While research is lacking in the U.S., climate and genetic origin have been observed as important factors affecting the development and success of *V. destructor* (De Jong et al., 1984; De Jong & Soares, 1997; Moretto et al., 1991; Moretto & Mello Jr., 1999). Within our study, A
lineage samples had the fewest instances of pathogen infection for the three species detected. Highly expressed defensive and hygienic grooming behaviors have been observed in A lineage honey bees in Brazil and Mexico, resulting in lower mite infestations, particularly when compared to C lineage bees (Moretto et al., 1993; Moretto & Mello Jr., 1999; Vandame et al., 2002). These efficient grooming behaviors could explain our study’s lower prevalence and intensity of *V. destructor* detected in the A lineage bees.

Furthermore, A lineage samples were most common amongst feral colonies in Utah; the low incidence of *varroa* mites could be due to Utah feral colonies being isolated for a long period of time from managed colonies. In comparison, average mite loads were highest in the SC U.S., where samples were primarily from managed colonies. Honey bee colonies within close proximity of each other, as seen in apiaries, are more likely to share pathogens, specifically *Nosema* sp. and *V. destructor*, compared to isolated colonies due to shared floral resources, robbing events, or other interactions (Fürst et al., 2014; Oertel, 1967; Renz & Rosenkranz, 2001; Youngsteadt et al., 2015). Geographic isolation should be further explored as well as infection tolerance to determine if colonies are truly tolerant or if they are simply isolated from pathogens. Additionally, feral colonies have been found to establish in nest cavities much smaller (25-50% less spacious) compared to beekeepers’ hives (Seeley & Morse, 1976). Smaller colonies with fewer brood cells may reduce mite populations by limiting reproductions opportunities.

It is necessary to examine factors associated with increased pathogen susceptibility to mitigate the effects of harmful pathogens in honey bee populations. Future research should revisit colonies and evaluate pathogen tolerance and resistance, specifically in feral populations, to observe whether colonies truly are persisting despite pathogen presence, as Seeley et al. (2007) observed in the New York Arnot Forest. We hope to broaden our survey in the future to
include remote sensing technology to measure hive conditions (colony weight, internal temperature, and humidity) and explore traits (i.e., use of stores during winter, honey production, varroa mite infestation levels, pathogen infection levels using Q-PCR, swarm frequency) associated with lineages.
Acknowledgements

We would like to acknowledge the Arkansas, Oklahoma, Missouri, Hawaii, and Utah beekeepers who agreed to participate in this study. We would also like to thank Jerry Shue and the Utah Department of Agriculture and Food for assisting with sample collection. This research was supported in part by the University of Arkansas, Arkansas Agricultural Experiment Station.
References


Fries, I. (2010). *Nosema ceranae* in European honey bees (*Apis mellifera*). *Journal of Invertebrate Pathology, 103* Suppl 1, S73–79. https://doi.org/10.1016/j.jip.2009.06.017


Kono, Y., & Kohn, J. R. (2015). Range and frequency of Africanized honey bees in California (USA). *PLoS ONE, 10*(9), e0137407. [https://doi.org/10.1371/journal.pone.0137407](https://doi.org/10.1371/journal.pone.0137407)


viruses acquired from an artificial diet. *PLOS ONE*, **15**(11), e0242688. https://doi.org/10.1371/journal.pone.0242688


Renz, M., & Rosenkranz, P. (2001). Infestation dynamics and reinvasion of *Varroa destructor* mites in honey bee colonies kept isolated and in groups. *Apidologie (France)*.


SAS Institute Inc. (2021). *JMP Pro* (v.16.0.0) [Windows].


<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2</td>
<td>F: 5’-GGCAGAATAAGTGCATTG-3’</td>
<td>A</td>
</tr>
<tr>
<td>H2</td>
<td>R: 5’-CAATATC ATTGATGACC-3’</td>
<td>A</td>
</tr>
<tr>
<td>NosemaSSU-1F</td>
<td>F: 5’-ACAATATGTATTAGATCTGATATA-3’</td>
<td>B</td>
</tr>
<tr>
<td>NosemaSSU-1R</td>
<td>R: 5’-TAATGATATGCTTAAG TTCAAG-3’</td>
<td>B</td>
</tr>
<tr>
<td>S. apis ITS-F</td>
<td>F: 5’-AATGCCAGAAGCACGTATCC-3’</td>
<td>C</td>
</tr>
<tr>
<td>S. apis ITS-R</td>
<td>R: 5’-GAACGAGATATACCTAAGCTGTTACAC-3’</td>
<td>C</td>
</tr>
<tr>
<td>Ms-160 F</td>
<td>F: 5’-TTGCA AAAGCTGTTTAGATGC-3’</td>
<td>C</td>
</tr>
<tr>
<td>Ms-160-R</td>
<td>R: 5’-TGACCAGAAATGTTTGCTGAA-3’</td>
<td>C</td>
</tr>
<tr>
<td>CBSSU rRNA F2</td>
<td>F: 5’-CTTTTGACGAACAAACTGCCCTATC-3’</td>
<td>D</td>
</tr>
<tr>
<td>CBSSU rRNA B4</td>
<td>R: 5’-AACCGAACGACTAAACCC-3’</td>
<td>D</td>
</tr>
<tr>
<td>L. passim18S-F</td>
<td>F: 5’-AGGGATATTTAAACCC ATCGAAAATCT-3’</td>
<td>E</td>
</tr>
</tbody>
</table>
Table 4.2. Percent of each *Apis mellifera* COI–COII haplotype within the three source types. Highlighted color indicates what management sources the haplotypes were detected in. A total of 44 haplotypes were detected in this study.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Managed (%)</th>
<th>Swarm (%)</th>
<th>Feral (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2j</td>
<td>2</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C2s</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C15</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C35</td>
<td>2</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>M17a</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>M4c''</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>M6</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C27</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.00</td>
</tr>
<tr>
<td>M2</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.00</td>
</tr>
<tr>
<td>A4a</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A16</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A26</td>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A26c</td>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A29a</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>M3a</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>M4n</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>M5</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>M70</td>
<td>35</td>
<td>88.57</td>
<td>11.43</td>
<td>0.00</td>
</tr>
<tr>
<td>A1d</td>
<td>26</td>
<td>3.85</td>
<td>0.00</td>
<td>96.15</td>
</tr>
<tr>
<td>A26a</td>
<td>15</td>
<td>6.67</td>
<td>0.00</td>
<td>93.33</td>
</tr>
<tr>
<td>C12</td>
<td>14</td>
<td>78.57</td>
<td>0.00</td>
<td>21.43</td>
</tr>
<tr>
<td>C26</td>
<td>2</td>
<td>50.00</td>
<td>0.00</td>
<td>50.00</td>
</tr>
<tr>
<td>C29</td>
<td>2</td>
<td>50.00</td>
<td>0.00</td>
<td>50.00</td>
</tr>
<tr>
<td>C30</td>
<td>6</td>
<td>33.33</td>
<td>0.00</td>
<td>66.67</td>
</tr>
<tr>
<td>C33</td>
<td>6</td>
<td>66.67</td>
<td>0.00</td>
<td>33.33</td>
</tr>
<tr>
<td>M4</td>
<td>2</td>
<td>50.00</td>
<td>0.00</td>
<td>50.00</td>
</tr>
<tr>
<td>A1b</td>
<td>9</td>
<td>0.00</td>
<td>44.44</td>
<td>55.56</td>
</tr>
<tr>
<td>A1e</td>
<td>26</td>
<td>0.00</td>
<td>19.23</td>
<td>80.77</td>
</tr>
<tr>
<td>A4</td>
<td>7</td>
<td>0.00</td>
<td>14.29</td>
<td>85.71</td>
</tr>
<tr>
<td>A4''</td>
<td>9</td>
<td>11.11</td>
<td>11.11</td>
<td>77.78</td>
</tr>
<tr>
<td>C1</td>
<td>239</td>
<td>73.22</td>
<td>8.37</td>
<td>18.41</td>
</tr>
<tr>
<td>C2</td>
<td>80</td>
<td>16.25</td>
<td>11.25</td>
<td>72.50</td>
</tr>
<tr>
<td>C11</td>
<td>94</td>
<td>53.19</td>
<td>6.38</td>
<td>40.43</td>
</tr>
<tr>
<td>C19</td>
<td>27</td>
<td>37.04</td>
<td>18.52</td>
<td>44.44</td>
</tr>
<tr>
<td>C31</td>
<td>39</td>
<td>53.85</td>
<td>5.13</td>
<td>41.03</td>
</tr>
<tr>
<td>M3</td>
<td>9</td>
<td>11.11</td>
<td>33.33</td>
<td>55.56</td>
</tr>
<tr>
<td>M7</td>
<td>46</td>
<td>50.00</td>
<td>30.43</td>
<td>19.57</td>
</tr>
<tr>
<td>O1</td>
<td>7</td>
<td>28.57</td>
<td>28.57</td>
<td>42.86</td>
</tr>
<tr>
<td>O2</td>
<td>7</td>
<td>14.29</td>
<td>14.29</td>
<td>71.43</td>
</tr>
<tr>
<td>O5</td>
<td>8</td>
<td>12.50</td>
<td>12.50</td>
<td>75.00</td>
</tr>
<tr>
<td>O5&quot;b</td>
<td>4</td>
<td>50.00</td>
<td>25.00</td>
<td>25.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>766</strong></td>
<td><strong>47.39</strong></td>
<td><strong>10.57</strong></td>
<td><strong>42.04</strong></td>
</tr>
</tbody>
</table>

Note: Due to rounding error, percentages may not sum to 100%.
Table 4.3. Percent of each *Apis mellifera* COI–COII haplotype within the three regions. Highlighted color indicates which region the haplotypes were detected in. A total of 44 haplotypes were detected in this study.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>SC U.S. (%)</th>
<th>HI (%)</th>
<th>Utah (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2j</td>
<td>2</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C2s</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C15</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C26</td>
<td>2</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C27</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C29</td>
<td>2</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C30</td>
<td>6</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C35</td>
<td>2</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>M2</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>M4</td>
<td>2</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>M5</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>M6</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>M17a</td>
<td>1</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>O5&quot;b</td>
<td>4</td>
<td>100</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>A26d</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.00</td>
</tr>
<tr>
<td>M4c&quot;&quot;&quot;</td>
<td>1</td>
<td>0.00</td>
<td>100</td>
<td>0.00</td>
</tr>
<tr>
<td>M70</td>
<td>35</td>
<td>0.00</td>
<td>100</td>
<td>0.00</td>
</tr>
<tr>
<td>A1b</td>
<td>9</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A1e</td>
<td>26</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A4a</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A16</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A26</td>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A26c</td>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>A29a</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>M3a</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>M4n</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>O2</td>
<td>7</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>C19</td>
<td>27</td>
<td>70.37</td>
<td>29.63</td>
<td>0.00</td>
</tr>
<tr>
<td>A1</td>
<td>23</td>
<td>60.87</td>
<td>0.00</td>
<td>39.13</td>
</tr>
<tr>
<td>A1d</td>
<td>26</td>
<td>88.46</td>
<td>0.00</td>
<td>11.54</td>
</tr>
<tr>
<td>A26a</td>
<td>15</td>
<td>40.00</td>
<td>0.00</td>
<td>60.00</td>
</tr>
<tr>
<td>A4</td>
<td>7</td>
<td>28.57</td>
<td>0.00</td>
<td>71.43</td>
</tr>
<tr>
<td>A4””</td>
<td>9</td>
<td>11.11</td>
<td>0.00</td>
<td>88.89</td>
</tr>
<tr>
<td>C12</td>
<td>14</td>
<td>85.71</td>
<td>0.00</td>
<td>14.29</td>
</tr>
<tr>
<td>C31</td>
<td>39</td>
<td>87.18</td>
<td>0.00</td>
<td>12.82</td>
</tr>
<tr>
<td>O5</td>
<td>8</td>
<td>75.00</td>
<td>0.00</td>
<td>25.00</td>
</tr>
<tr>
<td>M7</td>
<td>46</td>
<td>0.00</td>
<td>97.83</td>
<td>2.17</td>
</tr>
<tr>
<td>C1</td>
<td>239</td>
<td>36.40</td>
<td>50.63</td>
<td>12.97</td>
</tr>
<tr>
<td>C2</td>
<td>80</td>
<td>87.50</td>
<td>6.25</td>
<td>6.25</td>
</tr>
<tr>
<td>C11</td>
<td>94</td>
<td>81.91</td>
<td>2.13</td>
<td>15.96</td>
</tr>
<tr>
<td>C33</td>
<td>6</td>
<td>50.00</td>
<td>33.33</td>
<td>16.67</td>
</tr>
<tr>
<td>M3</td>
<td>9</td>
<td>77.78</td>
<td>11.11</td>
<td>11.11</td>
</tr>
<tr>
<td>O1</td>
<td>7</td>
<td>42.86</td>
<td>28.57</td>
<td>28.57</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>766</td>
<td>49.61</td>
<td>30.55</td>
<td>19.84</td>
</tr>
</tbody>
</table>

Note: Due to rounding error, percentages may not sum to 100%.
Table 4.4. Percent of samples positive for *Nosema* sp. and *Lotmaria passim* within lineages, regions, and management source.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>n</th>
<th>Nosema sp. (%)</th>
<th>L. passim (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>125</td>
<td>4.80</td>
<td>9.60</td>
</tr>
<tr>
<td>C</td>
<td>516</td>
<td>19.19</td>
<td>14.15</td>
</tr>
<tr>
<td>M</td>
<td>99</td>
<td>29.29</td>
<td>22.22</td>
</tr>
<tr>
<td>O</td>
<td>26</td>
<td>7.69</td>
<td>11.54</td>
</tr>
<tr>
<td>Total</td>
<td>766</td>
<td>17.75</td>
<td>14.36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>n</th>
<th>Nosema sp. (%)</th>
<th>L. passim (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC</td>
<td>380</td>
<td>10.00</td>
<td>9.21</td>
</tr>
<tr>
<td>UT</td>
<td>152</td>
<td>5.26</td>
<td>11.84</td>
</tr>
<tr>
<td>HI</td>
<td>234</td>
<td>38.46</td>
<td>24.36</td>
</tr>
<tr>
<td>Total</td>
<td>766</td>
<td>17.75</td>
<td>14.36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>n</th>
<th>Nosema sp. (%)</th>
<th>L. passim (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>322</td>
<td>8.39</td>
<td>8.07</td>
</tr>
<tr>
<td>M</td>
<td>363</td>
<td>24.52</td>
<td>19.56</td>
</tr>
<tr>
<td>S</td>
<td>81</td>
<td>24.69</td>
<td>16.05</td>
</tr>
<tr>
<td>Total</td>
<td>766</td>
<td>17.75</td>
<td>14.36</td>
</tr>
</tbody>
</table>

Note: Due to rounding error, percentages may not sum to 100%.

Table 4.5. Percent of honey bee samples with *Varroa destructor* presence and average number of *V. destructor* detected per 100 honey bees from each genetic lineage. Sampled out of 416 colonies.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>n</th>
<th>Percent Positive (%)</th>
<th>Avg. No. Mites/100 Bees</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>26</td>
<td>7.41</td>
<td>0.93</td>
</tr>
<tr>
<td>C</td>
<td>329</td>
<td>25.84</td>
<td>2.28</td>
</tr>
<tr>
<td>M</td>
<td>55</td>
<td>16.36</td>
<td>1.89</td>
</tr>
<tr>
<td>O</td>
<td>5</td>
<td>20.00</td>
<td>1.25</td>
</tr>
<tr>
<td>Total</td>
<td>415</td>
<td>23.32</td>
<td>2.13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>n</th>
<th>Percent Positive (%)</th>
<th>Avg. No. Mites/100 Bees</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC</td>
<td>190</td>
<td>43.16</td>
<td>3.48</td>
</tr>
<tr>
<td>UT</td>
<td>51</td>
<td>4.57</td>
<td>2.06</td>
</tr>
<tr>
<td>HI</td>
<td>175</td>
<td>13.73</td>
<td>0.69</td>
</tr>
<tr>
<td>Total</td>
<td>415</td>
<td>23.32</td>
<td>2.13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>n</th>
<th>Percent Positive (%)</th>
<th>Avg. No. Mites/100 Bees</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>80</td>
<td>8.75</td>
<td>0.91</td>
</tr>
<tr>
<td>M</td>
<td>303</td>
<td>29.37</td>
<td>2.66</td>
</tr>
<tr>
<td>S</td>
<td>33</td>
<td>3.03</td>
<td>0.30</td>
</tr>
<tr>
<td>Total</td>
<td>415</td>
<td>23.32</td>
<td>2.13</td>
</tr>
</tbody>
</table>

Note: Due to rounding error, percentages may not sum to 100%.
Table 4.6. Pairwise comparison of *Nosema* sp. infection rate between genetic lineages. Significance indicates an association.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>n</th>
<th>p-value</th>
<th>Adj. p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-M</td>
<td>615</td>
<td>0.03</td>
<td>0.09</td>
</tr>
<tr>
<td>C-O</td>
<td>542</td>
<td>0.20</td>
<td>0.39</td>
</tr>
<tr>
<td>C-A</td>
<td>641</td>
<td>&lt;0.001</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>M-O</td>
<td>125</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>M-A</td>
<td>224</td>
<td>&lt;0.001</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>O-A</td>
<td>151</td>
<td>0.63</td>
<td>0.63</td>
</tr>
</tbody>
</table>

*Significance level of α=0.05.

Table 4.7. Pairwise comparison of mean rank scores of *varroa* loads between management source types.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Score Mean Difference</th>
<th>Std. Error</th>
<th>Z</th>
<th>p-value</th>
<th>Adj. p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-S</td>
<td>53.67</td>
<td>16.31</td>
<td>3.29</td>
<td>&lt;0.001</td>
<td>0.003*</td>
</tr>
<tr>
<td>M-F</td>
<td>40.97</td>
<td>11.19</td>
<td>3.66</td>
<td>&lt;0.001</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>F-S</td>
<td>12.69</td>
<td>18.40</td>
<td>0.69</td>
<td>0.49</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Significance level of α=0.05.

Table 4.8. Pairwise comparison of mean rank scores of *varroa* loads between regions.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Score Mean Difference</th>
<th>Std. Error</th>
<th>Z</th>
<th>p-value</th>
<th>Adj. p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC-HI</td>
<td>76.66</td>
<td>9.33</td>
<td>8.21</td>
<td>&lt;0.001</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>SC-UT</td>
<td>55.75</td>
<td>14.03</td>
<td>3.97</td>
<td>&lt;0.001</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>UT-HI</td>
<td>20.89</td>
<td>14.17</td>
<td>1.47</td>
<td>0.14</td>
<td>0.42</td>
</tr>
</tbody>
</table>

*Significance level of α=0.05.
Figure 4.1. Average number of *varroa* mites per 100 bees by region, ± standard error.

Figure 4.2. Average number of *varroa* mites per 100 bees by management source, ± standard error.
Appendix 4.1. Average number of mites per 100 bees by genetic lineage, ± standard error.

Table 4.2. Pairwise comparison of *Lotmaria passim* infection rate between management source types. Significance indicates an association.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>n</th>
<th>p-value</th>
<th>Adj. p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-S</td>
<td>444</td>
<td>0.811</td>
<td>1</td>
</tr>
<tr>
<td>M-F</td>
<td>685</td>
<td>&lt;0.001</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>F-S</td>
<td>403</td>
<td>0.030</td>
<td>0.089</td>
</tr>
</tbody>
</table>
Appendix 4.3. Pairwise comparisons for the odds ratios (OR) of lineage 1’s being infected with *Nosema* sp. compared to lineage 2. Odds ratios (OR) greater than 1 indicate that the odds of *Nosema* infection among lineage 1 are greater than the odds of *Nosema* sp. infection in lineage 2.

<table>
<thead>
<tr>
<th>Lineage 1</th>
<th>Lineage 2</th>
<th>OR</th>
<th>$\chi^2$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>O</td>
<td>0.55</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>A</td>
<td>M</td>
<td>0.11</td>
<td>21.13</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>A</td>
<td>C</td>
<td>0.22</td>
<td>12.21</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>2.53</td>
<td>1.54</td>
<td>0.2151</td>
</tr>
<tr>
<td>C</td>
<td>M</td>
<td>0.51</td>
<td>6.99</td>
<td>0.008*</td>
</tr>
<tr>
<td>C</td>
<td>A</td>
<td>4.55</td>
<td>12.21</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>M</td>
<td>O</td>
<td>4.91</td>
<td>4.25</td>
<td>0.039*</td>
</tr>
<tr>
<td>M</td>
<td>C</td>
<td>1.94</td>
<td>6.99</td>
<td>0.008*</td>
</tr>
<tr>
<td>M</td>
<td>A</td>
<td>8.85</td>
<td>21.13</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>O</td>
<td>M</td>
<td>0.20</td>
<td>4.25</td>
<td>0.039*</td>
</tr>
<tr>
<td>O</td>
<td>C</td>
<td>0.40</td>
<td>1.54</td>
<td>0.215</td>
</tr>
<tr>
<td>O</td>
<td>A</td>
<td>1.80</td>
<td>0.48</td>
<td>0.487</td>
</tr>
</tbody>
</table>
Chapter 5: Preliminary survey of cockroaches in honey bee colonies: Species identification and pathogen screening

Abstract

Honey bees are negatively impacted by several parasite, pathogen, and pest species. Cockroaches are commonly found in honey bee colonies, particularly abandoned and weakened colonies, and are known to feed on dead insects. While cockroaches are generally not considered a direct pest to honey bees, little is known about which species occur in colonies and whether they can vector harmful honey bee pathogens. Our study is the first to identify cockroaches found in U.S. honey bee colonies to species. Arkansas honey bee colonies were sampled for cockroaches; using taxonomic keys, samples were identified as the Pennsylvania wood roach, Parcoblatta pennsylvanica. Molecular diagnostics identified the samples as Parcoblatta sp. Parcoblatta sp. gut extractions were screened for the presence of honey bee DNA using PCR primers specific for Hymenoptera. Several samples were positive, indicating that wood roaches found in honey bee colonies may feed on deceased honey bees. If honey bees are infected with pathogens in a colony, roaches may feed on dead infected bees and spread the pathogen amongst adjacent colonies. DNA samples were also screened for the honey bee pathogen Nosema ceranae and parasitic trypanosome Lotmaria passim, none were found positive. This study illustrates the need for future studies to investigate the potential impacts of cockroaches in honey bee colonies.
Introduction

Honey bees, *Apis mellifera* L., in the U.S. have suffered population fluctuations and health declines over the last several decades (Potts et al., 2010; vanEngelsdorp & Meixner, 2010). Among the direct stressors impacting honey bee health are pests, parasites, and pathogens (Evans & Schwarz, 2011; Forfert et al., 2015). While several harmful disease-causing agents have been identified in honey bee colonies, lesser-studied species are of interest.

Honey bee colonies can be home to numerous organisms, including insects, mites, rodents, and reptiles (Caron, 1997). The small hive beetle, *Aethina tumida* Murray, and *varroa* mite, *Varroa destructor* Anderson and Truman, are harmful parasites known to cause substantial harm and often considered ubiquitous in honey bee colonies (Ellis et al., 2002; Ellis & Munn, 2005). *Aethina tumida* is a known vector of *Paenibacillus* larvae which is responsible for American foulbrood (Bailey & Ball, 1991; Schäfer et al., 2010). Additionally, *Lotmaria passim* Schwarz, *Crithidia mellificae* Langridge and McGhee, *Nosema ceranae* Fries, *Apis mellifera* filamentous virus (AmFV), and Deformed wing virus (DWF) have been detected in *A. tumida* suggesting vector potential (Cilia et al., 2018; de Landa et al., 2021; Eyer et al., 2009; Nanetti et al., 2021). *Varroa destructor* is a known vector of several honey bee viruses, including the DWV complex and Acute bee paralysis virus (ABPV) complex (Di Prisco et al., 2011; Kumar et al., 2022; Martin, 2001; Posada-Florez et al., 2020; Ryabov et al., 2019). Bacterial pathogens and *L. passim* have also been isolated from *V. destructor* (De Rycke et al., 2002; Hubert et al., 2015; Quintana et al., 2021).

Several other organisms are commonly detected in honey bee colonies that act as occasional pests (Bailey & Ball, 1991), one of which is the cockroach (Order: Blattodea). Cockroaches have been observed as minor pests of honey bee colonies in the U.S., occurring
abundantly in weak colonies, between the inner and outer hive covers, where they have been observed to feed on dead bees, pollen, honey, and comb (Argo, 1874; Caron, 1997; Haydak, 1963). Furthermore, cockroaches have also been reported in honey bee colonies in Ireland (Truman, 1905), South America (Stejskal, 1955), and Australia (Beuhne, 1911). Cockroaches are less common in healthy honey bee colonies with a large population of workers, and the damage they inflict is considered negligible (Caron, 1997).

Despite their pervasiveness in honey bee colonies, virtually no recent research has been conducted to explore which species of cockroaches occur in honey bee colonies or what implications they may have on honey bee health (de Landa et al., 2021; Neumann & Elzen, 2004; Rosenkranz et al., 2010). Specifically, the role of cockroaches as a potential biological vector should be considered. A 1955 study conducted by Stejskal (1955) explored associated insects found in honey bee colonies with gregarine infection as potential biological vectors, several of which were cockroach species. The study found cephaeline gregarines, along with wax, pollen, and dead bees in dissected cockroach guts, deeming cockroaches responsible for introducing the gregarines into honey bee colonies via their feces (Stejskal 1955).

Cockroaches are adaptable insects capable of thriving in various environments where food, moisture, and warmth are available (Baumholtz et al., 1997; Ifeanyi & Olawumi, 2015). Honey bee colonies provide this desirable environment: dark and warm with an adequate food source and protection from potential predators (Argo, 1874; Stejskal, 1955). Additionally, several cockroach species have shown to be partially herbivorous, feeding on pollen, nectar, and flowers, suggesting that they may be attracted to pollen stores in honey bee colonies (Bell et al., 2007). There have been reports of two genera of cockroaches occurring in and around honey bee
colonies, *Periplaneta* (Family: Blattidae) and *Parcoblatta* (Family: Ectobiidae) (Bell et al., 2007; Pokhrel, 2009).

*Periplaneta* is a cockroach genus composed of nearly 50 species, none are endemic to the U.S. Four species have been introduced into the U.S.: *Periplaneta americana* L. which is cosmopolitan in its distribution (Bell & Adiyodi, 1982); *P. australasiae* F., which is common in southeastern coastal U.S. (Barcay, 2011); *P. fuliginosa* Serville, which is common in the Southern U.S. (Appel & Rust, 1985); and *P. brunnea* Burmeister, which is established in the Southern U.S. and California (Bell & Adiyodi, 1982; Edmunds, 1957; Rehn, 1945). Members of *Periplaneta* are often considered household pests, although they reside outdoors when conditions are favorable. In particular, *P. australasiae* and *P. fuliginosa* are commonly found in woodpiles, under bark, and leaf litter; and typically enter homes only when outdoor conditions are unfavorable (Appel & Rust, 1985). Cockroaches, presumably *P. fuliginosa*, have been found frequently in honey bee hives in North Carolina; however, species identification remains unconfirmed and warrants further investigation into their role in hive sanitation (Bell et al., 2007). Also, *P. americana* is considered to be a serious year-round predatory pest of honey bee colonies in Nepal (Pokhrel, 2009).

The genus *Parcoblatta* includes several species of woodland cockroaches native to the U.S. *Parcoblatta* species are rarely considered domestic pests as they naturally inhabit outdoor areas with occasional home invasions. They feed on decaying organic matter, sap, mushrooms, and living and dead insects. *Parcoblatta* sp. are commonly found under organic matter, including rotting logs, loose tree bark, or leaf litter; additionally, they have been found in abandoned honey bee hives (Bell et al., 2007; Blatchley, 1920; Tvedten, 2007). Interestingly, MacDonald & Matthews (1983) suggest that nymphal *Parcoblatta* benefit the colony cycle of *Vespula*

118
squamosa Drury by scavenging debris in the colony and suppressing fungal and protozoan populations. Regarding dispersion, Parcoblatta males can fly short distances (30m) (Cornwell, 1968). All females within the Parcoblatta genus in the U.S. are wingless apart from *P. caudelli* Hebard, which possess wings and can fly (Lawson, 1967).

Twelve species of Parcoblatta occur in the U.S. (Hebard 1917, Pratt 1988). These species include *P. bolliana* Saussure and Zehnter and *P. divsa* Saussure and Zehntner, which are widely occurring across the U.S.; *P. caudelli*, *P. fulvescens* Saussure and Zehntner, *P. lata* Brunner, *P. pensylvanica* DeGeer, *P. uhleriana* Sassure, *P. virginica* Brunner, and *P. zebra* Hebard, which occur across the Eastern and southeastern U.S.; *P. americana* Scudder occurring in the Western U.S.; *P. desertae* Rehn and Hebard occurring in Texas; and *P. notha* Rehn and Hebard occurring in Arizona (Atkinson et al., 1990; Beccaloni, 2019; Hebard, 1917; Pratt, 1988).

As scavengers, cockroaches are often found to feed on human and animal feces, sewage, and garbage (Bell et al., 2007; Fotedar et al., 1991). Cockroaches are known and presumed indirect mechanical vectors and disease reservoirs of numerous pathogens, including bacteria, protozoa, and viruses (Baumholtz et al., 1997; Fotedar et al., 1989, 1991; Kasprzak & Majewska, 1981; Majewska, 1986). When cockroaches contact contaminated surfaces or food sources they may spread pathogens to new areas with their body, feces, saliva, and vomit (Ash & Greenberg, 1980; Fotedar et al., 1989; Fotedar & Banerjee, 1993; Zurek & Schal, 2004).

Cockroaches have been proven to cause food contamination, allergic skin reactions, and asthmatic reactions in humans (Baumholtz et al., 1997). While there has been research examining cockroaches and their effects on humans, little research has been done looking at their role within honey bee colonies (Baumholtz et al., 1997; Horn & Hanula, 2002). If cockroaches can host honey bee pathogens and spread them via feces, as suggested by Stejskal (1955), an
infection could quickly spread within a colony. Honey bees are known to remove foreign material, allo-groom, and feed colony mates; these close intra-colony interactions may facilitate the rapid spread of pathogens within a colony (Chen et al., 2008; Evans & Spivak, 2010; Rothenbuhler, 1964; Schmid-Hempel, 2017).

Two microsporidian pathogens are known to infect honey bees, *N. ceranae* and *N. apis* Zander, via ingestion of spores (Fries, 2010; Smith, 2012; Webster et al., 2004). Over the past decade, *N. ceranae* has surpassed, and likely displaced, *N. apis* as the dominant *Nosema* species in U.S. honey bees (Chen et al., 2008; Fries, 2010; Smith, 2012). Both *Nosema* species are known to attack the midgut of adult honey bees; however, *N. ceranae* has been documented to be more virulent and detrimental to *A. mellifera* than *N. apis*. (Huang et al., 2015; Mayack & Naug, 2009; vanEngelsdorp & Meixner, 2010). *Nosema ceranae* often results in reduced colony productivity, suppressed immunity, and overall weakening of colonies, potentially leading to colony mortality (Martín-Hernández et al., 2009; Taric et al., 2020). Additionally, *N. ceranae* DNA has been identified in *varroa* mite and small hive beetle samples, suggesting vector potential (Cilia et al., 2018; Uroš et al., 2014).

Two species of obligate Trypanosomatid parasites, *C. mellificae* and *L. passim*, are known to colonize the rectum of honey bees (Schwarz et al., 2015). Both species have been detected in the U.S. While transmission is not well understood, the trypanosomes have shown to influence colony health by altering honey bee immunity, physiology, behavior, and lifespan (Ravoet et al., 2013; Runckel et al., 2011; Schwarz et al., 2015; Strobl et al., 2019; Williams et al., 2019; Xu et al., 2018). *Lotmaria passim* DNA has been detected in *V. destructor* and their hemolymph, suggesting passive transport of the pathogen or vectoring potential (Quintana et al.,
2021). With little knowledge of how trypanosomes are transmitted to honey bees, exploration is needed into potential vectors.

While cockroaches are known to occur in honey bee colonies, the species and their role within colonies are largely unknown. This study aims to identify cockroach species found in Arkansas honey bee colonies and determine if the cockroaches are feeding on dead honey bees or carrying honey bee pathogens. This information allows for a better understanding of honey bee pathogen transmission and provides a basis for understanding the role of cockroaches in honey bee colonies.

**Materials and Methods**

**Collection**

Adult and nymphal cockroaches were collected from active and abandoned honey bee colonies located in Washington County, Arkansas, USA. Ten cockroach samplings occurred in Farmington, AR and 13 samplings in Fayetteville, AR. Cockroaches were collected from honey bee colonies in 2018 in May (n=1) and November (n=1), as well as May (n=1), September(n=6), and October (n=1) of 2019. Fayetteville sampling occurred in September of 2019 (n=13). Samples were stored in 70% ethanol at room temperature, and voucher specimens are maintained in the Insect Genetic Lab, University of Arkansas, Fayetteville, AR, USA.

**Identification**

Cockroaches were sorted by life stage and sex and identified to genus, or if possible, species level. Styli and paired cerci are present on adult male cockroaches and absent from nymphs and females. Female cockroaches and nymphs could not definitively be identified to species level using morphological features. Body measurements were taken from adult male cockroaches. Distinct morphological characteristics and geographic location were used to
distinguish roaches using Hebard (1917), Bell (1981), Blatchley (1920), and Peck & Beninger (1989) taxonomic keys. Abdominal and pronotal features, tegmina length, and, in limited cases, size and coloration, were the primary characteristics used to distinguish cockroaches. Cockroach samples were photographed prior to DNA extraction.

DNA was extracted from individual cockroach abdomens and dissected alimentary canals following a salting-out protocol with in-house reagents per Sambrook & Russell (2001). This consisted of combining and homogenizing the cockroach tissue with 300µL of cell lysis solution in 1.5 ml Eppendorf tubes. Samples were stored in a -80°C freezer for at least one hour, followed by a 5-minute incubation in an 80°C water bath. Each sample received 100 µL of protein precipitation solution and was centrifuged at 13.2 X 1000 rpm for 3 minutes. After, 300 µL of the supernatant and 300 µL chilled 100% isopropanol alcohol were dispensed in new 1.5 ml labeled tubes and centrifuged at 13.2 X 1000 rpm for 4 minutes. The supernatant of each sample was poured off, and the tubes were blotted dry, leaving a small DNA pellet. Following, 300 µL of 70% chilled ethanol was added to each tube and centrifuged at 13.2 X 1000 rpm for 4 min. The supernatant was again discarded, and tubes were blotted and placed, uncapped, on a 65°C heat block for 30 min to allow for ethanol evaporation. The extraction product was then re-suspended in a 50 µL Tris: EDTA solution and left at room temperature for at least 12 hours, followed by storage at -20°C.

Successful DNA extraction of the cockroaches was confirmed using universal COI insect paired primers LepF and LepR (Table 5.1) under the following thermocycler conditions: 2 minutes at 95°C, 39 cycles of 94°C for 45 seconds, 48°C for 1 minute, and 72°C for 1 minute, followed by a final extension of 72°C for 5 minutes. These primers amplify a 658 bp target amplicon. Detection of PCR amplicons for this and subsequent molecular diagnostic assays was
done by subjecting PCR products to electrophoresis on a 2% agarose gel and visualized using a BioDoc-it™ Imaging System (UVP, Inc., Upland, CA).

Molecular species-level identification was achieved via DNA sequencing. PCR product successfully amplified underwent purification and concentration using VWR centrifugal devices (VWR, Radnor, PA). Forward and reverse primers LepF and LepR were combined with the purified DNA and sent to Eurofins Genomics (Diatherix, Huntsville, AL) for direct sequencing bidirectionally. Consensus sequences with the primer ends removed were obtained using Geneious v6.1.8 (Biomatters Ltd., Auckland, New Zealand) and compared using an NIH BLAST search of DNA sequences available on NCBI GenBank (National Center for Biotechnology Information). This was used to confirm identity and percent similarity between our samples and those on GenBank. Our sample sequences and mtDNA COI sequences of Parcoblatta species found on GenBank were used for phylogenetic analysis. Multiple alignment of sequences was done using Geneious with a cost matrix of 65% similarity, a gap open penalty of 12, and a gap extension penalty of 3. The outgroup taxon was Periplaneta fuliginosa (GenBank accession number MH184322). Bayesian phylogenetic analysis was conducted with the MrBayes (Ronquist and Huelsenbeck, 2003) plug-in within Geneious v6.1.8 with 100,000 burn-in and 1,000,000 replications using a HKY85 model.

Samples with confirmed DNA presence underwent a multiplex qPCR using the StepOne Real-Time PCR System (Applied Biosystems, Waltham, MA) to detect for Hymenopteran DNA and N. ceranae presence or absence and then for Hymenoptera DNA and L. passim presence or absence (Bourgeois et al., 2010; Xu et al., 2018). Multiplex qPCR was conducted with 12.5μL reaction volumes using 0.25μL of each primer and probe (Table 5.1), 6ul TaqMan Fast Universal Master Mix (Applied Biosystems, Waltham, MA), 4.5ul dH2O, and 0.5 μL DNA.
Each reaction included a negative (dH2O) and a known positive. Conditions for \textit{N. ceranae} PCR followed an initial 20 seconds at 95°C followed by 40 cycles of 1 second at 95°C and 20 seconds at 63°C. PCR conditions for \textit{L. passim} followed an initial activation of the Taq DNA polymerase at 95°C for 10 minutes, followed by 40 cycles of 15 seconds of denaturation at 95°C and a 1-minute extension at 60°C. Type culture from the American Type Culture Collection (ATCC, Manassas, VA) was obtained for \textit{L. passim} (PRA-422 ATCC), and a previously sequenced, \textit{N. ceranae} sample were included to serve as positive controls.

Results

No members of the \textit{Periplaneta} genus were identified in our study. All of the cockroaches sampled were determined to be members of the \textit{Parcoblatta} genus (Figure 5.1). A total of 23 adult male cockroach samples were identified as \textit{P. pennsylvanica} using taxonomic keys (Bell, 1981; Blatchley, 1920; Hebard, 1917). A total of 14 cockroach samples were sequenced for species confirmation, of these, 12 belonged to one haplotype (WR Parcoblatta haplotype 1), and the other two a second haplotype (WR Parcoblatta haplotype 2). Between the two haplotypes, a total of eight nucleotides were different from the 623 bp sequence. The two haplotypes were most similar to two sequences from GenBank. The WR Parcoblatta haplotype 2 was most similar to \textit{Parcoblatta} sp. HM385635 (99.8% match), and Parcoblatta sp. HM385634 (99.8% match), While WR Parcoblatta haplotype 1 was 98.9% similar to these two GenBank sequences. The Bayesian phylogenetic analysis revealed that the two wood roach haplotypes from our study formed a common clades with the two \textit{Parcoblatta} sp. sequences previously mentioned. This clade formed a distinct sister clade with four \textit{Parcoblatta pennsylvanica} GenBank sequences (Figure 5.1).
Among the key characteristics that distinguish *P. pennsylvanica* from other species is the white or pale yellow margins on the edges of the pronotum, thorax, and front half of the wings (Figure 5.2). The primary distinguishing trait in males is the concave base of the median and first dorsal abdominal segment and the two paired ridge structures which are lacking in the similar *P. divisa* (Cochran, 1999). Additionally, *P. pennsylvanica* is darker in coloration and larger in size than other *Parcoblatta* species in the U.S. (males are 20.7 mm on average) (Atkinson et al., 1991; Bell, 1981; Peck & Beninger, 1989).

Q-PCR results indicated Hymenoptera DNA presence in three dissected guts of the cockroaches sampled. Neither the dissected guts nor the cockroach abdominal samples were positive for *N. ceranae* and *L. passim*.

**Discussion**

This is the first study to identify cockroaches to species in U.S. honey bee colonies. Our results of identifying *P. pennsylvanica* in honey bee colonies are predictable as it is among the six common *Parcoblatta* species occurring in the Southeastern U.S. (Figure 5.1). Furthermore, *P. pennsylvanica* is the most common *Parcoblatta* species in the U.S. (Cochran, 1999).

Cockroach species such as *P. fuliginosa* have been speculated to occur in honey bee colonies, additional research is necessary to determine if other species occur in U.S. honey bee colonies.

Additionally, this is the first study since 1955 to explore roaches in honey bee colonies and their vector potential of harmful honey bee pathogens (Stejskal, 1955). Previous literature has observed *P. pennsylvanica* feeding on dead insects; based detecting of Hymenoptera DNA in three of the gut samples, these roaches likely fed on dead honey bees (Blatchley, 1920). Because *Parcoblatta sp.* has been observed to fly up to 30m, there is potential for wood roaches to travel between honey bee colonies, especially when bee yards have multiple colonies within close
proximity. Cockroaches have been found to spread gregarines in honey bee colonies; therefore, it is possible that they could carry pathogens from one colony to another (Stejskal, 1955). Pathogen spread could be exacerbated by migratory beekeeping, in which colonies are shipped several thousand kilometers across the country and intermingled with other colonies (Alger et al., 2018).

While our study did not detect *N. ceranae* nor *L. passim* in the cockroaches sampled, the small sample size limits potential findings. Ideally, pathogen detection would be performed on cockroaches found in colonies with known infection of pathogens. Future research should increase sampling efforts and broaden the sampling range to better determine *P. pennsylvanica* as a vector of honey bee pathogens. Additionally, future studies should expand the number of pathogens screened. For example, *Spiroplasma apis* and *S. melliferum* are bacterial pathogens thought to spread via floral surfaces. Cockroaches could act as a passive vector of bacterial pathogens.

Future research efforts should focus on exploring the cockroach’s role within honey bee colonies. While literature has suggested a negative role (Argo, 1874; Bell et al., 2007; Blatchley, 1920; Stejskal, 1955), MacDonald & Matthews (1983) suggest that *Parcoblatta* sp. are beneficial recyclers within Southern yellow jacket colonies. Future research could also involve inoculating roaches with pathogens and introducing them to honey bee colonies negative for pathogen presence to determine whether roaches can transmit the pathogens. Monitoring for cockroaches in honey bee colonies could be used to augment the existing Cooperative Agricultural Pest Survey (CAPS) honey bee survey to determine if these organisms should be included in nationwide surveys.
Acknowledgements

This research was supported in part by the University of Arkansas, Arkansas Agricultural Experiment Station.
References


Lawson, F. A. (1967). Ecological and collecting notes on eight species of *Parcoblatta* (Orthoptera: Blattidae) and certain other cockroaches. *Journal of the Kansas Entomological Society, 267–269*.


Crithidia mellificae as a new contributory factor to winter mortality. *PLOS ONE*, 8(8), e72443. https://doi.org/10.1371/journal.pone.0072443


Table 5.1. List of primers used for molecular detection of cockroach DNA, Hymenoptera DNA, and parasites and pathogens. A: Modified from Hajibabaei et al., 2006 B: Xu et al., 2018 C: Bourgeois et al., 2010

<table>
<thead>
<tr>
<th>Primer/Probe</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LepF</strong></td>
<td>F: 5′-ATTCAACCAATCATAAAAGATATTGG-3′</td>
<td>A</td>
</tr>
<tr>
<td><strong>LepR</strong></td>
<td>R: 5′-TGATTTTTTGACATCCAGAGTATTTA-3′</td>
<td>A</td>
</tr>
<tr>
<td><strong>Hymenoptera</strong></td>
<td>F: 5′-TAACCTGGCATTATGTTAGTACGTC-3′</td>
<td>B</td>
</tr>
<tr>
<td><strong>Hymenoptera</strong></td>
<td>R: 5′-CCTCGACACTCAGTGAAGAGGC-3′</td>
<td>B</td>
</tr>
<tr>
<td><strong>Hymenoptera</strong></td>
<td>Probe: 5′-[HEX]AGCTCCTTCAA[BHQ1a-Q]-3′</td>
<td>B</td>
</tr>
<tr>
<td><strong>L. passim</strong></td>
<td>F: 5′-CGAGCTCATAAAATAATGTAAGCATAAATAAG-3′</td>
<td>B</td>
</tr>
<tr>
<td><strong>L. passim</strong></td>
<td>R: 5′-TTTTAGCAATATTTTAGCAAGCTCAGG-3′</td>
<td>B</td>
</tr>
<tr>
<td><strong>L. passim</strong></td>
<td>Probe: 5′-[HEX]TTGGTGTTTGCTATGT[MGB] -3′</td>
<td>B</td>
</tr>
<tr>
<td><strong>N. ceranae</strong></td>
<td>F: 5′-AAGAGTGAGACCTATCAGCTAGTTG-3′</td>
<td>C</td>
</tr>
<tr>
<td><strong>N. ceranae</strong></td>
<td>R: 5′-CCGTCTCTCAGGGCTCTTCTC-3′</td>
<td>C</td>
</tr>
<tr>
<td><strong>N. ceranae</strong></td>
<td>Probe: 5′-[JOE]ACCGTTACCGTCACAGCCTTGT-3′</td>
<td>C</td>
</tr>
</tbody>
</table>
Figure 5.1. Bayesian phylogenetic tree of partial COI sequence data from Arkansas cockroaches compared to sequences downloaded from Genbank.
Figure 5.2. Adult male *Parcoblatta pennsylvanica* samples morphologically identified using the paired ridge structures located on the first and median abdominal segment (A) and the marginal lightening along the pronotum and wings (B).
Chapter 6: Conclusion

This dissertation examines several factors associated with honey bee health. Because *Apis mellifera* L. is such an important economic pollinator species, further research into improving its health is necessary. Research to date has heavily focused upon commercial honey bee colonies, leaving knowledge gaps pertaining to swarms, feral, and hobbyist-managed honey bee colonies. Specifically, there is a lack of research pertaining to the genetic diversity and parasite and pathogen occurrence in these colonies. Furthermore, pests occur within honey bee colonies which lack proper species identification or an understanding of vectoring potential.

This dissertation investigated the mitochondrial genetic diversity and pathogen prevalence of honey bee colonies from managed, feral, and swarm populations. We sampled honey bee colonies from Arkansas, Hawaii, Missouri, Oklahoma, and Utah. The genetic diversity of Utah feral colonies and swarms and Arkansas hobbyist-managed and unmanaged colonies were characterized using mitochondrial DNA (mtDNA) sequencing in chapter 2 and chapter 3. In unmanaged colonies in Utah, we found a surprising amount of haplotype diversity, specifically within the Africanized lineage. Two of the detected haplotypes are unique to Utah in the U.S. In Arkansas, we found an unexpected amount of genetic diversity compared to previous COI-II surveys, specifically commercial queen breeder studies. This suggests that feral populations of honey bees have a unique reservoir of genetics compared to managed populations which could provide diversified genetic stocks for future breeding programs. This diversification could potentially improve immunity to harmful pathogen species in honey bee populations.

In chapter 4, honey bee genetic variation and pathogen (*Varroa destructor* Anderson and Truman, *Nosema ceranae* Fries, *Lotmaria passim*, *Spiroplasma melliferum*, and *S. apis*) occurrence was examined in managed, feral, and swarmed honey bee populations from three
areas of the U.S. Research has shown that the maintenance of genetic variation at the colony level is important to overall colony health, and may reduce pathogen susceptibility (Tarpy, 2003; Tarpy & Seeley, 2006). Due to long-term selective breeding of a limited genetic stock of C lineage honey bee queens, introduction of feral genetics from differing haplotypes and genetic lineages may increase the potential for genetic material associated with useful apicultural traits in future breeding programs.

In the 5th chapter, a commonly occurring pest in honey bee colonies was examined, the cockroach. This is the first time the species of cockroach in U.S. honey bee colonies has been explored. We found that the species occurring in the Arkansas honey bee colonies samples was *Parcoblatta pennsylvanica*. Also, in this chapter we identified honey bee DNA in the wood roach, verifying that the roaches do feed on dead bees or honey bee feces. This led us to explore whether the wood roaches collected from Arkansas honey bee colonies had DNA from harmful honey bee pathogens which could potentially link them as a passive vector. While none of the screened pathogens were detected in our wood roach samples, further research is necessary to understand their role in honey bee colonies.

This study is among the first to comprehensively explore genetic origin, management source, and regionality as factors of pathogen infection in U.S. honey bee populations. Such studies are necessary to understand and improve honey bee populations in the U.S.