AN INVESTIGATION OF GENETIC VARIABILITY IN *LUCILIA CUPRINA* AND *MUSCA DOMESTICA* UTILIZING PHYLOGENETIC AND POPULATION GENETIC APPROACHES

by

Laura Catherine Doll

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Forensic and Investigative Sciences at IUPUI

Indianapolis, Indiana

August 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Christine Picard, Chair

Department of Biology

Dr. Randall Roper

Department of Biology

Dr. Aaron Tarone

Department of Entomology, Texas A&M University

Approved by:

Dr. John Goodpaster

I would like to dedicate this work to my family and friends who have always supported my dreams and education.

ACKNOWLEDGMENTS

I would first and foremost like to thank my advisor, Dr. Christine Picard, for her continued support of my education and guidance in becoming a better scientist. I will be forever grateful for all the opportunities that have been and will be presented to me due to my time at IUPUI. Thank you also to Donna Roskowski for guiding me in my teaching and supporting me in rough times. Thank you to Dr. Aaron Tarone and Dr. Randall Roper for serving on my committee and offering guidance and additional perspectives. Thank you to Dr. Jeffrey Wells for assisting with data analysis. Thank you to Douglas Seemann for collecting fly specimens in Arizona. Thank you to my lab mates Charity Owings, Anne Andere, Khadija Moctar, and Ionna Koltsidou, not just for helping me in the lab and providing guidance, but for being good friends.

I would like to thank my parents, my brother, and my grandma for always supporting my education and pushing me to be the best I can be. Thank you also, of course, to my dogs Fiona, Toby, and Riley, for offering unconditional love when I need it the most. Thank you to my core group of friends who have stuck by me through thick and thin: Courtney Burke, Sydney Goyette, Catherine Moore, Erica Taylor, Rachel Daniels, and Haley Jo Keller. Thank you to the friends I have made at IUPUI for going through this experience with me and for all the fun times we have had.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Forensic entomology is a subdiscipline of entomology that involves the use of insect behavior and developmental data to aid in criminal investigations. Genetic data has become increasingly important to the field as there has been a push for DNA-based species identification methods of forensically relevant insects. Genetic data can also elucidate population structure and relatedness of these insects, and such knowledge can contribute to the development of more specific datasets for insects in different regions. The first study presented here investigated the phylogenetics of sister species *Lucilia cuprina* and *Lucilia sericata* to identify possible subspecies divisions and issues with DNA-based identifications in the United States. The initial aim of this study was to identify genetic differences between specimens of *L. cuprina* that preferred live versus carrion flesh. Flies collected from Indiana, USA and South Africa were sequenced and analyzed. Upon sequencing of the genes *COI, Period*, and *28s,* our results indicated that *L. cuprina* from Indiana possess a unique combination of nuclear and mitochondrial haplotypes that suggest a unique lineage, possibly indicating modern hybridization with *L. sericata.* The inability of both nuclear and mitochondrial genes to distinguish between *L. cuprina* and *L. sericata* raises questions about the capabilities of DNA-based species identifications within this genus. Additionally, the inability of these genes to distinguish between specimens that preferred live versus carrion flesh highlights a need for continued research of these behavioral differences. The second study presented here investigated the population structure and relatedness of house flies in the American southwest in relation to a civil lawsuit where neighbors of a poultry farm alleged that flies were emanating from the farm to their homes. *Musca domestica* (house fly) specimens were collected from the chicken farm and from locations in varying directions and distances from the farm. Amplified fragment length polymorphism (AFLP) analysis was performed and the data were used in a number of analyses. Population reallocation simulations generally indicated that samples from different locations were not genetically different enough from other locations to allocate to their true origin population over others. Kinship analysis showed differences in samples collected in a later season that indicate a genetic bottleneck over time. Population structure analysis indicated the presence of two intermixing genetic populations in the dataset. AMOVA revealed that the majority of genetic variation laid within, rather than among, populations. A Mantel test revealed

no significant correlation between genetic and geographic distances. These results indicate that the *M. domestica* population in this region of southwestern America is large and intermixing, with no clear genetic distinctions between specimens collected at the poultry farm versus the surrounding locations. In regard to the civil lawsuit, it was not possible to conclude that the flies did not emanate from the poultry farm. In a broader perspective, these data can be utilized to develop pest management strategies in this region. Overall, the data from both studies presented here will be useful to forensic investigations, development of more specific and detailed data and identification techniques, and pest control measures.

CHAPTER 1. INTRODUCTION

1.1 Forensic Entomology

Entomology is the study of insects [1]. The discipline is quite widespread, having applications from pest control and ecological research to human and animal health. One subdiscipline of entomology is forensic entomology, which is the study of insects as it pertains to legal matters [2]. Forensic entomologists study the invasion and succession of arthropods found on corpses [3]. Insect behavior and developmental data are then used to aid in criminal investigations. The most recognized use of forensic entomology is the calculation of postmortem intervals (PMI) and minimum postmortem intervals ($_{m}$PMI) of a victim based on the presence of the insects found on the corpse [4] . However, forensic entomology can be useful in a number of other circumstances, including elucidating body relocation, abuse and neglect in living victims, wound identification, toxicological analyses, and linkage of suspects to crimes [1]. Additionally, insect data can be forensically relevant in nonviolent circumstances, such as cases of insects transmitting pathogens to residences and other human-inhabited areas.

1.1.1 Forensically Relevant Insects

Because of their known roles in insect succession on a corpse, certain families of flies and beetles are considered forensically relevant or insects of forensic importance. Flies (Order Diptera) involve a number of families that are forensically relevant, with the two major examples being blow flies (Diptera: Calliphoridae) and flesh flies (Diptera: Sarcophagidae). There are over 1000 species of blow flies that are found worldwide, albeit with different distributions. Calliphorids are not only attracted to decaying tissue (carrion), but also to dung, some vegetation, and in some species, open wounds on living animals (referred to as myiasis) [1]. They are not only important forensically, but their role of breaking down carcasses is invaluable to nutrient cycling and community ecology [5]. Blow flies are among the first insects to find and colonize remains.

Flesh flies contain over 2000 species that generally prefer warm climates. Unlike blow flies, flesh flies are well-known to be attracted to nectar and feed on decomposing vegetation [1]. Still, they can be among the first colonizers of carrion. Flesh flies tend to be more prominently

found on remains located indoors. Like blow flies, flesh flies have been known to inhabit and feed on open wounds [3]. In addition to Calliphorids and Sarcophagids, a number of additional fly families are considered forensically relevant including Muscids (Family Muscidae), skipper flies (Family Piophilidae), dung flies (Family Scathopagidae), among others [1, 3].

In addition to the many families of flies that are forensically important, a number of families of beetles (Order Coleoptera) are also forensically relevant. The Order Coleoptera contains about a third of known insect species. Carrion beetles (Family Silphidae) contains over 1500 species, with approximately forty-six being widespread in North America [1]. Adult carrion beetles may arrive during an early stage of decomposition and feed not only on the decaying carcass, but also feed on fly eggs and larvae. Carrion beetle larvae are present during and indicative of later stages of decomposition. Additional forensically relevant beetles include rove beetles (Family Staphylinidae), clown beetles (Family Histeridae), and checkered beetles (Family Cleridae), among others [1, 3]. Of all forensically relevant insects, the current studies focus on *Lucilia cuprina* (Diptera: Calliphoridae) and *Musca domestica* (Diptera: Muscidae).

1.1.2 Insect Succession

When a human or animal dies, insects are capable of finding and colonizing the body within minutes [2]. Different insect species typically colonize and leave the decaying body in a predictable order, and as such the process is referred to as ecological succession [6]. The typical insect succession is described as follows: blow flies (Diptera: Calliphoridae) and flesh flies (Diptera: Sarcophagidae) are the first colonizers attracted to the fresh carcass. These flies ovipost (blow flies) or larviposit (flesh flies) on the carcasses and the maggots mature using the carrion flesh as a protein source. When maggots are high in activity and form maggot masses, gases begin escaping from the body and different species of predacious beetles (Order Coleoptera) arrive. Once the flesh has been consumed and the body becomes dry, blow flies are no longer attracted [3]. Because of this predictable succession, the composition of species on a decaying body at any given time is indicative of its stage of decay. Studying the species involved in succession and understanding their genetics and development allows forensic entomologists to deduce the time needed for the insects found on a body to have reached their current state of development.

1.2 Genetics in Forensic Entomology

1.2.1 DNA-Based Species Identification

Population genetics is the study of the origin, amount, and distribution of genetic variation in target populations using genomic data [7]. Studying this information is useful in elucidating factors such as relatedness, genetic structure, adherence to Hardy-Weinberg equilibrium, and population structure [8]. In forensic entomology, DNA data and its associated uses have become increasingly important due to the rise of DNA-based species identifications of larvae and adult insects found at crime scenes. Identifying the species of insect found on a body is an essential role of the forensic entomologist, and this identification will dictate downstream analyses and PMI determinations [9]. However, the insects found on a corpse are generally in a larval stage and lack the major distinguishing morphological differences that they will exhibit once they reach adulthood. Morphological keys for insects in their larval stage are difficult to use for the non-expert [10], making successful larval species identifications based solely on morphology a challenging task. Because of this issue, many researchers have suggested and investigated the use of molecular methods to identify or confirm species identities [11, 12].

Mitochondrial DNA (mtDNA) has played a large role in studies investigating insect species identification due to its haploidy, high copy number, and availability of conserved primers [12, 13]. Most studies utilizing mtDNA focus on *cytochrome c oxidase* subunits I and II (*COI* and *COII*), which roughly coincide with the proposed universal animal DNA barcoding region [14]. However, many mtDNA regions are highly conserved, and as such, mtDNA analysis may not be able to resolve closely related species [10, 12]. A number of nuclear DNA markers have been utilized in studies on DNA-based identification on insects, including but not limited to *28S* rRNA, *NADH dehydrogenase* subunit 5, and *Period* [12, 15, 16]. Though many have seen promising results, there is currently no agreement on which locus or loci to use in forensic entomological identifications. However, many studies have emphasized the importance of utilizing multiple loci rather than a single locus [17-19].

1.2.2 Population Assignment

It is known that fly populations of many species can vary genetically and phenotypically in different regions and under different conditions. A 2014 study found that populations of the

blow fly *Cochliomyia macellaria* in three regions of Texas exhibited significant differences in pupal mass as well as immature and pupal development times [20]. Flies from a single population reared under varying conditions (i.e. food moisture, substrate) experience significant differences in development time [21]. Studies of this type emphasize that genotype, phenotype, and/or environmental effects can lead to inter and intra-species variations. Due to the possible differences that may be present, it is important to investigate to what population a fly specimen belongs in order to utilize the most accurate dataset for postmortem interval (PMI) determination. Additionally, population genetics can be utilized to infer the relatedness of insects found on a corpse. Multiple studies have indicated patterns of local relatedness in blow flies at a single location or bait [22, 23]. Such information could be used to genetically identify unrelated individuals, which could indicate that a corpse had been moved from its original location. Further, investigating differences in populations allows development of new, more specific developmental datasets to be used in forensic investigations.

1.2.3 Phylogenetic Analyses

Phylogenetics is the reconstruction of genealogies and evolutionary relationships to describe the descent and dynamics of species [24]. Phylogenetic data are essential to modern forensic entomology because DNA-based identification methods must be able to distinguish between closely related species. If DNA-based methodologies were unable to discriminate between species, which is the goal, their use would be moot. Phylogenetic research on forensically relevant insect species provides the basis that makes genetically distinguishing between species and proper species determinations possible.

Studies of this nature tend to have two approaches: a general approach, and an approach investigating a specific characteristic. A general approach may simply analyze the genetic data to see what the results indicate and understand whether or not species are able to be resolved, such as in a 2001 study on Calliphorids [25]. Such studies provide a general basis for the use of certain markers and may identify overall evolutionary patterns. A specific approach entails the use of phylogenetic analyses to investigate a specific characteristic or behavior rather than a general overview. One such behavior that is particularly relevant to forensically important insects and the present study is the investigation of the evolution of myiasis, or larval infestation of open wounds on live animals. The evolution of this behavior within Calliphoridae is unclear

[26], especially since some species are known to exhibit different behaviors in different regions [18, 27]. Phylogenetic analyses of individuals with and without this behavior aim to resolve genetic distinctions and evolutionary patterns that explain these differences.

1.3 Conclusion

The goal of this research was to utilize the afore described population genetic and phylogenetic approaches to identify genetic variation within the blow fly *L. cuprina* and the filth fly *M. domestica.* The data generated from this work may be utilized to increase the specificity of forensic entomological testing and could influence pest control management strategies in certain regions. We present the first inclusion of *L. cuprina* sequence data from specimens originating from Indiana, which expands upon previous knowledge of the status of this fly in the United States. Additionally, to our knowledge, we present the first use of AFLP (amplified fragment length polymorphism) analysis on the house fly *M. domestica,* which could be used as a proof of concept and influence the direction of future studies.

CHAPTER 2. PHYLOGENETIC ANALYSIS OF THE BLOW FLY *LUCILIA CUPRINA*

2.1 Introduction

Lucilia cuprina Wiedmann (Diptera: Calliphoridae), or the Australian sheep blow fly, is species of blow fly primarily known for its role as a pest in certain regions. Like most blow flies, female *L. cuprina* oviposit onto carrion in the environment so larvae have a protein source for development [26]. However, this behavior is known to vary in different parts of the world for this species [27, 28]. In some geographic areas such as Australia and New Zealand, *L. cuprina* primarily exhibits parasitic behavior in the form of myiasis [29], also commonly referred to as sheep strike or fly strike. In these areas, *L. cuprina* larvae infest the living tissue and open wounds of animals, with their primary victim being domestic sheep [28]. Flystruck sheep experience rapid breathing and loss of appetite leading to weight loss, and will likely die by blood poisoning if no treatment is received [30]. In addition to costing sheep their health, myiasis comes at an extreme economic cost – in Australia, where *L. cuprina* is responsible for more than 90% of cases [31], it is estimated that myiasis prevention and treatment cost the wool industry $173 million per year [32]. Estimates that expand to include Australasian regions estimate damages up to $320 million annually [33]. Despite the disastrous impacts resulting from parasitic *L. cuprina* in the Oceanic region and its surroundings, the behavior of *L. cuprina* differs in other areas of the world. The species is present and relatively widespread throughout North America, but seems to prefer carrion in this region rather than living flesh [34]. There is no record of *L. cuprina* being important to sheep myiasis in the United States [27]. It is believed that the Australian region is particularly susceptible to myiasis due to certain susceptible breeds of sheep and a suitable climate [33, 35]. Blowfly strike does occur in other regions, but is often primarily caused by different species (e.g. *Lucilia sericata* in northern Europe) [36].

The evolution of the myiasis behavior within blow flies, and the genus *Lucilia,* is not fully understood. Data suggests that primary obligate parasitism (that is, a preference for feeding on tissue of living hosts) likely arose independently at least five times throughout a number of taxa [26]. The situation is further complicated by both inter- and intraspecies variation in myiasis behavior within the genus *Lucilia* [27]. It has been suggested that multiple ectoparasitism evolution events coevolved in recent history, likely alongside the husbandry of sheep [27].

Proper insect identification is crucial in a forensic context [4, 9], especially if such an identification will be used in a PMI calculation that could implicate a suspect. With the range of behavior variation observed within *L. cuprina*, the ability to distinguish between variants remains a crucial task. Phylogeny studies as well as DNA-barcoding research thus far concerning *L. cuprina* have largely investigated the species in relation to its sister species, *L. sericata.* This is largely due to an occurrence of hybridization between the species seen in Hawaiian populations [15, 17, 18, 28]. Upon mtDNA sequence analysis of both *L. cuprina* and *L. sericata* from varying regions, *L. sericata* and Hawaiian *L. cuprina s*amples could not be differentiated [29]. Additional research has confirmed a paraphyletic relationship between the mtDNA of *L. sericata* and certain *L. cuprina* [18], specifically in regard to *COI*, a popular DNA-barcoding marker [11]. Because of the lack of *COI* reciprocal monophyly exhibited within the genus *Lucilia*, species identification based on this popular marker may not be as straightforward as was hoped [10].

Studies utilizing nuclear DNA markers have shown a clearer distinction between *L. cuprina* and *L. sericata.* Phylogenetic analyses using the mtDNA gene *COI* as well as nuclear genes *28s* rRNA and *Period* (*Per*) indicated that *L. sericata* and *L. cuprina* are sister species with 100% support, while the *COI* analyses showed *L. cuprina* as paraphyletic with respect to *L. sericata* [16]. However, combined analyses of all 3 genes still resolved *L. cuprina* and *L. sericata* as sister clades, highlighting an importance of using multiple genes in species identifications.

We aimed to investigate trends in the clustering of *L. cuprina* in relation to *L. sericata,* which we suspected could be indicative of a novel subspecies relationship as well as incorporate a new regional *L. cuprina* population recently found in Indiana [37]. *Lucilia cuprina* is known to have two subspecies, *L. cuprina cuprina* (Wiedmann) and *L. cuprina dorsalis* (Robineau-Desvoidy), which have varying distributions and are distinguishable by subtle morphological variation [34]. We suspected that the two protein source preferences of *L. cuprina* (carrion versus myiasis) may follow this, or a new, subspecies classification. *Lucilia cuprina cuprina* is distributed throughout the neotropical, oriental, and neacrtic regions [38], following the general distribution trend of carrion-preferring *L. cuprina*. *Lucilia cuprina dorsalis* has been found throughout sub-Saharan Africa and Australasian regions [39], following the general trend of *L. cuprina* that prefer myiasis. It has been previously suggested that the two apparent *L. cuprina COI* haplotypes reflect the two subspecies [17, 19]. The geographic trends of *L .c. dorsalis*

correspond to the mtDNA lineage that is more distantly related to *L. sericata,* suggesting that *L. c. cuprina* corresponds to the *L. sericata*-like mitochondrial lineage [19]. We hypothesized that *L. cuprina* may additionally cluster according to its preference for carrion or live flesh, which could follow the current subspecies division or imply additional distinctions. Such knowledge would allow development of a genetic test for subspecies and variations of *L. cuprina,* which could be an essential tool for research and forensic investigations.

2.2 Materials and Methods

2.2.1 Blow Fly Collection and Identification

Indiana specimens were collected in various local parks using a sweep net and decayed meat as bait (Table 2.1), while South African specimens were either collected from live sheep or from carcasses or using decayed meat bait (Table 2.2). The species of each collected specimen was confirmed using morphological keys [39, 40]. The main morphological feature used to distinguish between adult *L. cuprina* and *L. sericata* was the number of setae below the inner vertical seta [41]. A subset of specimens collected in Indiana that were identified as *L. cuprina* were sent to Dr. Terry Whitworth, the author of a number of widely used blow fly morphological keys [39]. Dr. Whitworth confirmed our identifications of *L. cuprina.* Though the specific specimens sent to Dr. Whitworth were not used in these analyses, the *L. cuprina* specimens described presently were identified in the same manner using the same features.

Table 2.1. *Lucilia* specimens collected in Indiana, U.S.A. using decayed meat bait.

Specimen ID	Site	City	Latitude	Longitude	Date	Species
IN-1	Broad Ripple Park	Indianapolis	39°52’17”	-86°07’51”	25-May-16	*Lucilia cuprina*
IN-2	Province Park	Franklin	39°28’37”	-86°06’39”	25-May-16	*L. cuprina*
IN-4	Province Park	Franklin	39°28’37”	-86°06’39”	23-Jun-16	*L. cuprina*
IN-5	Province Park	Franklin	39°28’37”	-86°06’39”	23-Jun-16	*L. cuprina*
IN-7	Province Park	Franklin	39°28’37”	-86°06’39”	7-Jul-16	*L. cuprina*
IN-9	University Park	Greenwood	39°36’36”	-86°03’02”	13-Oct-17	*L. cuprina*
IN-10	University Park	Greenwood	39°36’36”	-86°03’02”	13-Oct-17	*L. cuprina*
IN-11	Military Park	Indianapolis	39°46’16”	-86°10’08”	13-Oct-17	*L. cuprina*
IN-12	Province Park	Franklin	39°28’37”	-86°06’39”	26-Jul-17	*L. cuprina*
IN-13	Province Park	Franklin	39°28’37”	-86°06’39”	26-Jul-17	*L. cuprina*
IN-14	Province Park	Franklin	39°28’37”	-86°06’39”	26-Jul-17	*L. cuprina*
INLc-1	University Park	Greenwood	39°36’36”	-86°03’02”	23-Jun-16	*L. cuprina*
INLc-2	Province Park	Franklin	39°28’37”	-86°06’39”	13-Oct-16	*L. cuprina*

Table 2.1 continued

INLs-1	Military Park	Indianapolis	39°46’16”	-86°10’08”	19-Apr-17	*Lucilia sericata*
INLs-2	Military Park	Indianapolis	39°46’16”	-86°10’08”	19-Apr-17	*L. sericata*
INLs-3	Broad Ripple Park	Indianapolis	39°52’17”	-86°07’51”	28-Jun-17	*L. sericata*
INLs-4	Broad Ripple Park	Indianapolis	39°52’17”	-86°07’51”	28-Jun-17	*L. sericata*
INLs-5	Northwest Park	Greenwood	39°37’43”	-86°08’37”	17-May-17	*L. sericata*
INLs-6	Province Park	Franklin	39°28’37”	-86°06’39”	16-Aug-17	*L. sericata*
INLs-7	Province Park	Franklin	39°28’37”	-86°06’39”	16-Aug-17	*L. sericata*
INLs-8	University Park	Greenwood	39°36’36”	-86°03’02”	17-May-17	*L. sericata*
INLs-9	University Park	Greenwood	39°36’36”	-86°03’02”	17-May-17	*L. sericata*
INLs-10	Northwest Park	Greenwood	39°37’43”	-86°08’37”	26-Sep-17	*L. sericata*

Table 2.2. *Lucilia* specimens collected in South Africa either from live sheep (S) or a dead animal carcass/decayed meat bait (C). S specimens originated from one of three live sheep (numbered).

Specimen ID	Site	Locality	Latitude	Longitude	Date	Species
S2-2	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*Lucilia sericata*
S2-3	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*Lucilia cuprina*
S2-4	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S2-6	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S2-7	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S2-8	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S2-9	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S2-10	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S2-11	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S2-12	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S2-13	Sheep 2	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
S3-1	Sheep 3	Local Farm	-33°17’55.79"	26°27'49.9"	5-Jan-15	*L. cuprina*
C1-1	Carcass 1	Rhodes University	-33°18’48.82"	26°31'7.19"	7-Jan-15	*L. sericata*
C1-2	Carcass 1	Rhodes University	-33°18’48.82"	26°31'7.19"	7-Jan-15	*L. cuprina*
C1-3	Carcass 1	Rhodes University	-33°18’48.82"	26°31'7.19"	7-Jan-15	*L. cuprina*
C1-4	Carcass 1	Rhodes University	-33°18’48.82"	26°31'7.19"	7-Jan-15	*L. sericata*
C1-5	Carcass 1	Rhodes University	-33°18’48.82"	26°31'7.19"	7-Jan-15	*L. sericata*

2.2.2 Polymerase Chain Reaction (PCR) Amplification

DNA was extracted from either the head for the adults or the 3-7th thoracic segments (avoiding internal organs) from larvae. All DNA was extracted using a DNeasy® Blood & Tissue Kit (QIAGEN®) following the manufacturer’s protocol. The genes *cytochrome oxidase I* (*COI*), *Period* (*Per*), and *28s* rRNA were chosen for analysis. *COI* is a popular DNA-barcoding marker not only in entomological investigations, but for all animals. *COI* is a mitochondrial gene and is ideal for analysis due factors such as low recombination, high copy number, and availability of universal primers [42]. *28s* rRNA is a nuclear gene that has been used widely in

phylogenetic studies of insects as it displays both conserved and variable regions, and diverse rates of genetic evolution [42]. The *Per* nuclear circadian clock gene has been utilized in a number of related studies due to the hypothesis that it could provide better phylogenetic resolution due its fast evolution [15]. Our use of these genes builds upon previous research and allows comparison of results.

The primers C1-J1709 and C1-N2353 were used to amplify a ~600bp region of the *COI* gene in all samples [15, 16, 43]. These particular *COI* primers were chosen due to use in related studies and the presence of a ~300bp overlap with the DNA barcoding region [14-16]. The primers Per5.F and Per5.R were used to amplify a ~730bp region of the *Per* gene in all samples [15, 16]. The primers 28s_F and 28s_R [15, 16, 18] were used to amplify a ~650bp region in domain 1-2 of the *28s* gene in all samples (Table 2.3). The 10µL volume polymerase chain reaction (PCR) consisted of the following: 5µL 2X PCR Master Mix (Promega™), 0.5µL 1X bovine serum albumin (BSA) (Promega™), 1µL 5µM forward and reverse primers, and 2.5µL genomic DNA. The PCR protocol was followed as described by Williams et al. (2016) [16]. PCR conditions were as follows: denaturation for 5 min at 95ºC, 36 cycles of 94ºC for 30s, 55ºC for 60s, and 72ºC for 30s, followed by a final extension period of 72ºC for 7 minutes. Amplification was verified using a 1% agarose gel for expected sized products.

Table 2.3. Primers used to amplify *COI, Period*, and *28s* genes.

Primer	Target Gene	5' to 3' Sequence
C1-J1709	*COI*	AATTGGGGGGTTTGGAAATTG
C1-N2353	*COI*	GCTCGTGTATCAACGTCTATTCC
Per5.F	*Period*	GCCTCCAGATACGGTCAAAC
Per5.R	*Period*	CCGAGTGTGGTTTGGAGATT
28s_F	*28s*	CCCCCTGAATTTAAGCATAT
28s_R	*28s*	GTTAGACTCCTTGGTCGGTG

2.2.3 DNA Sequencing

Amplicons were purified using 1µL ExoSAP-IT™ (Thermo Fisher Scientific™) and 2.5µL PCR product following the manufacturer's protocol. Subsequently, cycle sequencing was performed using BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific™) following the manufacturer's protocol. Sequencing products were then purified by ethanol

precipitation: 1.25µL 125mM EDTA was dissolved into the samples. 20µL ice cold 95% ethanol was added, and the plate incubated at room temperature for 15 minutes. Samples were then centrifuged in a refrigerated centrifuge (Sorvall Legend XTR Centrifuge; Thermo Scientific™) at 2500g for 30 minutes. 95% ethanol was promptly removed and followed by the addition of 20µL ice cold 70% ethanol. Samples were placed in the same centrifuge used previously and centrifuged at 2500g for 15 minutes. After removal of the 70% ethanol, the plate was inverted and centrifuged at 185g for 1 minute. Samples were resuspended in 10µL HiDi™ formamide (Thermo Fisher Scientific™), vortexed for 15 seconds, and denatured for 5 minutes at 95°C using a Mastercycler Pro thermocycler (Eppendorf®). Sequence separation and detection was performed using a 3500 genetic analyzer (Thermo Fisher Scientific™).

In addition to this sequencing, whole genomes of several Indiana and South African samples were sequenced using Illumina 2 X 150bp paired end sequencing. Genomes were assembled using CLC Genomics Workbench v9.0.1 (QIAGEN®). Local BLAST was conducted on each genome to identify contigs of interest (Picard, unpublished).

2.2.4 Sequence Analysis

Sequences were visualized and manually edited using Sequence Scanner v1.0 (Applied Biosystems®) using a minimum Phred score of 20 for the beginning and end of the sequences, with the replacements of internal ambiguous bases with Ns, and aligned using Clustal Omega [44]. Alignments were previewed and trimmed using BioEdit v7.1.3.0 [45]. Additional DNA sequences of *COI, Per,* and *28s* in *L. cuprina, L. sericata, Phormia regina,* and *Chrysomya rufifacies* were obtained from GenBank® (National Center for Biotechnology Information; NCBI) (Table 2.4). Further, local BLAST was conducted on previously assembled genomes of *P. regina* [46] and *C. rufifacies* (Picard, unpublished), as well as a *L. cuprina* genome from Florida, USA (Picard, unpublished), to identify contigs containing the genes of interest.

Table 2.4. Additional DNA sequences obtained from GenBank® (NCBI).

Species	Gene	Accession Number
Lucilia cuprina	*COI*	AB112863.1
		EU626528
		FJ650544.1
		AJ417705.1
		AJ417711.1
		AJ417710.1
		DQ453496.1
		FJ650560.1
	Period	JN792805.1
		JN792809.1
		JRES01000940.1
		JN792819.1
		JN792814.1
		JN792815.1
		JN792791.1
		JN792786.1
	28s	KY197796.1
		FJ650542.1
		EU626539.1
		FJ650541.1
		FJ650533.1
		EU626550.1
		JN792705.1
		FR719302.1
		AJ417709.1
Lucilia sericata	*COI*	AJ422212.1
		AJ417714.1
		AB112843.1
		AB112850.1
		AJ417713.1
	Period	JN792856.1
		JN792833.1
		JN792836.1
	28s	KR133394.1
		JN792779.1
		EU626535.1
		JN792746.1
		AJ300141.1
		AJ300140.1
		AJ300139.1
Chrysomya rufifacies	*COI*	NC_019634.1
	28s	JQ246609.1
Phormia regina	*COI*	AF295550.1
	28s	AF366685.1

2.2.5 Phylogenetic Analysis

Phylogenetic analyses of all sequences were conducted using PAUP* v.4.0a164 [47]. Parsimony analyses were run using default parameters. *Phormia regina* and *C. rufifacies* sequences were set as outgroups. The number of bootstrap replicates was set to 1000, and bootstrap 50% majority-rule consensus trees are reported below.

2.3 Results

2.3.1 *Cytochrome Oxidase I* (*COI*)

When examining the phylogenetic tree based on *COI* sequence (Figure 2.1), it becomes apparent that there is no clear distinction between specimens collected from carrion and specimens collected from live sheep. Not only can sheep (S) and carrion (C) specimens be found within a single clade, but in some instances, they share an identical sequence (i.e. Figure 2.1 branch 1) (Table 2.5). Because the samples are interspersed and not confined to a single clade, the relationship is described as paraphyletic. Additionally, a paraphyletic relationship is seen between sister-species *L. cuprina* and *L. sericata,* as each species is not confined to its own clade.

This study marks the first investigation of its kind to include *L. cuprina* specimens from Indiana, USA. The parsimony analysis indicates that Indiana *L. cuprina* possess a *L. sericata*-like mitochondrial haplotype rather than a *L. cuprina*-like haplotype. All Indiana *L. cuprina* specimens (Figure 2.1 branch 10) fall within an otherwise strictly *L. sericata* clade and share an identical sequence with some Indiana *L. sericata* specimens. The Indiana *L. cuprina* specimens possess the only *L. cuprina COI* sequences to fall within clades in the bottom half of the tree where all *L. sericata* specimens fall, with the exception of one South African *L. sericata* sample (Figure 2.1 branch 1). Additionally, one South African *L. sericata* sample (Figure 2.1 branch 15; sample C1-5 included in group Lser_C_SA) appears to be more distantly related to the other *L. sericata* and *L. cuprina* specimens.

A Florida *L. cuprina* specimen that was previously collected and sequenced by the laboratory (Figure 2.1 branch 6; sample FL_Genome is included in group Lcup_FL_USA) falls within a clade containing other Florida *L. cuprina* samples, several South African *L. cuprina* samples, and Hawaiian *L. cuprina* samples. These results and their proximity to *L. sericata*

clades (Figure 2.1) suggest that our Floridian sample and the other samples in the clade also possess a more *L. sericata*-like haplotype. This is supported by previous studies that have identified a *L. sericata*-like *COI* haplotype in Hawaiian *L. cuprina* [15, 17, 18, 28] and some South African *L. cuprina* [17].

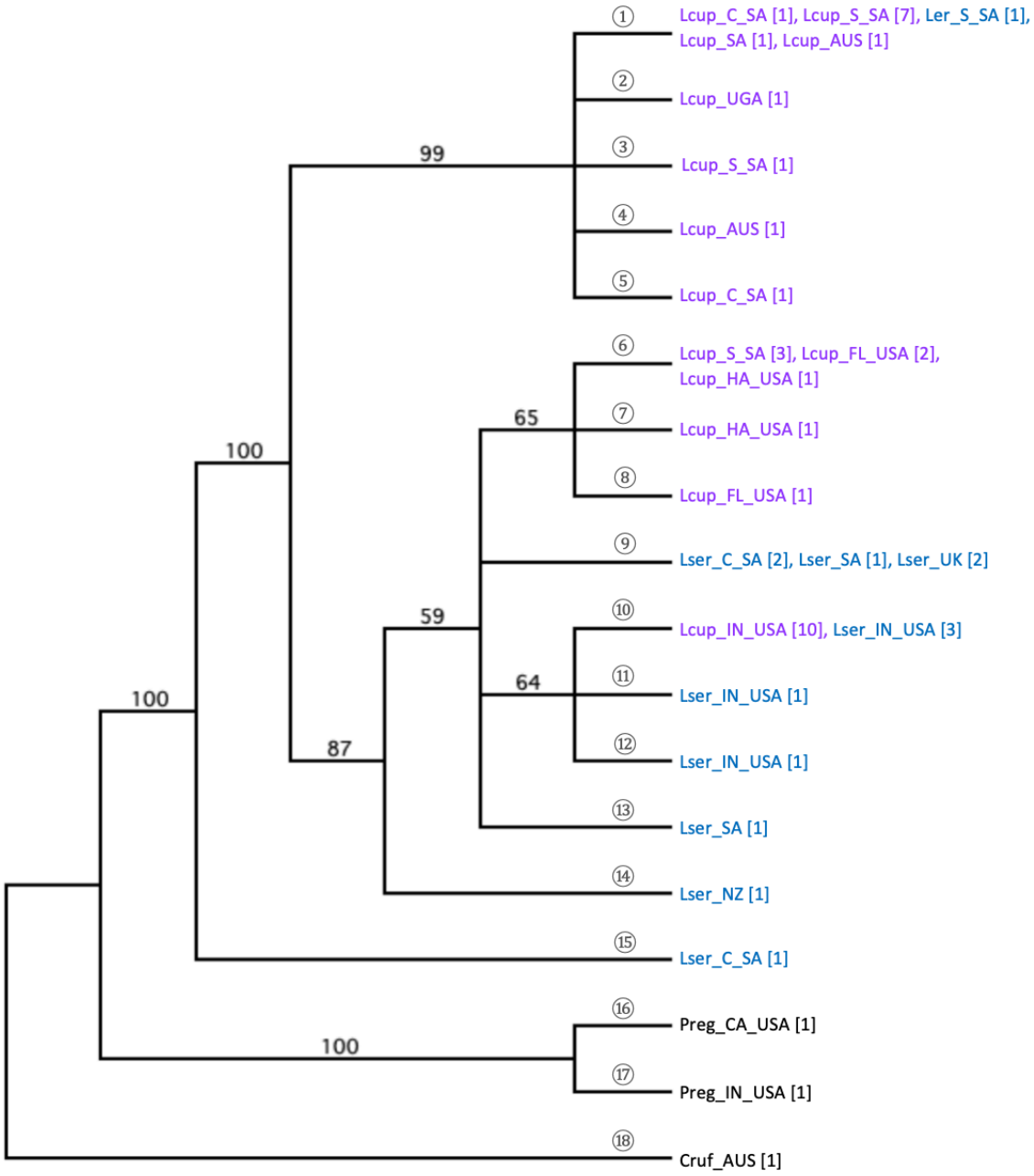

Figure 2.1 Maximum parsimony bootstrap (1,000 replicates) consensus tree based on *cytochrome oxidase I* (*COI*) sequence. Numbers on branches indicate percent bootstrap support. Bracketed numbers indicate the number of specimens included in each group. Group descriptions include species, location, and of those collected in South Africa, whether it came from live sheep or carrion. Each branch is labeled with a circled number for more in-depth description of samples included (Table 2.5). Lcup=*Lucilia cuprina* Lser=*Lucilia sericata* Preg=*Phormia regina* Cruf=*Chrysomya rufifacies* C=carrion S=sheep SA=South Africa UGA=Uganda USA=United States IN=Indiana FL=Florida CA=California HA=Hawaii AUS=Australia UK=United Kingdom NZ=New Zealand.

Table 2.5. Samples within each branch of the phylogenetic tree for *cytochrome oxidase I* (*COI*). Branch numbers refer to the circled number on the branch in the phylogenetic tree (Figure 2.1). Samples within each branch exhibited sequences of 100 percent identity. GenBank® (NCBI) accession numbers are listed if applicable.

Branch	Sample ID/Accession Number	Species	Locality
Branch 1	C1-3	*Lucilia cuprina*	South Africa
	S2-4	*L. cuprina*	South Africa
	S2-11	*L. cuprina*	South Africa
	S2-12	*L. cuprina*	South Africa
	S3-1	*L. cuprina*	South Africa
	S2-3	*L. cuprina*	South Africa
	S2-1	*L. cuprina*	South Africa
	S2-7	*L. cuprina*	South Africa
	S2-2	*Lucilia sericata*	South Africa
	AB112863.1	*L. cuprina*	Perth, Australia
	EU626528.1	*L. cuprina*	Noordhoek, South Africa
Branch 2	AJ417711.1	*L. cuprina*	Tororo, Uganda
Branch 3	AJ417710.1	*L. cuprina*	Queensland, Australia
Branch 4	S2-10	*L. cuprina*	South Africa
Branch 5	C1-2	*L. cuprina*	South Africa
Branch 6	S2-9	*L. cuprina*	South Africa
	FL_Genome	*L. cuprina*	Florida, USA
	S2-6	*L. cuprina*	South Africa
	S2-13	*L. cuprina*	South Africa
	FJ650544.1	*L. cuprina*	Florida, USA
	AJ417705.1	*L. cuprina*	Hawaii, USA
Branch 7	DQ453496.1	*L. cuprina*	Hawaii, USA
Branch 8	FJ650560.1	*L. cuprina*	Florida, USA
Branch 9	C1-1	*L. sericata*	South Africa
	AJ422212.1	*L. sericata*	Somerset, England, UK
	C1-4	*L. sericata*	South Africa
	AJ417714.1	*L. sericata*	Bristol, England, UK
	AB112843.1	*L. sericata*	Graaff-Reinet, South Africa
Branch 10	IN-10	*L. cuprina*	Greenwood, IN, USA
	INLs-10	*L. sericata*	Greenwood, IN, USA
	INLs-8	*L. sericata*	Greenwood, IN, USA
	INLs-4	*L. sericata*	Indianapolis, IN
	INLc-2	*L. cuprina*	Indianapolis, IN, USA
	IN-1	*L. cuprina*	Indianapolis, IN, USA
	IN-2	*L. cuprina*	Franklin, IN, USA
	IN-9	*L. cuprina*	Greenwood, IN, USA
	IN-11	*L. cuprina*	Indianapolis, IN, USA
	IN-12	*L. cuprina*	Franklin, IN, USA
	IN-13	*L. cuprina*	Franklin, IN, USA
	INLs-3	*L. sericata*	Indianapolis, IN, USA
	IN-14	*L. cuprina*	Franklin, IN, USA
	INLc-1	*L. cuprina*	Greenwood, IN, USA
Branch 11	INLs-5	*L. sericata*	Greenwood, IN, USA
Branch 12	INLs-9	*L. sericata*	Greenwood, IN, USA
Branch 13	AB112850.1	*L. sericata*	Graaff-Reinet, South Africa
Branch 14	AJ417713.1	*L. sericata*	South Island, New Zealand
Branch 15	C1-5	*L. sericata*	South Africa
Branch 16	AF295550.1	*Phormia regina*	California, USA
Branch 17	MINK00000000.1	*P. regina*	Indiana, USA
Branch 18	NC_019634.1	*Chrysomya rufifacies*	Western Australia

2.3.2 *Period* (*Per*)

Analysis of the nuclear gene *Period* (*Per*) indicates a lack of resolution due to paraphyletic relationships (Figure 2.2). Like in *COI,* carrion (C) and sheep (S) are not only located within the same clade, but in some instances, they share identical sequences (i.e. Figure 2.2 branch 7) (Table 2.6). Further, *L. sericata* and *L. cuprina* are again paraphyletic as they are not defined by separate clades. Interestingly, Indiana *L. cuprina* specimens again appear to be more closely related to *L. sericata* than other *L. cuprina* specimens.

The Florida *L. cuprina* sample falls within a branch and clade that strictly contains other *L. cuprina* specimens. This result further emphasizes the apparent uniqueness of Indiana *L. cuprina* from *L. cuprina* collected in other locations in the United States. While the Florida sample appears to have an *L. sericata*-like mitochondrial haplotype for *COI,* analysis of *Per* indicates it possesses a typical *L. cuprina* haplotype for this gene.

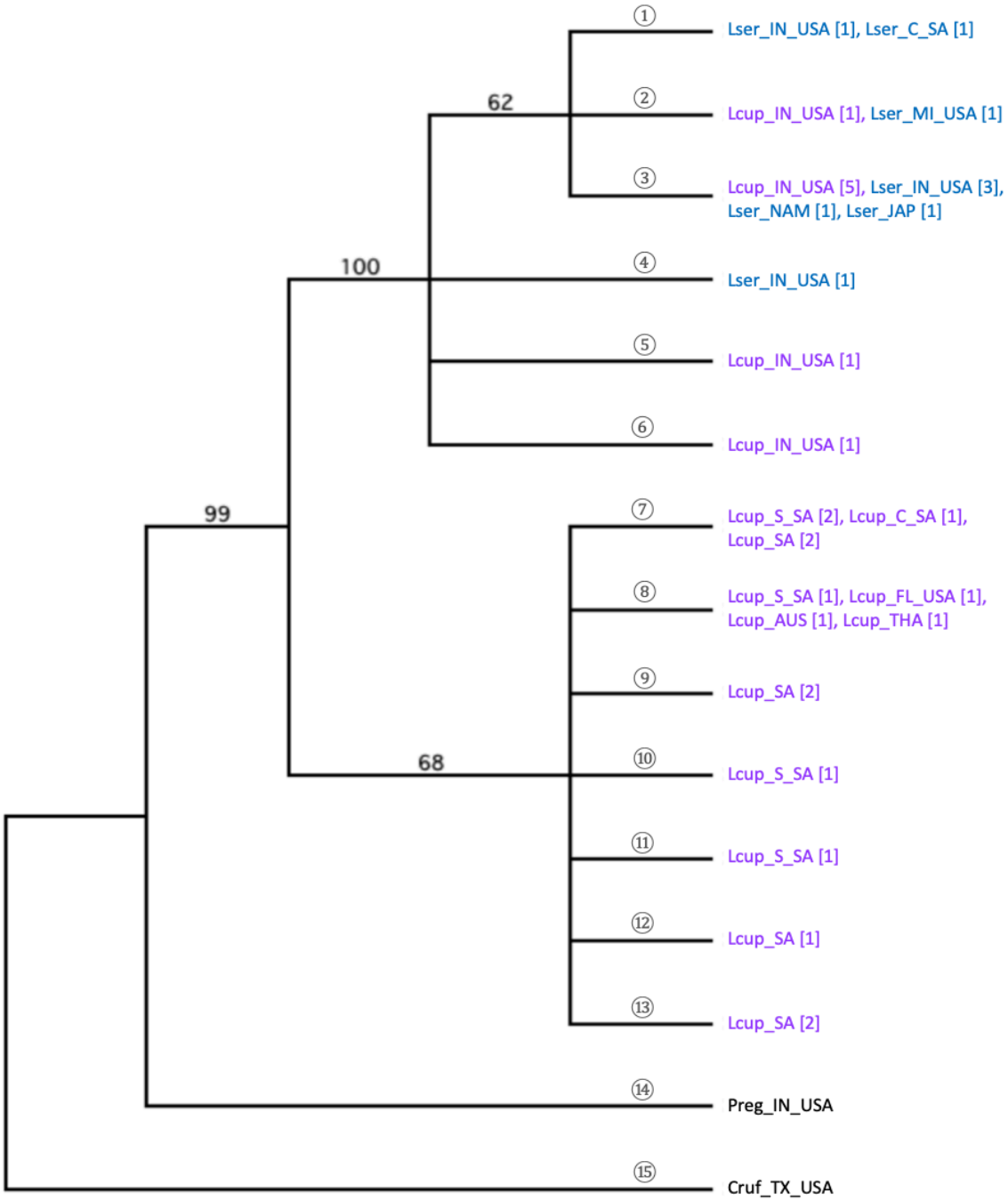

Figure 2.2. Maximum parsimony bootstrap (1,000 replicates) consensus tree based on *Period* (*Per*) sequence. Numbers on branches indicate percent bootstrap support. Bracketed numbers indicate the number of specimens included in each group. Group descriptions include species, location, and of those collected in South Africa, whether it came from live sheep or carrion. Each branch is labeled with a circled number for more in-depth description of samples included (Table 2.6). Lcup=*Lucilia cuprina* Lser=*Lucilia sericata* Preg=*Phormia regina* Cruf=*Chrysomya rufifacies* C=carrion S=sheep SA=South Africa USA=United States IN=Indiana MI=Michigan FL=Florida TX=Texas JAP=Japan NAM=Namibia THA=Thailand AUS=Australia.

Table 2.6. Samples within each branch of the phylogenetic tree for *Period* (*Per*). Branch numbers refer to the circled number on the branch in the phylogenetic tree (Figure 2.2). Samples within each branch exhibited sequences of 100 percent identity. GenBank® (NCBI) accession numbers are listed if applicable.

Branch	Sample ID/Accession Number	Species	Locality
Branch 1	INLs-6	*Lucilia sericata*	Franklin, IN, USA
	C1-4	*L. sericata*	South Africa
Branch 2	JN792856.1	*L. sericata*	Michigan, USA
	IN-11	*Lucilia cuprina*	Indianapolis, IN, USA
	IN-9	*L. cuprina*	Greenwood, IN, USA
Branch 3	INLs-3	*L. sericata*	Indianapolis, IN, USA
	IN-4	*L. cuprina*	Franklin, IN, USA
	INLs-7	*L. sericata*	Franklin, IN, USA
	INLc-2	*L. cuprina*	Franklin, IN, USA
	INLs-1	*L. sericata*	Indianapolis, IN, USA
	JN792833.1	*L. sericata*	Iwate, Japan
	JN792836.1	*L. sericata*	Possession Island, Namibia
	IN-2	*L. cuprina*	Franklin, IN, USA
	IN-7	*L. cuprina*	Franklin, IN, USA
	IN-10	*L. cuprina*	Greenwood, IN, USA
Branch 4	INLs-9	*L. sericata*	Greenwood, IN, USA
Branch 5	INLc-1	*L. cuprina*	Greenwood, IN, USA
Branch 6	IN-1	*L. cuprina*	Indianapolis, IN
Branch 7	C1-3	*L. cuprina*	South Africa
	JN792805.1	*L. cuprina*	Durban, South Africa
	JN792809.1	*L. cuprina*	Durban, South Africa
	S2-8	*L. cuprina*	South Africa
	S2-1	*L. cuprina*	South Africa
Branch 8	S9-9	*L. cuprina*	South Africa
	JRES01000940.1	*L. cuprina*	Australia
	FL_Genome	*L. cuprina*	Florida, USA
	JN792819.1	*L. cuprina*	Chang Mai, Thailand
Branch 9	JN792814.1	*L. cuprina*	Grahamstown, South Africa
	JN792815.1	*L. cuprina*	Grahamstown, South Africa
Branch 10	S2-3	*L. cuprina*	South Africa
Branch 11	S2-6	*L. cuprina*	South Africa
Branch 12	JN792791.1	*L. cuprina*	Cape Town, South Africa
Branch 13	JN792786.1	*L. cuprina*	Bloemfontein, South Africa
Branch 14	MINK00000000.1	*Phormia regina*	Indiana, USA
Branch 15	Cruf_Genome	*Chrysomya rufificaies*	Texas, USA

2.3.3 *28s*

Analysis of the nuclear *28s* rRNA gene (*28s*) presented the least variation of the three genes utilized in this study (Figure 2.3). Many samples from various species, locations, and collection methods (carrion versus sheep), share identical sequences (i.e. Figure 2.3 branches 1 and 3) (Table 2.7). While many of the Indiana *L. cuprina* specimens share a sequence with *L. sericata* specimens (Figure 2.3 branch 3), one Indiana *L. cuprina* specimens falls within a primarily *L. cuprina* clade. The same South African *L. sericata* specimen that was on outlier

falling in an otherwise all *L. cuprina* clade in the *COI* analysis (Figure 2.1) follows the same pattern in this analysis (Figure 2.3 branch 1). Additionally, the same South African *L. sericata* sample (C1-5) that appeared to be more distantly related to the other *L. sericata* and *L. cuprina* specimens in the *COI* analysis follows the same pattern in this analysis (Figure 2.3 branch 4).

Interestingly, a *C. rufifacies* sequence obtained from GenBank® (National Center for Biotechnology Information; NCBI) falls within the ingroup when analyzing *28s.* This further emphasizes the inability of *28s* to resolve not only sister species *L. cuprina* and *L. sericata,* but also its ability to resolve more distantly related species. The *28s* sequences analyzed seem to be more conserved within the family Calliphoridae and even within the order Diptera than those of *COI* and *Per,* as demonstrated by the results of importing each *28s* sequence BLASTn® (NCBI) for query against a nucleotide database. Based on the first 100 BLAST hits, the sequence haplotype represented by branch 1 shares 100% sequence identity with *L. cuprina* and a *L. cuprina x L. sericata* hybrid, as well as 99.43-99.73% identity with species within genera *Lucilia, Chrysomya, Calliphora, Rhyncomya,* and *Hemigymnochaeta* (Supplementary Table 1). Branch 2 shares 100% sequence identity with *Lucilia eximia, L. sericata, Lucilia thatuna, Lucilia mexicana, Lucilia cluvia,* and *Chrysomya chloropyga,* as well as 99.73% identity with species within genera *Calliphora, Lucilia, Chrysomya, Rhyncomy*a, and *Hemigymnochaeta.* Branch 3 shares 100% sequence identity with many species within *Calliphora, Chrysomya, Lucilia, Sarconesia, Chloroprocta,* and *Onesia*. Branch 4 shares 100% sequence identity with *Lucilia papuensis, Hemipyrellia fernandica, Hypopygiopsis infumata, and Hypopygiopsis violacea,* as well as 99.20-99.73% identity with species within genera *Hemipyrellia, Chrysomya, Calliphora, Lucilia, Melanomya,* and *Sarconesia.* Branch 5 shares 100% sequence identity with *C. rufifacies* and 99.47-99.73% identity with species within genera *Calliphora, Chrysomya, Lucilia, Sarconesia, Onesia*, and *Cordylobia*. Branch 6 shares 100% sequence identity with *P. regina* and *Chrysomya nigripes* as well as 99.73% identity with species within genera *Cochliomya, Protophormia, Compsomyia, Chrysomya, Lucilia, Hemilucilia, Rhinophora, Tachina, Sarconesia*, and *Chloroprocta*.

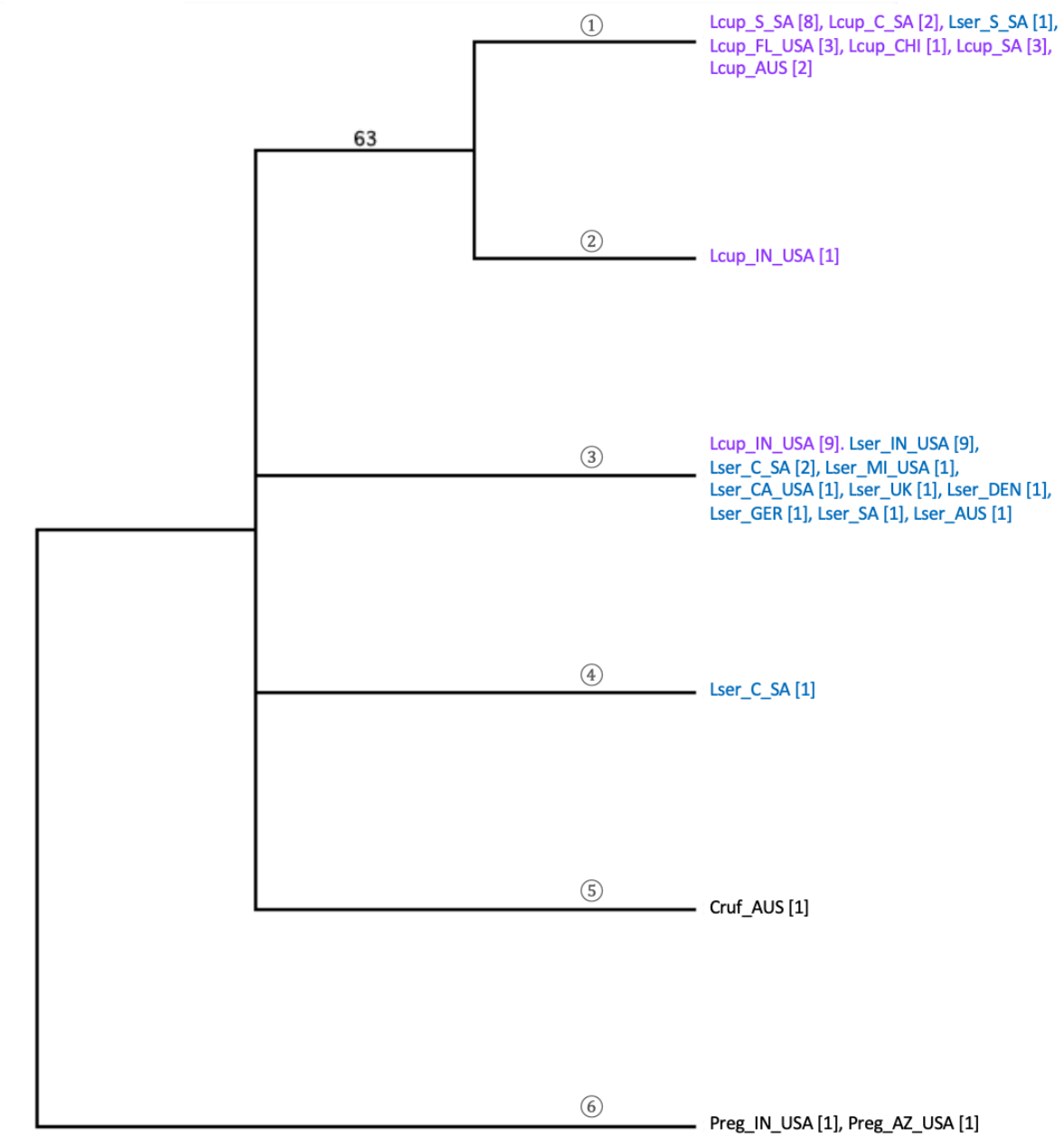

Figure 2.3. Maximum parsimony bootstrap (1,000 replicates) consensus tree based on *28s* sequence. Numbers on branches indicate percent bootstrap support. Bracketed numbers indicate the number of specimens included in each group. Group descriptions include species, location, and of those collected in South Africa, whether it came from live sheep or carrion. Each branch is labeled with a circled number for more in-depth description of samples included (Table 2.7). Lcup=*Lucilia cuprina* Lser=*Lucilia sericata* Preg=*Phormia regina* Cruf=*Chrysomya rufifacies* C=carrion S=sheep SA=South Africa USA=United States IN=Indiana FL=Florida CA=California AZ=Arizona CHI=China AUS=Australia GER=Germany DEN=Denmark UK=United Kingdom.

Table 2.7. Samples within each branch of the phylogenetic tree for (*28s*). Branch numbers refer to the circled number on the branch in the phylogenetic tree (Figure 2.3). Samples within each branch exhibited sequences of 100 percent identity. GenBank® (NCBI) accession numbers are listed if applicable.

Branch	Sample ID/Accession Number	Species	Locality
Branch 1	S9-9	*Lucilia cuprina*	South Africa
	S3-1	*L. cuprina*	South Africa
	S2-12	*L. cuprina*	South Africa
	FL_Genome	*L. cuprina*	Florida, USA
	S2-11	*L. cuprina*	South Africa
	S2-7	*L. cuprina*	South Africa
	S2-6	*L. cuprina*	South Africa
	S2-3	*L. cuprina*	South Africa
	S2-10	*L. cuprina*	South Africa
	S2-2	*Lucilia sericata*	South Africa
	C1-2	*L. cuprina*	South Africa
	KY197796.1	*L. cuprina*	China
	FJ650542.1	*L. cuprina*	Florida, USA
	EU626539.1	*L. cuprina*	Cape Town, South Africa
	FJ650541.1	*L. cuprina*	Florida, USA
	FJ650533.1	*L. cuprina*	Artois, CA, USA
	EU626550.1	*L. cuprina*	Noordhoek, South Africa
	C1-3	*L. cuprina*	South Africa
	JN792705.1	*L. cuprina*	Hornsby Heights, Australia
	FR719302.1	*L. cuprina*	Grahamstown, South Africa
	AJ417709.1	*L. cuprina*	Queensland, Australia
Branch 2	IN-5	*L. cuprina*	Franklin, IN, USA
Branch 3	INLs-3	*L. sericata*	Indianapolis, IN, USA
	INLs-6	*L. sericata*	Franklin, IN, USA
	INLs-5	*L. sericata*	Greenwood, IN, USA
	KR133394.1	*L. sericata*	Barsbüttel, Germany
	INLc-2	*L. cuprina*	Indianapolis, IN, USA
	JN792779.1	*L. sericata*	Michigan, USA
	EU626535.1	*L. sericata*	Bothasig, South Africa
	JN792746.1	*L. sericata*	Seaford, Australia
	AJ300141.1	*L. sericata*	Los Angeles, CA, USA
	AJ300140.1	*L. sericata*	Sjelland, Denmark
	AJ300139.1	*L. sericata*	North Somerset, England, UK
	INLs-10	*L. sericata*	Greenwood, IN, USA
	INLs-8	*L. sericata*	Greenwood, IN, USA
	INLs-7	*L. sericata*	Franklin, IN, USA
	INLs-4	*L. sericata*	Indianapolis, IN, USA
	INLs-1	*L. sericata*	Indianapolis, IN, USA
	IN-12	*L. cuprina*	Franklin, IN, USA
	IN-9	*L. cuprina*	Greenwood, IN, USA
	IN-4	*L. cuprina*	Franklin, IN, USA
	IN-2	*L. cuprina*	Franklin, IN, USA
	IN-1	*L. cuprina*	Indianapolis, IN, USA
	IN-7	*L. cuprina*	Franklin, IN, USA
	IN-11	*L. cuprina*	Indianapolis, IN, USA
	IN-14	*L. cuprina*	Franklin, IN, USA
	C1-1	*L. sericata*	South Africa
	INLs-2	*L. sericata*	Franklin, IN, USA
	C1-4	*L. sericata*	South Africa

Table 2.7 continued

Branch 4	C1-5	*L. sericata*	South Africa
Branch 5	JQ246609.1	*Chrysomya rufifacies*	Australia
Branch 6	MINK00000000.1	*Phormia regina*	Indiana, USA
	AF366685.1	*P. regina*	Arizona, USA

2.3.4 BLAST Results

The *COI* sequence of the majority of Indiana and South Africa *L. cuprina* and *L. sericata* specimens collected for this study resulted in top 100 BLAST hits within the genus *Lucilia*, with most hits being *L. cuprina* or *L. sericata.* A few *COI* sequence haplotypes contained a single *Hemipyrellia ligurriens* hit towards the bottom of the top 100 hits, albeit still with a percent sequence identity between 99 and 100%. The *28s* BLAST results varied widely among a number of genera as previously discussed. *Per* BLAST hits primarily consisted of *L. cuprina* and *L. sericata,* with lower hits within the top 100 straying to different species within Lucilia, as well as *He. fernandica.* In this instance, the *Hemipyrellia* species only shared at most 91.13% sequence identity with our *Per* sequences.

Due to its unique positions in the phylogenetic trees, the sequences of sample C1-5, which was identified by our laboratory to be *L. sericata,* were investigated. Upon BLAST analysis of this sample's *COI* sequence, the four top hits are *He. fernandica* with percent sequence identities of 97.71-97.99%. The fourth and fifth hits with lower percent identities are *He. pulchra,* and the sixth is *He. ligurriens.* BLAST analysis of the *28s* sequence of sample C1-5 resulted in many top hits with 100% sequence identity including *L. papeunsis, He. fernandica, Hy. infumata, and Hy. violacea. Lucilia sericata* appears as the 25th hit with 99.47% sequence identity. This sample did not amplify for *Per.* A quick, less in-depth phylogenetic analysis (www.Phylogeny.fr One-Click) of all previously analyzed sequences plus known *Hemipyrellia* sequences obtained from GenBank® places C1-5 in clades with known *He. fernandica* sequences (Figure 2.4) [48, 49].

Additionally, *L. sericata* sample S2-2 is the only *L. sericata* sample within an otherwise all *L. cuprina* clade in the trees for *COI* and *28s*, so its sequences were further investigated. S2-2 is the only sample we identified as *L. sericata* that BLASTs to *L. cuprina* rather than *L. sericata* for *COI.* Like many samples, S2-2 BLASTs to several species within a number of genera for *28s,*

though its top hits with 100% sequence identity are *L. cuprina* and *L. cuprina* x *L. sericata* hybrids.

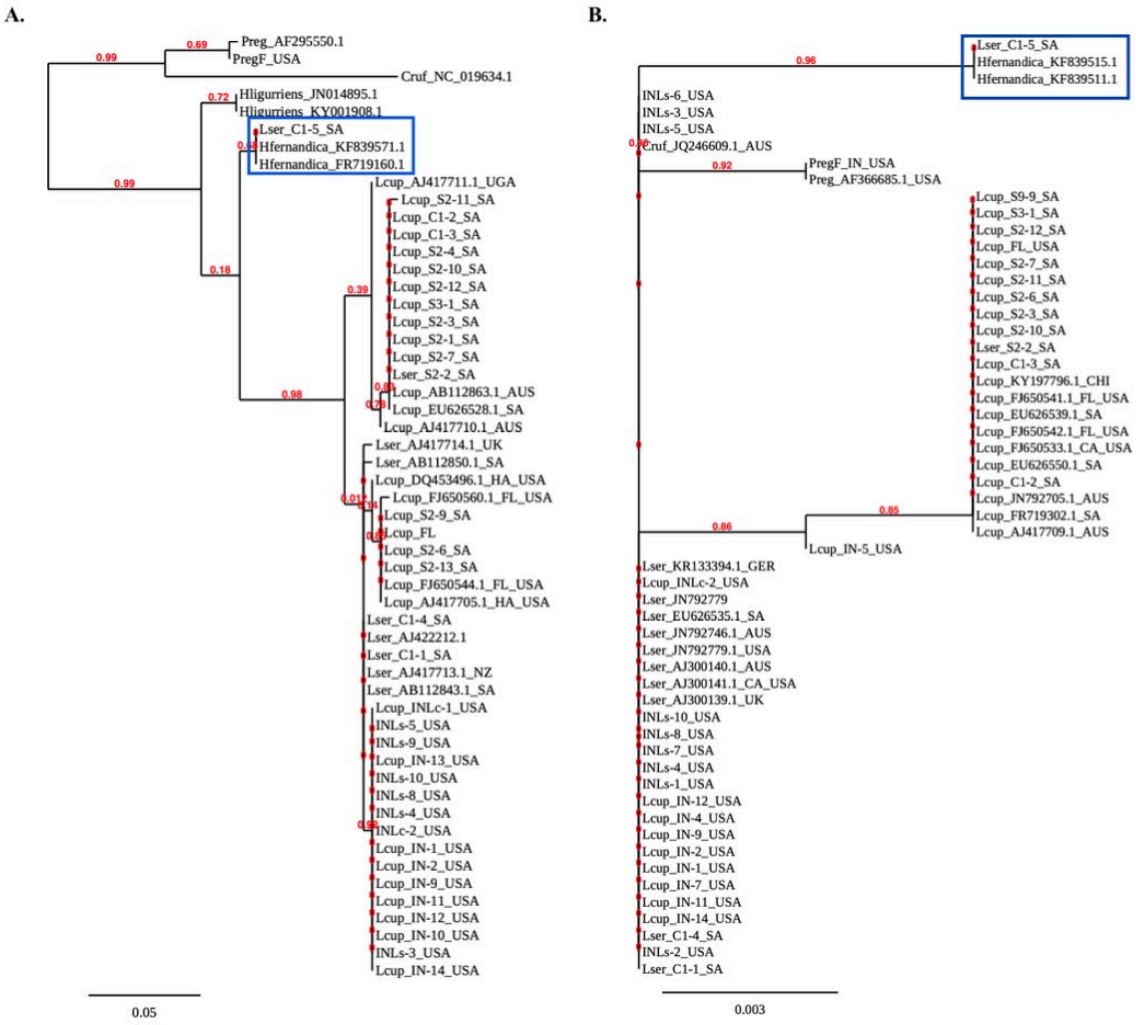

Figure 2.4. Phylogenetic trees resulted from a quick analysis utilizing www.Phylogeny.Fr (A.) *COI* (B.) *28s.* Clades of interest are outlined in blue. Both (A.) and (B.) contain phylogenetic trees in which sample C1-5 falls within a clade with known *He. fernandica* sequences. Specimens are either labeled with their internal ID or GenBank® accession number.

2.4 Discussion

The BLAST results indicate that, despite a possible outlier, all of our collected specimens were correctly identified to be either *L. cuprina, L. sericata,* or some hybrid of the two. Though some *COI* sequences resulted in a BLAST hit to *He. ligguriens, Hemipyrellia* is only found in the eastern hemisphere [16] and is mainly restricted to Asian regions [50], so it is unlikely that any of the Indiana or South African samples collected are truly *He. ligurriens.* Additionally, there are reports of intermixing *COI* haplotypes of *L. c. cuprina* and *He. ligurriens,* so the presence of this species in the BLAST results is not entirely surprising [19]. Further, BLAST results indicated that some of our *Per* sequences shared similarities with *He. fernandica. Hemipyrellia fernandica* is an Afrotropical species with morphology similar to, but distinguishable from, species within *Lucilia* [27]. *Hempyrellia fernandica* has been collected previously in South Africa [15], but since the *Lucilia* species overall had much higher percent identities than *He. fernandica*, our South African samples are not likely to be *Hemipyrellia.*

The exception to our *Lucilia* identifications may be a single sample (C1-5) that stands as an outlier in the analyses of *COI* and *28s*, and did not amplify for *Per.* This adult sample was identified as *L. sericata* by our laboratory, but its BLAST results as well as its location on the trees indicate some discrepancy. The phylogenetic trees indicate a more distant relationship between this sample and the other *Lucilia* samples analyzed. BLAST results indicate that sample C1-5 possesses sequence similarities to several species within the genus *Hemipyrellia* including *He. fernandica, He. pulchra, and He. ligguriens* for *COI.* For *28s,* the sequence was similar to *L. papeunsis, He. fernandica, Hy. infumata, and Hy. violacea. Hemipyrellia pulchra* is restricted to Asian regions [51, 52]. The genus *Hypopgiopsis* is only found in Asian and Australasian regions [16, 53]. The same is true of *L. papeunsis* [54, 55]. As discussed previously, this sample is not likely to be *He. ligurriens*. It is possible that it is *He. fernandica* based on a few samples that have been previously collected in South Africa.

Due to its uniqueness in our phylogenetic trees, and a quick phylogenetic analysis that placed it in a clade with *He. fernandica* rather than *Lucilia,* it is possible that this particular sample (C1-5) was misidentified to be *L. sericata* when it is truly *He. fernandica.* Because we are only examining a few genes and not the entire genome, it is difficult to conclude based on the two C1-5 sequences whether the initial morphological species identification was incorrect or not. It is possible that the sample truly is *L. sericata,* which would further highlight the dangers of

using these genes to make species identifications in a forensic context. Despite the possibility of sample C1-5 being misidentified based on its sequence data, it remains labeled as *L. sericata* in our trees and tables, as that was our conclusion based on morphology. This voucher and all others will be reexamined by an expert to confirm species identifications. Due to lack of laboratory access because of the COVID-19 pandemic, this species confirmation has yet to take place.

Additionally, sample S2-2, which we identified to be *L. sericata,* raised suspicion by being the only *L. sericata* sample within otherwise *L. cuprina* clades. Top BLAST hits were *L. cuprina* and a known *L. cuprina* x *L. sericata* hybrid for *28s*. However, as shown, *28s* sequences seemed to be more conserved and less species-specific. Top hits for the *COI* sequence of sample S2-2 were *L. cuprina* rather than *L. sericata.* However, several of our *L. cuprina* samples BLAST to *L. sericata* for *COI.* Further analysis is required to determine whether this specimen is truly *L. sericata,* a hybrid, or a misidentification. Like with sample C1-5, S2-2 remains labeled with our original species identification of *L. sericata* until reidentification can be completed by an expert.

Our *COI* analysis is concordant with the majority of related studies that conclude that *COI* cannot be relied on to resolve between *L. cuprina* and *L. sericata.* A 2007 study cautions against the use of *COI* to identify *Lucilia* species due to lack of reciprocal monophyly within the genus [10]. That is, if species are interspersed throughout a number of clades, using this information to make an identification can be uncertain. Not only does this study describe the same lack of reciprocal monophyly seen in our data, but it also describes a surprising issue of *L. cuprina* resemblance to *Hemipyrellia,* as seen with sample C1-5. Many researchers question the separation of the genus *Hemipyrellia* from *Lucilia* [10, 16]. Still, some research groups support a separation of *Hemipyrellia* and *Lucilia* [27]. However, this issue is minimal for studies and forensic applications in North America because no *Hemipyrellia* species exist in North America [19].

A number of additional studies conclude that *COI* analysis alone does not provide strong support for the differentiation between *L. cuprina* and *L. sericata.* A 2016 study reported paraphyly between *L. cuprina* and *L. sericata* regarding *COI,* and note that the power of *COI* data to discriminate between species increases when used in conjunction with additional genes [16]. Further, many studies reported *L. cuprina* specimens with *L. sericata*-like *COI* haplotypes, indicating the possibility of a hybrid status in these specimens that appear morphologically to be

L. cuprina [15, 18, 28]. The novel finding that Indiana *L. cuprina* specimens possess a unique *L. sericata*-like haplotype combination indicates that Indiana *L. cuprina* are of a unique lineage and perhaps suggests modern hybridization in this midwestern region. Since it is now evident that there is more variation to *COI* than previously believed, it is especially important to be cautious when using this gene for species identification. Because the presence of *L. cuprina* has only been identified in Indiana in the last few years, and the presence of this species in this region continues to be seen (i.e. *L. cuprina* is still collected in Indiana in 2019), then it is likely that this species has become locally adapted, and perhaps has done so due to a more recent hybridization. Without greater depth of sequencing, it is not possible to conclude at this point. However, this research highlights the need for more specimen sampling and sequencing.

A 2010 study argues that although *L. cuprina* is monophyletic in relation to *L. sericata* for *COI,* two distinct clades of *L. cuprina* appear that distinctly separate those with *L. sericata*-like haplotypes [19, 56]. Our *COI* data contradict this conclusion, as *L. cuprina* is located in two places within *L. sericata* clades. Were our Indiana *L. cuprina* samples omitted, our results would appear to generally follow the trend these studies describe. However, the addition of Indiana *L. cuprina* samples breaks this trend, again highlighting the apparent uniqueness of the Indiana specimens. The specimens that were utilized in the 2010 study were relatively coastal, collected in places such as Florida and California. Our data indicate that *L. cuprina* from the midwestern and landlocked state of Indiana possess different haplotypes than those presented in the aforementioned study. While the Indiana *L. cuprina* samples possess a *L. sericata*-like mitochondrial haplotype, they fall within a different clade than other *L. cuprina* samples with *L. sericata*-like haplotypes.

Our *28s* and *Per* sequences on their own were not able to differentiate between *L. cuprina* and *L. sericata,* nor could they distinguish between myiasis or carrion preference. *28s* sequences could not reliably be used to distinguish *Lucilia* from other blow flies, let alone distinguish *L. cuprina* and *L. sericata.* However, in our phylogenetic tree, Indiana *L. cuprina* were more closely related to *L. sericata* than other *L. cuprina.* Previous studies that investigated *28s* and *Per* generally reported results when concatenated with each other or with additional genes. For example, a study found that a concatenated analysis of *28s* and *Per* sequences was able to resolve *L. cuprina* and *L. sericata* [15, 16]. Further, they found that a concatenated analysis of *COI, 28s,* and *Per* resolved the two species. However, these analyses did not include specimens

from Hawaii, which had been previously identified to contain hybrids and may have influenced the results. In our single analysis, *Per* could not resolve species within *Lucilia.* However, our *Per* and *28s* data present new evidence that Indiana *L. cuprina* possess a *L. sericata*-like haplotype for this nuclear gene.

Genetic evidence of modern, natural interbreeding between *L. cuprina* and *L. sericata* was first reported in 2013 [15]. Two *L. cuprina* specimens with *L. sericata*-like *COI* haplotypes also possessed *28s* and *Per* haplotypes that differed from typical *L. cuprina.* These two unique samples were collected in Zimbabwe and Thailand. Our data provide the first indication that modern hybridization could have taken place in Indiana *L. cuprina*, which is likely the first evidence of such hybridization in North America. Again, a greater depth of sequencing is required to investigate this possibility.

While *L. cuprina* and *L. sericata* have been a focus of research due to their prevalence in North America and in forensic contexts, these species are not the only problematic pair within *Lucilia.* Sister-species *Lucilia caesar* and *Lucilia illustris* also possess varying levels of paraphyly that make molecular identification and discrimination difficult [19, 57, 58]. These *Lucilia* species exhibit paraphyly in several of the same genes as *L. cuprina* and *L. sericata* (*COI* and *28s*), as well as additional genes such as *ITS-2* and *bicoid* [57, 59]. The presence of molecular identification issues in this *Lucilia* species pair in addition to *L. cuprina* and *L. sericata* emphasizes a need for further sequence investigation of *Lucilia* as a whole. Interestingly, *L. caesar* and *L. illustris* are important agents of myiasis like *L. cuprina* and *L. sericata* [58, 60], correlating sister-species with paraphyletic relationships to myiasis behavior.

While the *Period* gene was hypothesized to aid in discriminating species within *Lucilia,* it is a circadian clock gene [15, 16, 61] and likely expresses differently under different environmental conditions. *Lucilia cuprina* is relatively new to Indiana [37] and will likely continue to move north due to changes in climate. As the species moves up in longitude and experiences changes in day and night cycles, *Per* may adapt locally and account for some inter-species sequence differences. *Period* was isolated in *L. cuprina* relatively recently and has only been investigated in a handful of studies [61], emphasizing a need for further investigation.

Though the current study indicates unique haplotype combinations in Indiana *L. cuprina,* the study has its limitations. While the *COI* primer pair used here has been used in similar studies and allowed comparison of results, there are a number of available *COI* primers available that

amplify different stretches of the gene that should be investigated in this context. Additionally, a concatenated analysis of the three genes utilized here may provide further insight. Further, a relatively small number of samples were utilized that were collected from similar urban parks, so future studies may benefit by expanding sample size and collection environments.

Based on incongruence found in Indiana specimens, our data indicate that overall, *COI, 28s,* and *Per* sequences are insufficient to distinguish between *L. cuprina* and *L. sericata.* Additionally, these three genes were unable to resolve differences between specimens collected from carrion versus from live sheep. Our original hypothesis was that carrion versus myiasis behavior preference would indicate a subspecies relationship within *L. cuprina.* Though it appears that this distinction is not apparent in the genes we analyzed, our results do imply a separation of lineages that is yet to be explained. While previous studies suggest that *L. c. cuprina* aligns with a *L. sericata*-like lineage and *L. c dorsalis* is more distantly related to *L. sericata,* our data suggest that Indiana *L. cuprina* belongs to a different *L. sericata*-like lineage.

2.5 Conclusions

Our results indicate that *L. cuprina* from Indiana are of a unique lineage, possessing a combination of *COI, 28s, and Per* haplotypes not previously reported in the United States. Not only is *L. cuprina* relatively new to Indiana [37], but it also appears to possibly be undergoing modern hybridization with *L. sericata* in the region. As *L. cuprina* likely continues to migrate to new areas within North America due to changes in climate, genetic differences such as the ones found in Indiana specimens may become more prominent.

Overall, our data support the previously reported conclusion that *COI* should not be used, at least not alone, as a species identifier within *Lucilia*. However, our study also indicates that *28s* and *Per* sequences are not sufficient to resolve *Lucilia* species or carrion versus myiasis behaviors. This study highlights the dangers that may be present within certain blow fly genera that must be accounted for when choosing DNA-based identification methods in research or forensic contexts. Future directions may include concatenated analyses of *COI, 28s,* and *Per*, as well as more specific and in-depth investigations of the unique genomes of *L. cuprina* from Indiana, USA. Additionally, further research should include additional *Hemipyrellia* sequences to better investigate the relationship between *Hemipyrellia* and *Lucilia.*

CHAPTER 3. POPULATION STRUCTURE OF HOUSE FLIES IN ARIZONA, USA

3.1 Introduction

The house fly *Musca domestica* L. (Diptera: Muscidae) is classified as a "filth fly" due to its propensity to thrive in "filthy" substrates such as feces, garbage, food waste, and sewage sludge [62, 63]. House flies are said to be synanthropic, as they live and thrive in environments closely associated with human activity. They can be found in abundance in areas such as restaurants, poultry farms, slaughter houses, and livestock farms, which is not only a nuisance, but a potential source of disease in humans and animals [64, 65]. House flies have been traced back to the origin of humanity [66] and their dispersal and ability to transport pathogens continues to be a topic of research. Today, it is known that the house fly can be a vector for over 100 pathogens including viruses, bacteria, fungi, protozoans, and metazoans [67]. Researchers aim to understand the house fly's status as a disease vector and quantify its dispersal capabilities in hopes to shed light on control and prevention measures. Because of their distribution and association with humans, house flies are considered forensically relevant [1].

A variety of methodologies have been previously utilized to study house fly populations. One such method includes mark/release/recapture, which typically involves specimens being chemically marked and released in the environment in question, followed by collections at various distances to quantify what percentage of marked flies are found again [68]. One such study conducted in Phoenix, Arizona indicated that *M. domestica* are capable of dispersing as far as 5 to 20 miles, but the majority of the population limits dispersal to 0.5 to 2.0 miles [69]. The authors cite that house flies have an inherent tendency to wander, though their dispersal may be influenced by factors including population pressure, differences in site attractiveness, geographical barriers, and preferential movement. An additional mark/release/recapture study indicated that house flies released at a poultry farm had a flight range potential of 7 km (4.35 miles), though this study was conducted in the Cameron Highlands region of Malaysia [70]. It is important to remember that studies have shown different *M. domestica* flight ranges in different regions [63, 69, 70].

A second potential method to study house fly populations as vectors of disease is to investigate the presence of antibiotic-resistant bacterial strains found within the flies. In 2010, a

research group suggested that the presence of antibiotic-resistant bacterial profiles in house flies in areas without routine antibiotic usage would suggest their dispersal from a facility with routine antibiotic use, such as a farming facilities [71]. Detection and comparison of *Enterococci* found in both house flies and pig manure revealed the flies contained the same multi-drug resistant clones as the manure, indicating the manure as the source of resistance in the flies [72].

An additional approach for investigating house fly populations is the use of population genetics. House flies are known to show an abundance of both phenotypic and genetic differences [73-75], making genetic comparisons of different populations probative. Early population genetic studies of *M. domestica* largely focused on allozymes [74, 76, 77] and mitochondrial markers [73, 78]. Though microsatellite markers have been developed for *M. domestica* [79-81], few studies have utilized them for analyses to date. One study that has utilized *M. domestica* microsatellites found significant genetic variation within, but little variation between, the populations surveyed, as well as a significant correlation between geographic and genetic distances [71]. A microsatellite survey of house flies from several global regions indicated that heterozygosity and allelic diversities were both homogeneous and substantial for all populations while mitochondrial diversities varied among regions [82].

To our knowledge, this study marks the first use of amplified fragment length polymorphism (AFLP) analysis methods on *M. domestica.* AFLP provides a relatively simple and inexpensive means of sampling an entire genome for polymorphism [83, 84]. The large quantity of loci data produced is useful for the inference of population structure [22, 85-88]. The current study investigates the population structure of *M. domestica* in the context of a civil lawsuit in Arizona, USA. Multiple plaintiffs whose homes fall within an approximate 7.0-mile radius of a local poultry and egg farm complained about flies emanating to their homes from the farm. Because of the house fly's ability to transmit pathogens, the homeowners filed a lawsuit against the poultry farm. The task at hand included utilizing genetic markers and population genetic analyses to determine whether or not house flies collected near the plaintiff's homes are genetically similar to or different from house flies collected from the poultry farm.

3.2 Materials and Methods

3.2.1 Arizona Samples

Fly Collection

Five fly specimens from each of 13 initial locations were collected and identified by board certified entomologist Douglas Seemann. The collection locations were analogous to the areas and distances associated with the plaintiffs of the case (Table 3.1). With the poultry farm as the starting point (site A), the collection sites ranged from 0.25 to 7.0 miles away from the farm in varying directions (Figure 3.1). The poultry farm was subsampled into sites Aa, Ab, and Ac. The majority of flies were collected over a span of two days with a sweep net using cat food as bait. Ten fly specimens were collected from each of two additional sites approximately seven months after the initial collections using the same bait and collection methods (Figure 3.2). These additional sites included a nearby dairy farm and a site near a University of Arizona research facility known as Biosphere 2.

Table 3.1. Summary of *Musca domestica* sampling locations within and surrounding the poultry farm in Arizona, USA.

Location ID	Collection Date	Site	Coordinates	Altitude
Aa	12/14/18	Lay House 3	33º29'25.78"N 112º57'00.39"W	1129'
Ab	12/14/18	Lay House 10	33º29'13.85"N 112º57'00.39"W	1125'
Ac	12/14/18	Pullet House	33º28'47.87"N 112º57'11.21"W	1115'
B	12/14/18	0.25 mile NE	33º29'37.03"N 112º56'40.56"W	1128'
C	12/14/18	0.5 mile NE	33º29'37.26"N 112º56'22.34"W	1125'
D	12/14/18	0.75 mile ESE	33º28'49.32"N 112º56'12.51"W	1097'
E	12/14/18	1.0 mile SE	33º28'29.24"N 112º56'12.64"W	1090'
F	12/14/18	1.5 mile SE	33º27'58.15"N 112º56'12.96"W	1072'
G	12/14/18	2.0 mile SW	33º27'59.80"N 112º58'48.65"W	1137'
H	12/15/18	3.0 mile ENE	33º30'50.25"N 112º53'59.68"W	1127'
I	12/15/18	4.0 mile SE	33º26'37.09"N 112º54'03.62"W	1019'
J	12/15/18	5.0 mile ESE	33º27'28.73"N 112º52'03.16"W	1059'
K	12/15/18	7.0 mile E	33º29'37.95"N 112º49'32.83"W	1139'
DF	7/14/19	Dairy Farm	33º32'06.66"N 112º53'08.65"W	1188'
BS	7/23/19	Biosphere 2	32º34'30.09"N 110º51'11.92"W	3848'

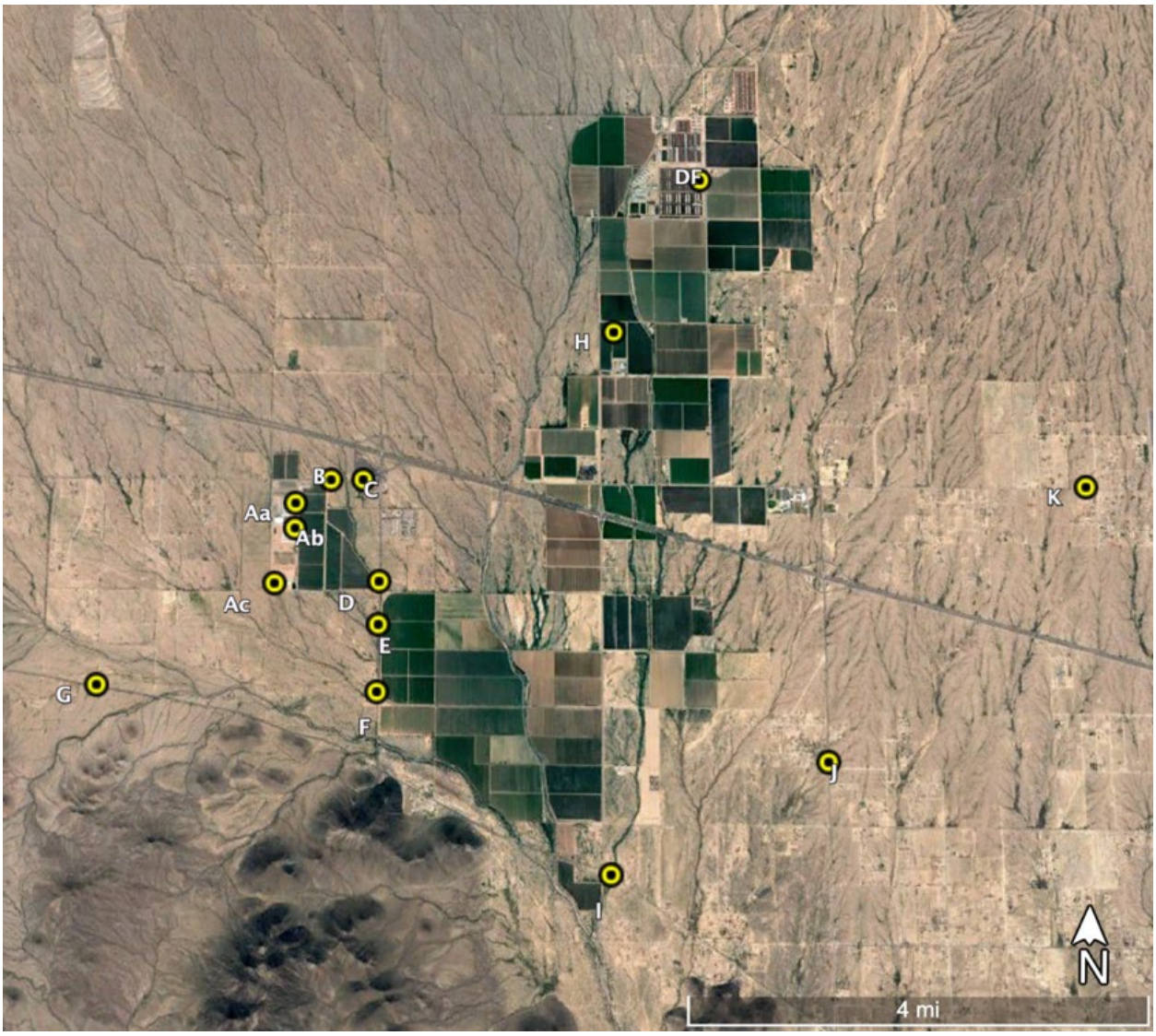

Figure 3.1. Map of fly collection sites within and surrounding the poultry farm in Arizona, USA.

Figure 3.2. Map demonstrating the distance in miles between the poultry farm (Aa) and the more distant sampling site Biosphere 2 (BS). The location of the nearby dairy farm (DF) is also marked. Several site markers were removed to reduce clutter due to the scale of the map.

Species Identification

Species identifications based on external morphology were performed by board certified entomologist Douglas Seemann. Of the 85 specimens collected, 83 were confirmed to be *M. domestica.* The two non- *M. domestica* specimens were identified as *Stomoxys calcitrans* and *Fania spp.* and were excluded from downstream analysis. Specimens were photographed, catalogued, preserved in ethanol, and securely shipped to Indiana University-Purdue University Indianapolis. A chain of custody transfer was documented upon receipt of the specimens.

DNA Extraction

DNA was extracted from fly heads using organic phenol-chloroform extraction. After decapitation, the heads digested in 200µL ChargeSwitch® lysis buffer (Invitrogen™) and 20µL 20mg/mL proteinase K (Invitrogen™) while incubating overnight at 56°C. To help expose and digest cells during this process, fly head tissue was smashed with sterile pipette tips and vortexed prior to incubation. 300µL phenol chloroform isoamyl alcohol (PCI) (25:24:1) (Thermo Fisher Scientific™) was added to each lysate. Sample tubes were inverted several times and centrifuged at 13,000rpm for 1 minute to allow separation into organic and aqueous layers. Extraction was continued following a standard organic phenol-chloroform protocol [89].

3.2.2 Purdue University Samples

Fly Collection

In order to provide an outgroup for comparative purposes, fly specimens collected in Indiana were added to the dataset (Table 3.2). These specimens were collected from an educational sheep barn at Purdue University in West Lafayette, Indiana (Figure 3.3). Flies were collected using a sticky trap and stored at -20°C.

Table 3.2. Summary of *Musca domestica* specimens collected in West Lafayette, Indiana, USA.

Specimen ID	Location ID	Collection Date	Site	Coordinates	Altitude
P2	Purdue	June '18	Purdue Sheep Barn	40º29'54"N 87º0'49.02"W	712'
P3	Purdue	June '18	Purdue Sheep Barn	40º29'54"N 87º0'49.02"W	712'
P5	Purdue	June '18	Purdue Sheep Barn	40º29'54"N 87º0'49.02"W	712'
P8	Purdue	June '18	Purdue Sheep Barn	40º29'54"N 87º0'49.02"W	712'
P11	Purdue	June '18	Purdue Sheep Barn	40º29'54"N 87º0'49.02"W	712'

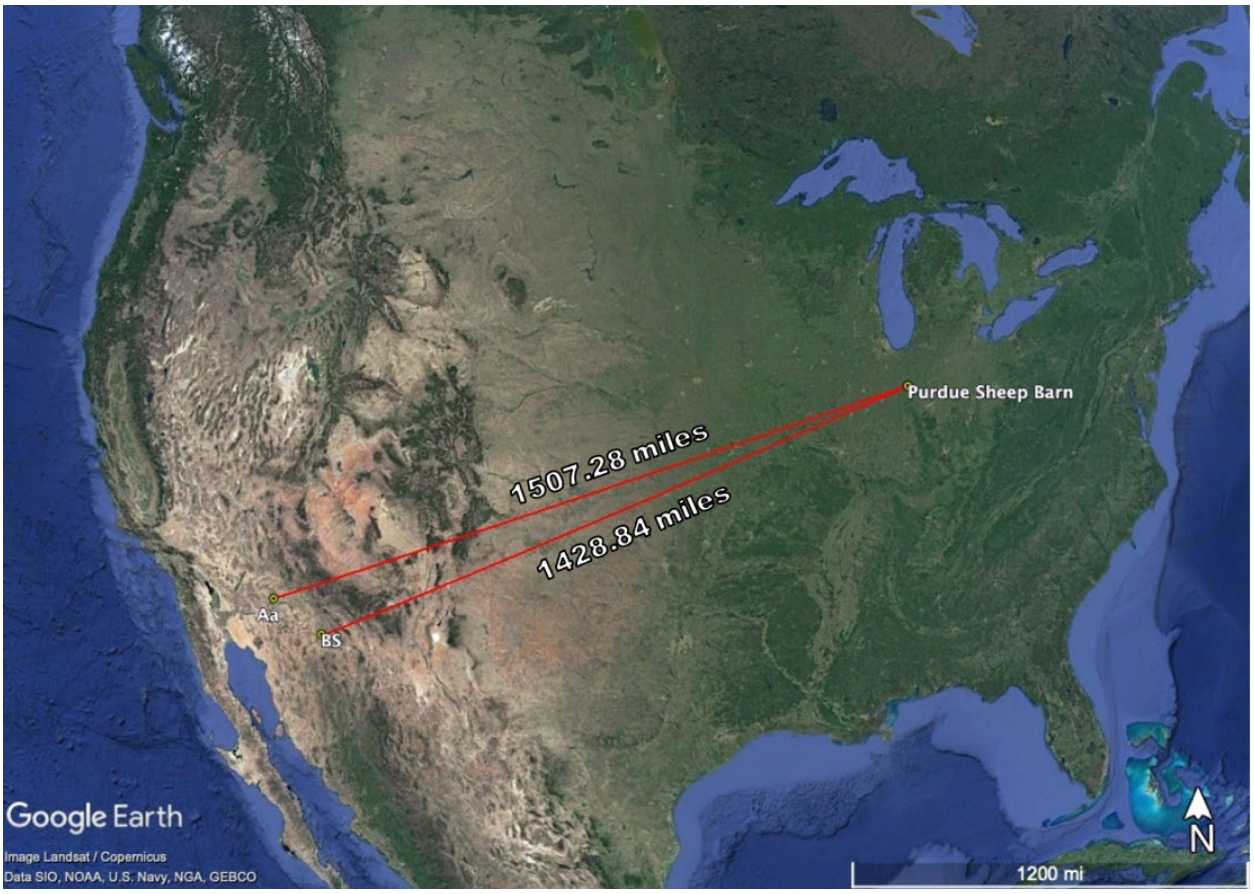

Figure 3.3. Map demonstrating the distance in miles between the Purdue sampling site and the poultry farm (Aa) and Biosphere 2 (BS) sampling sites. Several site markers were removed to reduce clutter due to the scale of the map.

Species Identification

Ten fly specimens were retrieved from the sticky trap and identified as *M. domestica* using morphological keys [90-94]. Species identifications of five samples were supported by the amplification and sequencing of the *cytochrome oxidase I* (*COI*) gene.

DNA Extraction

DNA was extracted from fly heads using organic phenol-chloroform extraction. After decapitation, the heads digested in 200µL ChargeSwitch® lysis buffer (Invitrogen™) and 20µL 20mg/mL proteinase K (Invitrogen™) while incubating overnight at 56°C. To help expose and digest cells during this process, fly head tissue was smashed with sterile pipette tips and vortexed prior to incubation. 300µL phenol chloroform isoamyl alcohol (PCI) (25:24:1) (Thermo Fisher

Scientific™) was added to each lysate. Sample tubes were inverted several times and centrifuged at 13,000rpm for 1 minute to allow separation into organic and aqueous layers. Extraction was continued following a standard organic phenol-chloroform protocol [89].

Polymerase Chain Reaction (PCR) Amplification

PCR of the *COI* gene was performed using forward primer TL2-N-3014 (Pat) and reverse primer C1-J-2183 (Jerry) [13]. The 10µL volume PCR reaction consisted of the following: 5µL 2X PCR Master Mix (Promega™), 0.5µL 1X bovine serum albumin (BSA) (Promega™), 1µL 5µM forward and reverse primers, and 2.5µL genomic DNA. PCR conditions were as follows: initial denaturation at 94°C for 60s, 5 cycles of 94°C for 30s, 47°C for 40s, and 72°C for 60s, 30 cycles of 94°C for 30s, 52°C for 40s, and 72°C for 60s, followed by a final extension step at 72°C for 10 minutes [95]. Amplifications were performed using a Mastercycler Pro thermocycler (Eppendorf®). Amplification was verified using a 1% agarose gel for expected sized products.

Sequencing

Amplicons were purified using 1µL ExoSAP-IT™ (Thermo Fisher Scientific™) and 2.5µL PCR product following the manufacturer's protocol. Subsequently, cycle sequencing was performed using BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific™) following the manufacturer's protocol. Sequencing products were then purified by ethanol precipitation: 1.25µL 125mM EDTA was dissolved into the samples. 20µL ice cold 95% ethanol was added, and the plate incubated at room temperature for 15 minutes. Samples were then centrifuged in a refrigerated centrifuge (Sorvall Legend XTR Centrifuge; Thermo Scientific™) at 2500g for 30 minutes. 95% ethanol was promptly removed and followed by the addition of 20µL ice cold 70% ethanol. Samples were placed in the same centrifuge used previously and centrifuged at 2500g for 15 minutes. After removal of the 70% ethanol, the plate was inverted and centrifuged at 185g for 1 minute. Samples were resuspended in 10µL HiDi™ formamide (Thermo Fisher Scientific™), vortexed for 15 seconds, and denatured for 5 minutes at 95°C using a Mastercycler Pro thermocycler (Eppendorf®). Sequence separation and detection was performed using a 3500 genetic analyzer (Thermo Fisher Scientific™).

Sequence Processing and Species Confirmation

Sequences were edited and trimmed manually using Sequence Scanner Software v1.0 (Applied Biosystems®). Trimmed sequences were imported into BLASTn® (National Center for Biotechnology Information; NCBI) for query against a nucleotide database. Five of the fly samples collected from Purdue University were confirmed to be *M. domestica* with query coverages ranging from 78-90% and e-values of 3e-87 to 5e-73.

3.2.3 Amplified Fragment Length Polymorphism (AFLP)

Preparation

The AFLP protocol described here is modified from a previously published protocol [22]. All primers and adapters were purchased from Integrated DNA Technologies (IDT) except for FAM-labeled *Eco*RI+A which was purchased from Applied Biosystems® Custom Oligo Synthesis Service (Table 3.3). Concentrated primers were diluted to 25pg/µL using nuclease-free water. *Pst*I and *Eco*RI adapters were prepared by combining 3µg of each adapter, 6 µL 10X Buffer H (Promega™), and nuclease-free water to a total volume of 120µL. Adapter preparation was concluded by heating the solutions in a Mastercycler Pro thermocycler (Eppendorf®) at 65°C for 10 minutes, 37°C for 10 minutes, and 25°C for 10 minutes.

Table 3.3. Sequences of adapters and primers for AFLP amplifications.

Name	5'-3' Sequence
*Eco*RI adapter 1	CTCGTAGACTGCGTACC
*Eco*RI adapter 2	AATTGGTACGCAGTCTAC
*Pst*I adapter 1	CTCGTAGACTGCGTACATGCA
*Pst*I adapter 2	TGTACGCAGTCTACG
*Eco*RI+A	GACTGCGTACCAATTCA
*Eco*RI+AFAM	6FAMGACTGCGTACCAATTCA
*Pst*I+A	GACTGCGTACATGCAGACA
*Pst*I+AAC	GACTGCGTACATGCAGACAAC
*Pst*I+ACG	GACTGCGTACATGCAGACACG
*Pst*I+AGT	GACTGCGTACATGCAGACAGT
*Pst*I+ATC	GACTGCGTACATGCAGACATC

Protocol

Genomic DNA was digested using 2 U *Pst*I, 2 U *Eco*RI, 5µL 10X Buffer H (Promega™), 0.5µL bovine serum albumin (BSA) (Promega™), 35µL DNA, and nuclease-free water to a total volume of 50µL. Digests were conducted in a Mastercycler Pro thermocycler (Eppendorf®) at 37°C for 3 hours followed by 15 minute denaturation step at 70°C. Adapters were ligated to the digested DNA using 1µL of each adapter, 1µL T4 DNA Ligase Buffer (New England Biolabs®) 1U T4 DNA Ligase (New England Biolabs®), and nuclease-free water to a total volume of 10µL. This solution was added directly to the 50µL of digested DNA to a new total volume of 60µL and held at room temperature for 3 hours. Samples were agitated gently each hour. Pre-selective amplification was conducted by combining 10µL 2X PCR Master Mix (Promega™), 1µL 25pg/µL *Eco*RI+A primer, 1µL 25pg/µL *Pst*I+A primer, 6µL nuclease-free water, and 2µL ligated DNA for a total volume of 20µL. Samples were amplified using a thermocycler and a program consisting of the following conditions: initial denaturation at 94°C for 2 minutes, 26 cycles of 94°C for 1 minute, 56°C for 1 minute, and 72°C for 90 seconds, and a final extension at 72°C for 5 minutes. Amplified samples were then diluted with 100µL nuclease-free water. Selective amplification was conducted by combining 1µL diluted amplicon, 10µL 2X PCR Master Mix (Promega™), 1.5µL 25pg/µL *Eco*RI+AFAM primer, 6µL nuclease-free water, and

1.5µL of one of the four 25pg/µL *Pst*I selective primers for a total volume of 20µL. Samples were amplified using a thermocycler and a program consisting of the following conditions: initial denaturation at 94°C for 2 minutes, 10 cycles of touchdown PCR at 94°C for 30 seconds, 65-56°C for 30 seconds, and 72°C for 90 seconds, 27 cycles of 94°C for 30 seconds, 56°C for 30 seconds, and 72°C for 90 seconds, and a final extension at 72°C for 10 minutes. 1µL of this PCR product was added to 9.7µL HiDi™ formamide (Thermo Fisher Scientific™), and 0.3µL GeneScan™ 600 LIZ™ dye Size Standard v2.0 (Applied Biosystems®). Samples were run on a 3500 genetic analyzer (Thermo Fisher Scientific™) using the following conditions: 50-cm 8-capillary array, POP-7™ Polymer (Applied Biosystems®), dye set G5, oven temperature of 60°C, injection voltage of 1.6 kV, injection time of 8 seconds, run voltage of 19.5 kV, and run time of 1330 seconds.

3.2.4 Analysis

Data Processing

Raw data from the 3500 genetic analyzer (Thermo Fisher Scientific™) were imported into GeneMarker® AFLP/Genotyping Software v2.4.0 (SoftGenetics LLC®). Default settings for AFLP analysis were used with the following modifications: local southern sizing method, 100-600bp analysis range, and 200rfu minimum peak intensity. Genotype data for each primer set were exported into Microsoft® Excel. Each dataset was sorted and cropped to include only alleles that were observed in either >5% or <95% of individuals [22, 96]. This resulted in datasets containing 267 polymorphic alleles from the four *Pst*I primers.

Statistical Analyses

Statistical and population genetic analyses were conducted utilizing a variety of downloadable software. AFLPOP v1.2, a population assignment program for AFLP data, was used to perform population reallocation simulations to estimate to which population the genotype of each sample belongs [97]. Spatial Analysis of Genetic Diversity (SPAGeDi 1.5) was used to calculate coefficients of kinship (*r*), or the probability of alleles being identical-by-descent (IBD),

by using pairwise comparisons and an assumed inbreeding coefficient of 0 [98, 99]. Structure v2.3.4, a software for population genetics inference, was used to investigate genetic clustering and structure [100]. Structure parameters included a burn-in of 1,000 and MCMC of 50,000 with k = 1-16 and five iterations. Output data from Structure was imported into Structure Harvester in order to determine the most likely number of populations present in the dataset [101]. Genetic Analysis in Excel (GenAlEx 6.503) was used to perform Analysis of Molecular Variance (AMOVA) as well as a Mantel Test [102, 103]. AMOVA uses genetic data to detect population differentiation and generate F-statistic analogs that describe heterozygosity in the dataset [104]. The Mantel Test investigates possible relationships between geographic and genetic distances [105].

3.3 Results

3.3.1 Population Assignment

When treating each sampling site as a putative separate population, individual specimens routinely did not reallocate back to their source population. Using an allocation threshold minimum log likelihood difference (MLD) of 3, 77.27% (N=68) of specimens did not allocate to any specific population (Table 3.4). An MLD of 3 requires a specimen to be 1000X more likely to allocate to a specific population over another to be allocated to that specific population [97]. This result indicates that the *M. domestica* specimens were largely not genetically different enough to justify allocation to one population over another. Only 7.95% of specimens (N=7) reallocated to their original population. The only populations that had specimens reallocate to their original population were BS, DF, and Purdue. Additionally, 14.77% of specimens (N=13) reallocated to the incorrect source population. Three specimens collected from the poultry farm (populations Aa-Ac) allocated to populations collected from sites up to 5.0 miles from the farm (Figure 3.4). No specimens collected from non-poultry farm sites allocated to the poultry farm. After lowering the MLD thresholds to 2 and 1 (specimens would be required to be 100X and 10X more likely to allocate to a specific population over another, respectively), which makes the reallocation criteria less stringent, more specimens allocated to some population, but still did not

allocate to their true origin population (Tables 3.5-3.6; Figures 3.5-3.6). This further highlights that the different source populations were genetically similar and difficult to differentiate.

Table 3.4. Population reallocation results under minimum log likelihood (MLD) 3. Total N=88.

	Number of Samples	**Percentage of Samples**
Reallocated to No Population	68	77.27
Reallocated to Correct Population	7	7.95
Reallocated to Incorrect Population	13	14.77

Table 3.5. Population reallocation results under minimum log likelihood (MLD) 2. Total N=88.

	Number of Samples	**Percentage of Samples**
Reallocated to No Population	52	59.09
Reallocated to Correct Population	10	11.36
Reallocated to Incorrect Population	26	29.55

Table 3.6. Population reallocation results under minimum log likelihood (MLD) 1. Total N=88.

	Number of Samples	**Percentage of Samples**
Reallocated to No Population	34	38.64
Reallocated to Correct Population	11	12.50
Reallocated to Incorrect Population	43	48.86

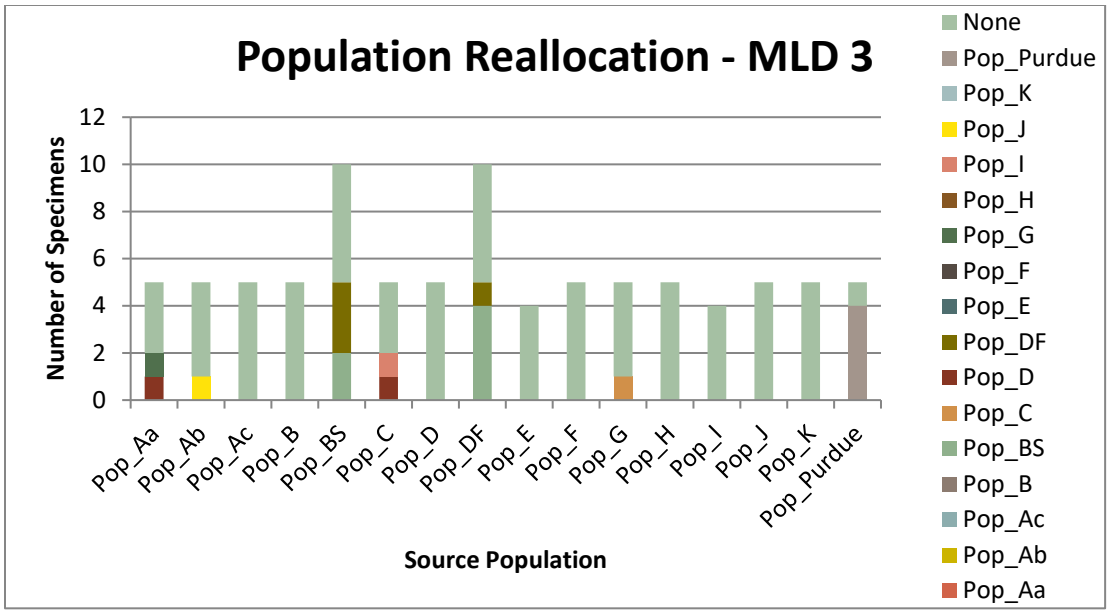

Figure 3.4. Population reallocation results under minimum log likelihood (MLD) 3. Under MLD 3, a specimen is required to be 1000X more likely to allocate to a specific population over another to be allocated to that specific population.

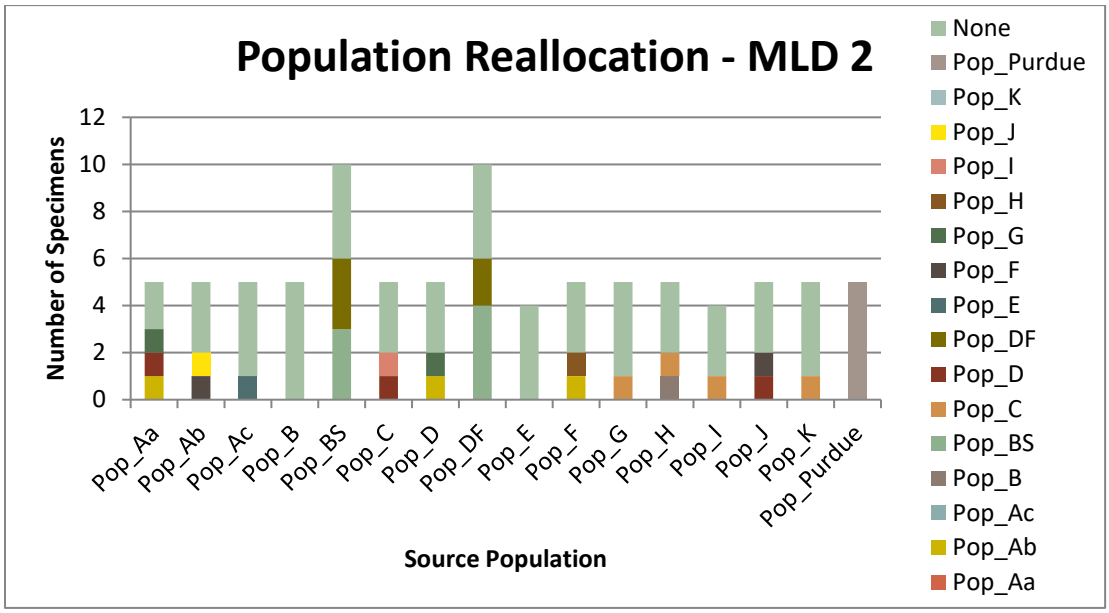

Figure 3.5. Population reallocation results under minimum log likelihood (MLD) 2. Under MLD 2, a specimen is required to be 100X more likely to allocate to a specific population over another to be allocated to that specific population.

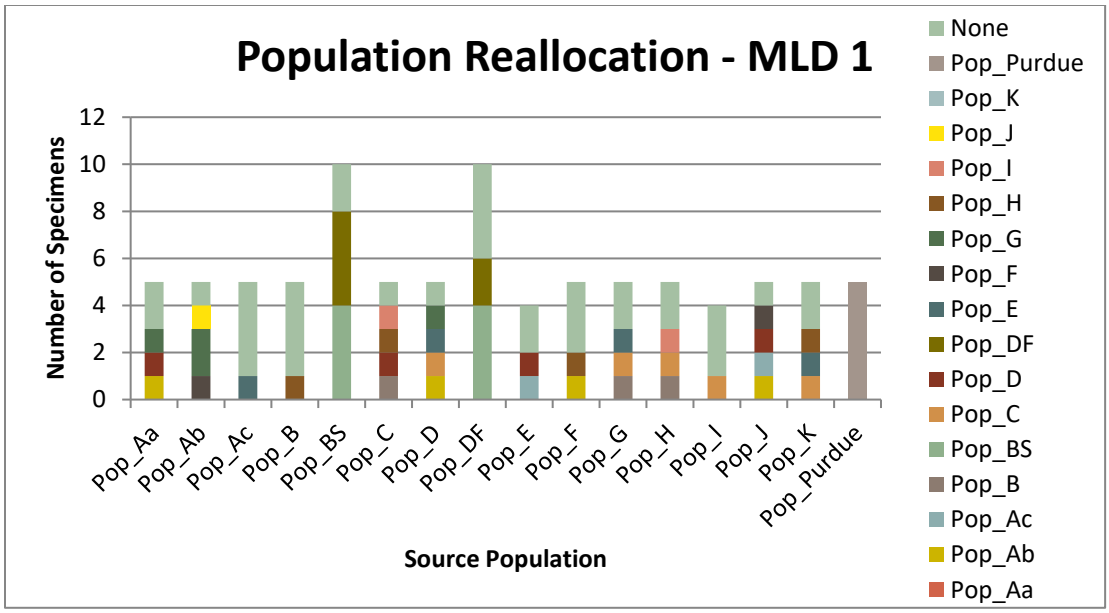

Figure 3.6. Population reallocation results under minimum log likelihood (MLD) 1. Under MLD 1, a specimen is required to be 10X more likely to allocate to a specific population over another to be allocated to that specific population.

3.3.2 Kinship

Kinship coefficients (*r*) were generated to describe the relatedness of samples using an assumed inbreeding coefficient of 0. Kinship coefficients among individuals from singular collection sites varied widely, with averages ranging from *r* = -0.002463 (site Aa) to *r* = 0.20213442 (site BS) (Table 3.7). The highest levels of kinship within the entire dataset were seen between a BS specimen and DF specimen (*r* = 0.388362), and a C specimen and D specimen (*r* = 0.368251). The lowest levels of kinship were seen between a D and DF specimen (*r* = -0.183518) and an Aa and BS specimen (*r* = -0.18213). With BS and DF excluded, the highest values were observed between a C specimen and D specimen (*r* = 0.437324) and a C and H specimen (*r* = 0.341764). The lowest levels of kinship with BS and DF excluded were between a D and K specimen (*r* = -0.170749) and an Aa and P specimen (*r* = 0.160084). When focusing on kinship coefficient values between poultry farm (sites Aa, Ab, and Ac) specimens and other collection sites excluding P, BS, and DF, the largest values were between an Aa specimen and D specimen (*r* = 0.213165 and *r* = 0.170299), and an Ab and F specimen (*r* = 0.153571).

A theoretical *r* value of 0.50 would reflect full sibship while a value of 0.25 would be expected for half-siblings [106] – as relatedness becomes more distant, *r* values become smaller. An *r* value of 1.0 would indicate theoretical identical twins. Average kinship values of specimens from sites BS and DF indicate a relationship of roughly half-siblings, while Purdue indicates kinship roughly analogous to those of first cousins (theoretical value 0.125). Within all other sites from and surrounding the poultry farm, specimens appear not to be closely related overall (Figure 3.7).

Table 3.7. Average, maximum, and minimum kinship coefficient values (*r*) among samples within each collection site. Poultry farm collection sites Aa, Ab, and Ac are both reported individually and combined as site A.

Site	Average	Maximum	Minimum
A	0.02130691	0.165988	0.-150765
Aa	-0.0002463	0.122749	-0.099194
Ab	0.0531922	0.130362	0.003661
Ac	0.040953	0.090714	-0.007113
B	0.0450903	0.088416	-0.014727
BS	0.20213442	0.305044	0.103499
C	0.0172792	0.106803	-0.12146
D	0.0100391	0.110251	-0.07937
DF	0.18604849	0.3049	0.105654
E	0.01639833	0.084825	-0.045612
F	0.0125672	0.040292	-0.018749
G	0.0272776	0.055807	0.002081
H	0.0103839	0.061266	-0.094741
I	0.053197	0.077498	0.031386
J	0.0009026	0.73189	-0.089139
K	0.0345174	0.099477	0.000501
Purdue	0.1439805	0.237096	0.048911

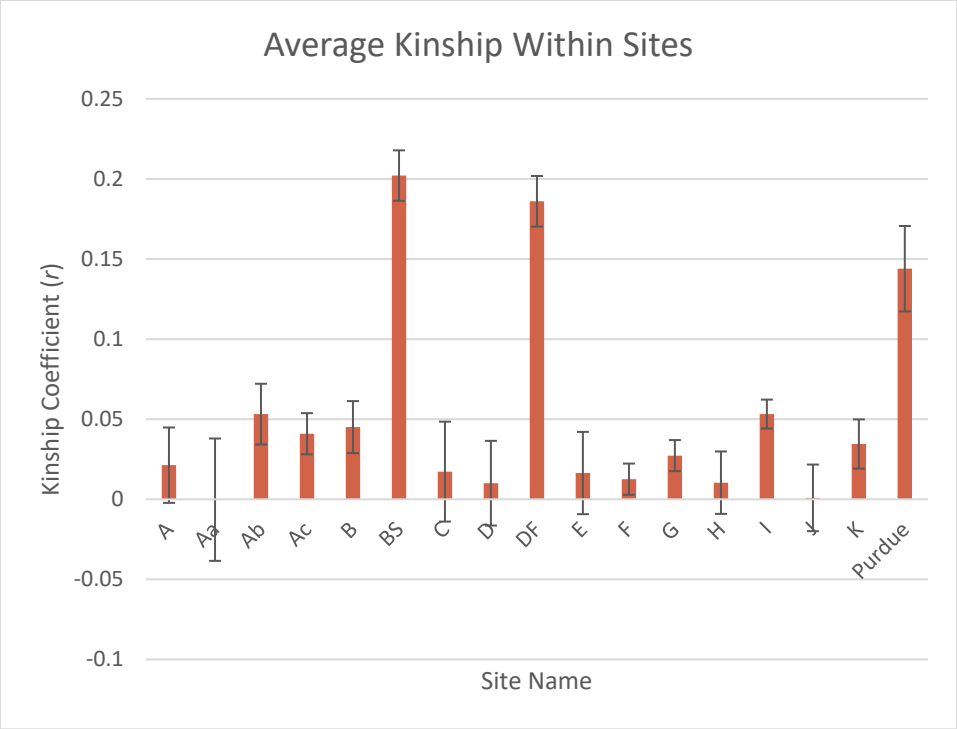

Figure 3.7. Average kinship coefficient values (*r*) among samples within each collection site. Poultry farm collection sites Aa, Ab, and Ac are both reported individually and combined as site A. Error bars represent standard error.

3.3.3 Population Number and Structure

The Structure program was run with poultry farm sites Aa, Ab, and Ac remaining separate, and additionally run with the sites combined into a singular site A. Upon noting the very clear difference in shared alleles between specimens collected in the first round of collections (from sites Aa-B, C, D, E-K) and specimens collected approximately seven months later (from sites BS and DF) (Figures 3.8-3.9), Structure was run two additional times excluding those specimens collected later (Figures 3.10-3.11). Regardless of the combination of sites Aa-Ac and the inclusion or exclusion of sites BS and DF, the datasets were always concluded to contain two genetically different populations (K=2) using Structure Harvester. If each of the collection sites were different genetic populations, K values would have been between 12-16 (dependent on site A combination and BS/DF inclusion or exclusion).

When observing the proportion of shared alleles demonstrated by the Structure bar plots, it is evident that specimens collected from sites BS and DF are genetically similar to each other, but largely differ genetically from specimens collected in and around the poultry farm during the

first collection. Sites BS and DF are additionally more genetically similar to the outgroup Purdue site than any of the other collection sites. Though K=2, flies collected from sites within the poultry farm (Aa-Ac) do not represent a distinct genetic population different from all other collection sites surrounding the farm. When the poultry farm (site A) remains separated into three individual sites of Aa, Ab, and Ac, site Aa appears to contain individuals from both genetic populations while sites Ab and Ac largely contain individuals from a single genetic population. This second genetic population, represented in red in the bar plots, also represents the majority of flies collected from the sites analogous to the plaintiff's homes in the civil lawsuit. Additionally, it is clear that the two genetic populations are experiencing admixture, meaning they are likely interspersing and breeding together. This is represented by a vertical bar in the plot, of which each represents an individual (Figures 3.12-3.13), containing a mixture of the two colors (which represent the two genetic populations).

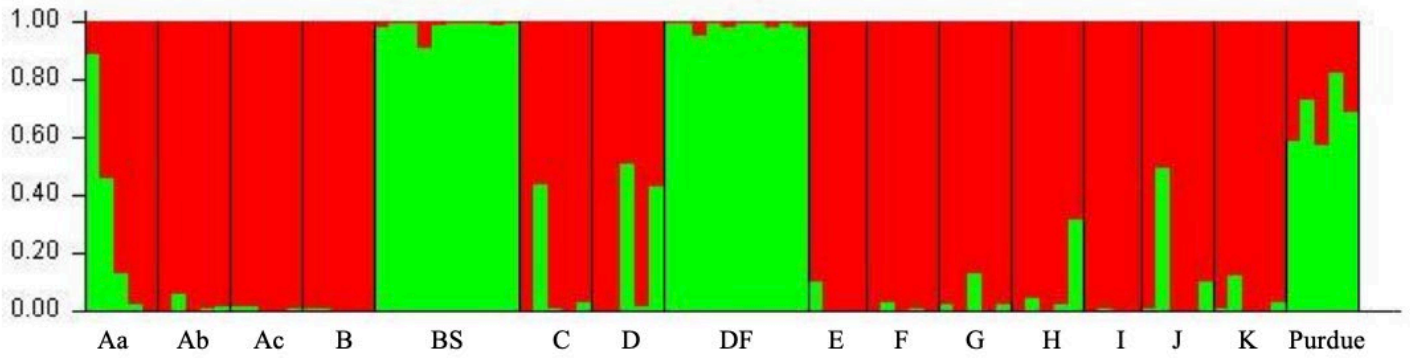

Figure 3.8. Structure bar plot summarizing population substructure by collection site. Site names are given on the x-axis while proportion of shared alleles is given on the y-axis. K=2.

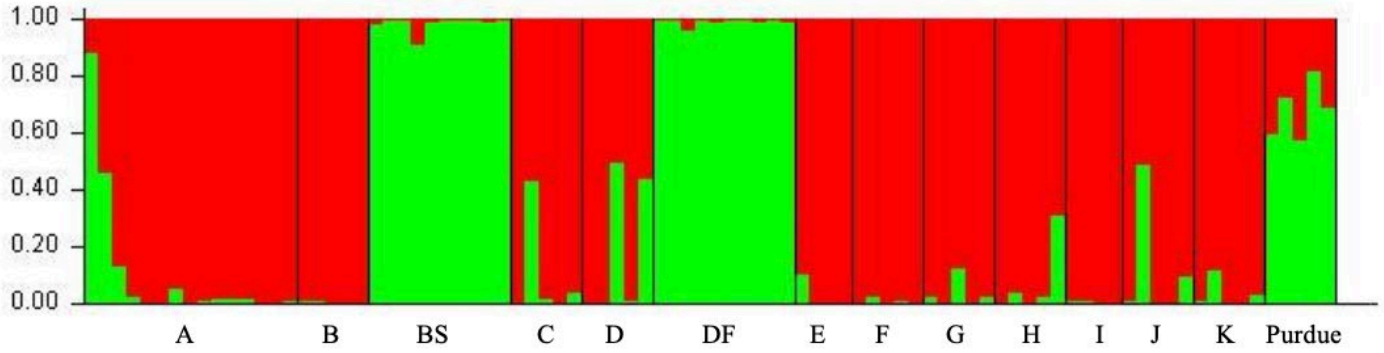

Figure 3.9. Structure bar plot summarizing population substructure by collection site. Sites Aa, Ab, and Ac have been combined into site A. Site names are given on the x-axis while proportion of shared alleles is given on the y-axis. K=2.

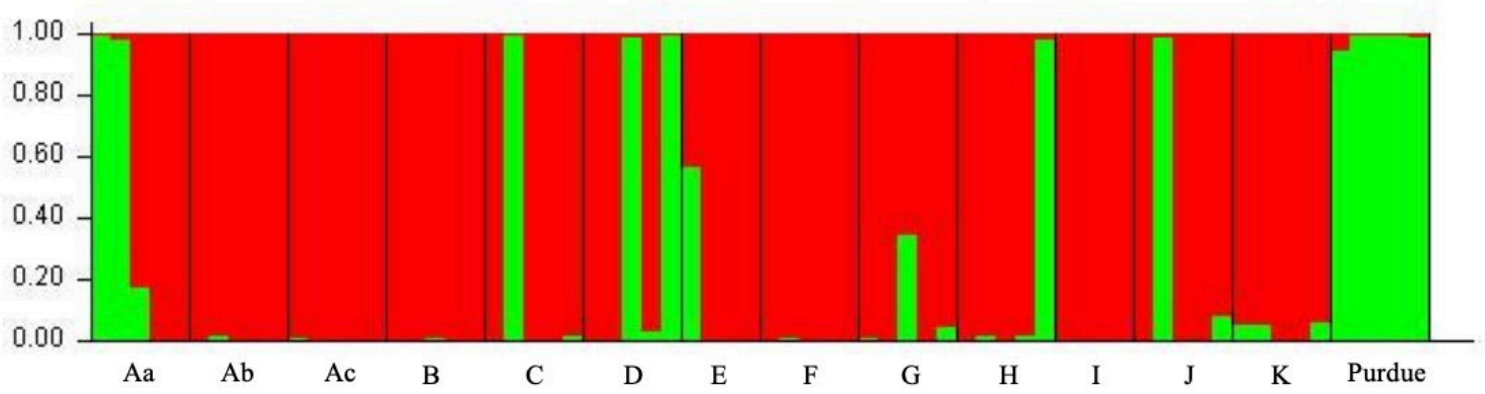

Figure 3.10. Structure bar plot summarizing population substructure by collection site. Sites BS and DF have been excluded. Site names are given on the x-axis while proportion of shared alleles is given on the y-axis. K=2.

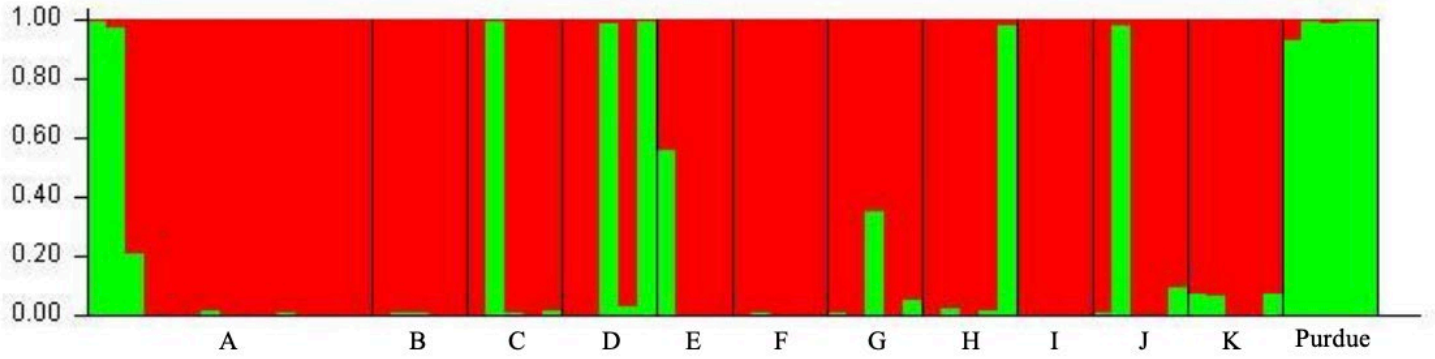

Figure 3.11. Structure bar plot summarizing population substructure by collection site. Sites Aa, Ab, and Ac have been combined into site A. Sites BS and DF have been excluded. Site names are given on the x-axis while proportion of shared alleles is given on the y-axis. K=2.

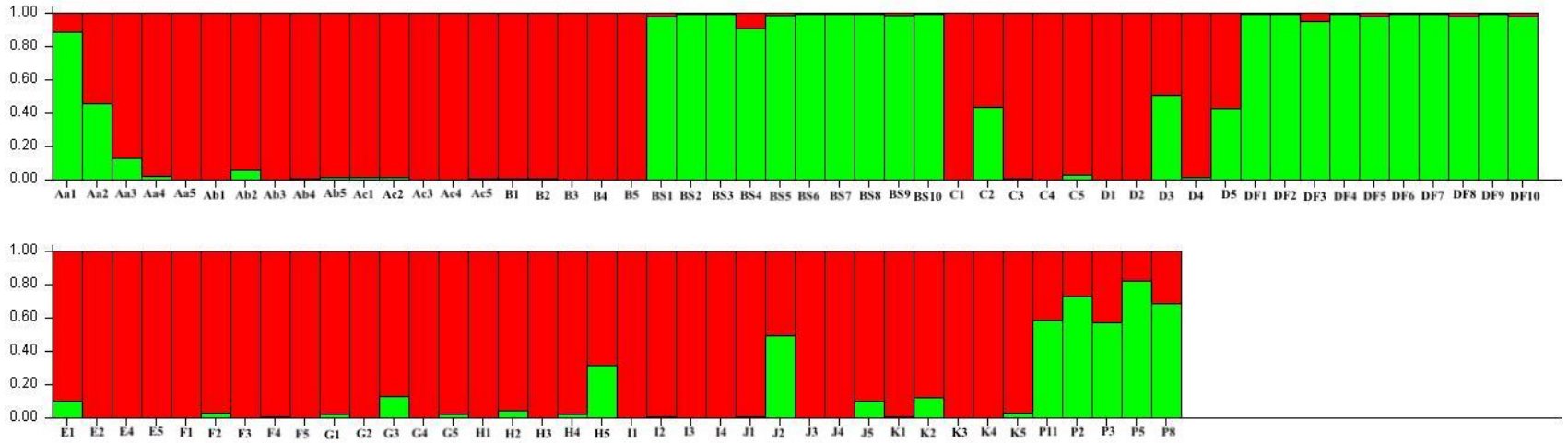

Figure 3.12. Structure bar plot summarizing population substructure by collection site. Individual sample IDs categorized by collection site are given on the x-axis while proportion of shared alleles is given on the y-axis. Each vertical bar represents an individual fly specimen. K=2.

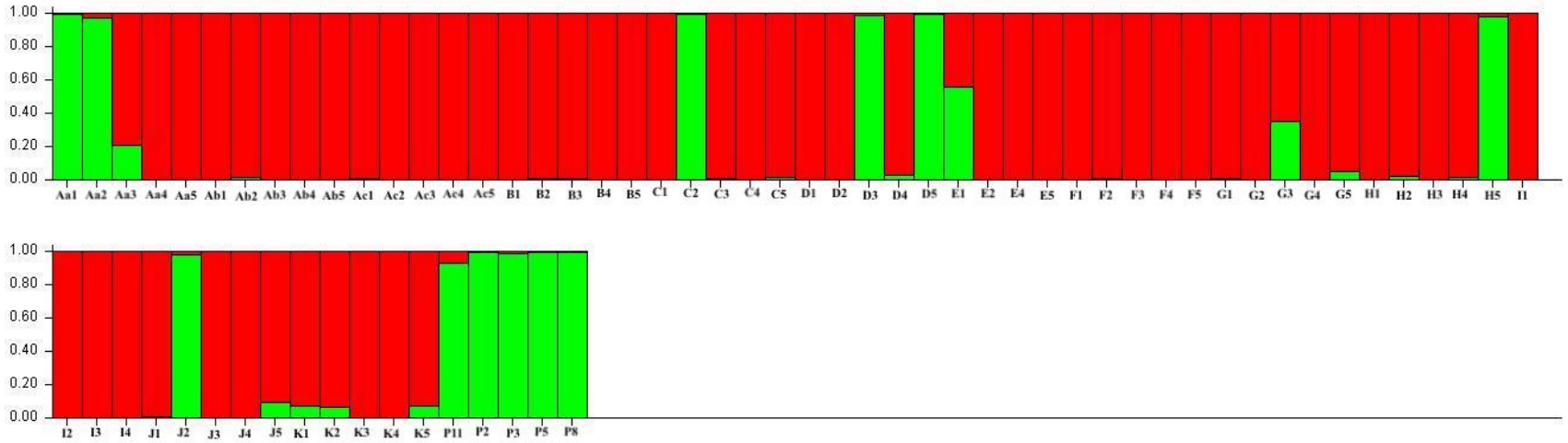

Figure 3.13. Structure bar plot summarizing population substructure by collection site. Individual sample IDs categorized by collection site are given on the x-axes while proportion of shared alleles is given on the y-axes. Samples from sites BS and DF have been excluded. Each vertical bar represents an individual fly specimen. K=2.

3.3.4 Genetic Variance and Geographic vs. Genetic Distances

Analysis of molecular variance (AMOVA) revealed that when treating each collection site as an individual population, the majority of genetic variance lay within, rather than among, populations. This indicates that rather than flies from each collection site differing from each other genetically, the greatest amount of genetic differences were present within the set of flies collected at each site. When specimens from sites BS and DF were included, genetic variance within populations was 87% while variance among populations was 13% (Tables 3.8-3.9). After excluding sites BS and DF, genetic variance within populations was 97% while variance among populations was 3% (Tables 3.10-3.11). After excluding sites BS, DF, and Purdue, AMOVA regarding only the initial collection sites revealed that 100% of genetic variance lay within populations and 0% lay among populations (Table 3.12). Results are based on 999 permutations. Mantel Tests found no relationship between genetic and geographic distances (Table 3.13; Figures 3.14-3.16).

Table 3.8. Summary of AMOVA results describing genetic variance among and within populations (collection sites). In this dataset, site A was separated into sites Aa, Ab, and Ac. Sites BS and DF were included.

Source	df	SS	MS	Estimated Variance	Percentage of Total Variance	*P*
Among Populations	15	1069.123	71.275	5.888	13%	0.001
Within Populations	72	2815.400	39.103	39.103	87%	
Total	87	3884.523		44.991	100%	

Table 3.9. Summary of AMOVA results describing genetic variance among and within populations (collection sites). In this dataset, sites Aa, Ab, and Ac were combined into site A. Sites BS and DF were included.

Source	df	SS	MS	Estimated Variance	Percentage of Total Variance	*P*
Among Populations	13	976.856	75.143	5.807	13%	0.001
Within Populations	74	2907.667	39.293	39.293	87%	
Total	87	3884.523		45.100	100%	

Table 3.10. Summary of AMOVA results describing genetic variance among and within populations (collection sites). In this dataset, site A was separated into sites Aa, Ab, and Ac. Sites BS and DF were excluded.

Source	df	SS	MS	Estimated Variance	Percentage of Total Variance	*P*
Among Populations	13	592.665	45.590	1.347	3%	0.002
Within Populations	54	2108.600	39.048	39.048	97%	
Total	67	2701.265		40.395	100%	

Table 3.11. Summary of AMOVA results describing genetic variance among and within populations (collection sites). In this dataset, sites Aa, Ab, and Ac were combined into site A. Sites BS and DF were excluded.

Source	df	SS	MS	Estimated Variance	Percentage of Total Variance	*P*
Among Populations	11	500.398	45.491	1.118	3%	0.001
Within Populations	56	2200.867	39.301	39.301	97%	
Total	67	2701.265		40.419	100%	

Table 3.12. Summary of AMOVA results describing genetic variance among and within populations (collection sites). In this dataset, site A was separated into sites Aa, Ab, and Ac. Sites BS and DF, and Purdue were excluded, meaning this data represents initial collection sites only.

Source	df	SS	MS	Estimated Variance	Percentage of Total Variance	*P*
Among Populations	12	467.235	38.936	0.000	0%	0.512
Within Populations	50	1949.400	38.988	38.988	100%	
Total	62	2416.635		38.988	100%	

Table 3.13. Summary of Mantel Test results for each dataset. “Site A Separated” refers to collection sites Aa, Ab, and Ac remaining separate while “Site A Combined” refers to combing those three sites into a single site A. “Includes BS, DF” refers to the inclusion of specimens collected from sites BS and DF. “Excludes BS, DF” refers to these specimens being removed. The last row represents a dataset solely consisting of specimens from the initial collection sites (excludes BS, DF, and Purdue).

	SSx	SSy	SPxy	Rxy	*P*
Site A Separated; Includes BS, DF	1304172.918	2075366512.811	9536952.000	0.183	0.030
Site A Combined; Includes BS, DF	1304172.918	2075366512.811	9536952.000	0.183	0.014
Site A Separated; Excludes BS, DF	549902.126	1582814359.288	12655491.624	0.429	0.001
Site A Combined; Excludes BS, DF	1304172.918	2075366512.811	9536952.000	0.183	0.014
Site A Separated; Excludes BS, DF, and Purdue	416350.213	24716.760	-4176.680	-0.041	0.400

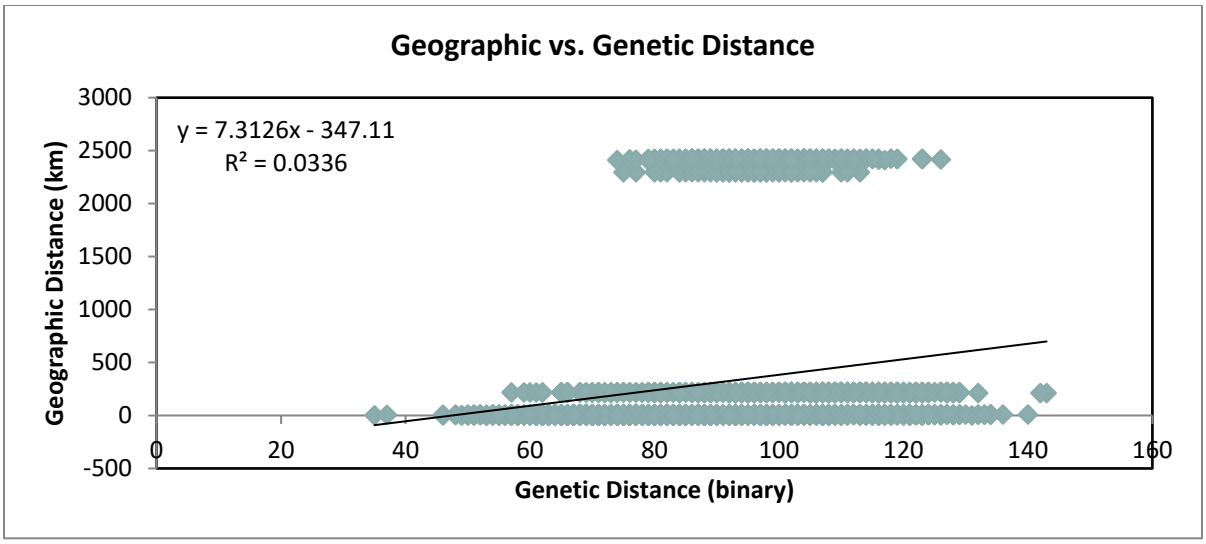

Figure 3.14. Relationship between genetic and geographic distance based on Mantel Test results. This graph reflects a dataset including specimens from sites BS and DF.

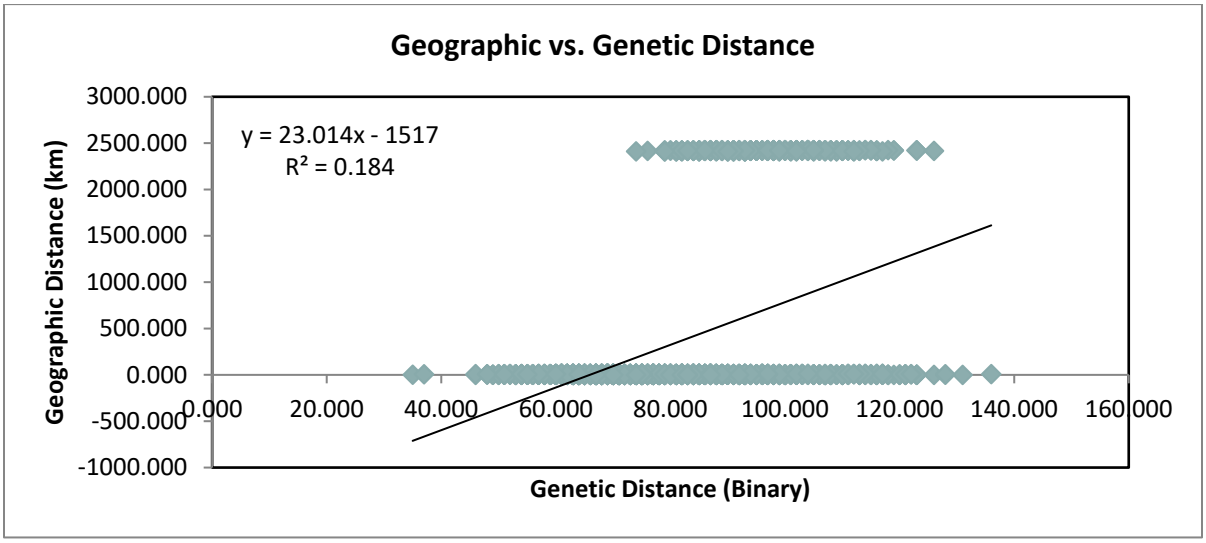

Figure 3.15. Relationship between genetic and geographic distance based on Mantel Test results. This graph reflects a dataset excluding specimens from sites BS and DF.

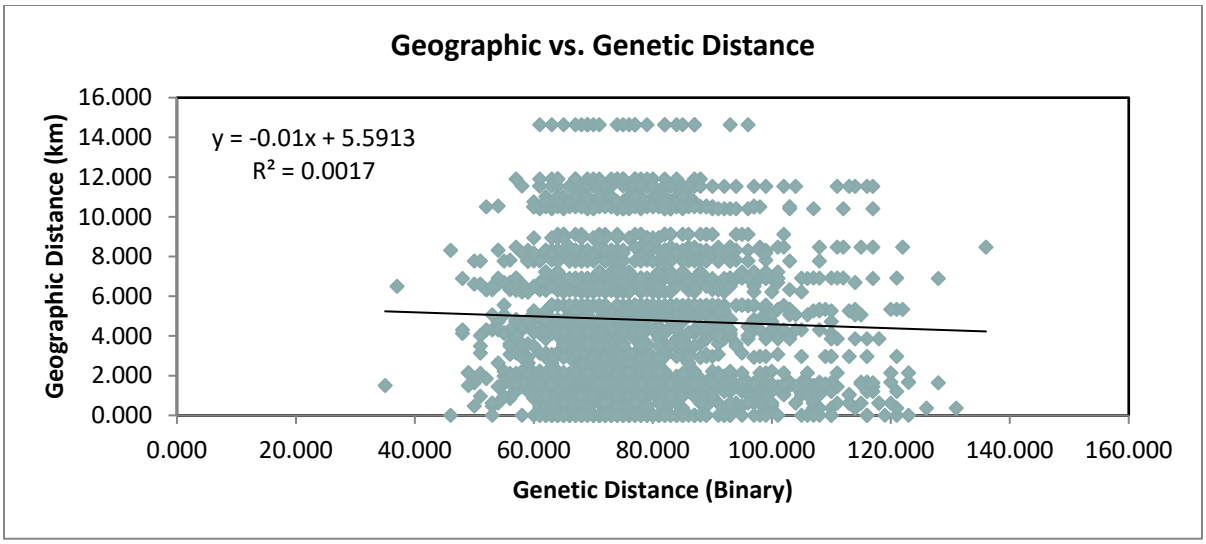

Figure 3.16. Relationship between genetic and geographic distance based on Mantel Test results. This graph reflects a dataset excluding specimens from sites BS and DF, and Purdue, leaving only data from the initial collection sites.

3.4 Discussion

The overall conclusion based on the variety of analyses presented here is that house flies collected in and around the poultry farm are likely all the same breeding population. Flies from the farm and flies from up to 7.0 miles away from the farm could not be differentiated using a population genetic approach.

The fact that the Biosphere 2 (BS) and Dairy Farm (DF) specimens differed so greatly from the other sites and were very similar to each other indicates that the time difference (approximately 7 months) had a large effect on the genetic structure of the house fly populations in this area. This is especially true when considering that collection site H was only 1.68 miles away from site DF. On the initial collection dates of December 14 and 15, 2018, it was the end of autumn in Arizona, with winter beginning shortly after on the 21st. On the second collection dates of July 14 and 23, 2019, it was the middle of summer. The closest weather station with data on these dates indicates that on December 14 and 15, 2018, the high temperatures were 66°F and 70°F, low temperatures were 43°F and 50°F, 0.00 inches precipitation on both dates, and max wind speeds of 7mph and 8mph, respectively [107]. On July 14 and 23, 2019, weather station data indicates high temperatures of 109°F and 106°F, low temperatures of 93°F and 83°F, 0.00 inches precipitation both days, and max wind speeds of 16mph and 13mph, respectively.

These rather extreme differences likely heavily influenced the house fly populations being sampled at each time period. It is known that house fly egg development and oviposition cycles are temperature dependent [108]. A 1986 study that investigated seasonal breeding structure concluded that in autumn, genetic drift causes subpopulations to differentiate due to overwintering, as populations breed slowly and become reduced in size [76]. Populations increased dramatically in June as summer began, and differentiation decreased rapidly. This cycle of overwintering to population expansion acts as a population bottleneck or "boom and bust," causing fluctuations in gene frequencies between seasons [77]. In the same 1986 study, it was observed that younger flies exhibited more distinct differentiation while older flies were panmictic [76]. This indicates that in the current study, not only could the season and temperature have influenced the genetic differences observed, but also the ages of the flies collected could have contributed to differences. Had additional specimens been collected from the thirteen original sites on the second date, their genetics may have differed significantly from the original specimens collected seven months prior.

House fly populations in commercial farming facilities are not only subject to population bottlenecks due to overwintering, but also due to pest control measures used to drastically reduce population size [71]. A number of control measures exist for *M. domestica,* ranging from chemical insecticides to bacterial agents and essential oils [109] . Unfortunately, no information concerning fly-control measures at the poultry farm or nearby dairy farm were provided.

While this study indicates an ability to investigate genetic populations of house flies using molecular methods, it is not without its limitations. First, the present study utilized small sample sizes of 5 or 10 individual specimens per collection site. Our intent was to establish a proof of concept that amplified fragment length polymorphism (AFLP) analysis could be an effective study tool for *M. domestica*, since studies on this species seem to have focused on other methodologies. As such, small sample sizes were initially utilized to avoid extended fly collections only to find that the data generated were not probative. Second, though AFLP data can be ideal for population genetic studies as previously discussed, AFLP as a method contains inherent limitations. One such limitation is that AFLP generates dominant markers, meaning that markers are either deemed present or absent, and homologous alleles are not discriminated [110]. As such, relatedness predictions utilizing AFLP data are estimates rather than precise results. If such a limitation would be considered an issue, perhaps AFLP could be used as an initial

screening method, and for more in-depth analyses, one could progress to more precise methodologies as needed. However, despite its limitations, AFLP continues to be a popular molecular technique, and suited the needs for our study.

Overall, the results of this study indicate that the house fly population in this region outside of Phoenix, Arizona predominantly consists of intermixing genetic populations breeding together with no clear geographic distinctions. Specimens were largely not distinctly different enough to reallocate to a specific source population, and if they did allocate, it was often not to their true source population. Kinship coefficient (*r*) values were low within initial collection sites, indicating that flies in one location were no more closely related to each other than flies in another location. Analysis of genetic structure indicated that the number of genetically distinct populations in the dataset was two (K = 2), with flies collected from the poultry farm appearing genetically similar to flies collected from other sites during the initial collections. The second population of the two mainly reflects the outgroup collected from Purdue, as well as differences observed in specimens from Biosphere 2 (BS) and Dairy Farm (DF). The majority of genetic variance was seen within, but not between, sampling sites, which indicates that the majority of the time, none of the collection sites provided specimens that were genetically distinct from other collection sites. Rather, the genetic distinctions were within the collection sites themselves. Additionally, no significant relationships were found that indicated geographic distances impacted genetic distances or vice versa.

3.5 Conclusions

Musca domestica population genetic analyses were conducted in regard to a civil lawsuit in Arizona, USA. Based on the results, it is not possible to conclude that house flies found in and around the plaintiffs' homes are genetically different from house flies collected from the poultry farm. However, at the same time, it is not possible to conclude that flies collected at the plaintiff's homes emanated directly from the poultry farm. Future studies may benefit from collecting specimens from the same sites at different time periods throughout different seasons. Additionally, to specifically test dispersal capabilities in a case study such as this, it may be pertinent to conduct a mark/release/recapture-based experiment in addition to the population genetic analyses presented here. Should any flies be proven to disperse from the poultry farm to

a nearby home, culturing and studying of bacterial strains found on the specimens may be prudent to assess the transmission potential of pathogens and antibiotic resistance. Nonetheless, this study indicates that house flies in this region outside of Phoenix, Arizona are largely part of a widespread, intermixing population.

CHAPTER 4. CONCLUSIONS

Though a case report utilizing entomology can be found as far back as 13th century China, the majority of the use and research of forensically relevant insects is relatively recent [111]. The first modern use of forensic entomology took place in 1855 in France [112], and after that report, forensic entomology slowly made its way to popularity. Researchers and taxonomists began investigating forensically relevant insects in the first half of the twentieth century [1], and, as demonstrated by the present studies and the plethora of those cited, research on forensically relevant insects remains prevalent today.

Though the field of forensic entomology continues to grow, it is not without its issues. In short, a few issues, or rather variations, that may affect a corpse's decomposition and therefore its PMI calculation, include differences in insect attraction to corpses, effects of larval aggregations, differences in insect developmental thresholds and rates, and the impacts of climate, geographical region, season, and habitat [1]. It is because of these variations that research on forensically relevant insect continues and allows us to learn new information each and every day. Each new piece of information uncovered helps forensic entomologists account for these variations and utilize more precise data when examining insects in a forensic context.

The first study presented here investigated behavioral differences and possible subspecies divisions within the blow fly *L. cuprina*. *Lucilia cuprina* is present in North America and is attracted to carrion, and therefore may be used to estimate a PMI in a criminal context. However, because of morphological similarities of *L. cuprina* and its sister-species *L. sericata,* as well as known hybridization between the two, morphological identifications as well as DNA-based identifications may be flawed and lead to misrepresentations of the PMI. Our data indicate that the variations within *L. cuprina* are more complicated than previously thought and may have raised more questions than answers. *Lucilia cuprina* specimens collected in Indiana, USA possess a novel combination of nuclear and mitochondrial haplotypes which may be indicative of modern hybridization. These results highlight the difficulties of utilizing DNA-based identification techniques within the genus *Lucilia* and suggest that known data concerning *L. cuprina* from other regions may not be applicable to Indiana specimens in a forensic context. It is important to research the extent to which North American *L. cuprina* are related to *L. cuprina* that are involved in myiasis, because if it were discovered that North American *L. cuprina* could

and did strike living animals, the set of assumptions utilized in PMI calculations would need to be altered. If American *L. cuprina* are found to be involved in myiasis, developmental datasets would have to be reevaluated and studies redone to account for newly discovered behavioral differences in the region. Consider the possibility of a scenario where a victim in North America is alive, possesses a wound that is infested by *L. cuprina* maggots who are feeding on the flesh, and then the victim subsequently dies, and the maggots continue to feed. If the forensic entomologist on the case were to assume that North American *L. cuprina* are not involved in myiasis, he or she may conclude that the maggots must have been present on the body longer to have ingested the amount of flesh that is missing, or the maggots may appear older due to increased size, when truly they had already been feeding on the victim prior to his death. In either of these situations, without knowing that the maggots could be involved in myiasis, a PMI calculation may be skewed due to an incorrect set of assumptions about the species and its behavior. If presented in court, a skewed PMI calculation could implicate an innocent suspect, or exonerate a guilty one.

The second study presented here investigated the population structure and relatedness of house fly (*M. domestica*) specimens collected in and around a poultry farm in Arizona, USA. House flies are considered forensically relevant because they are prevalent in areas of human activity can be attracted to carrion [1]. House flies were particularly forensically relevant in our study due to a civil suit filed against the poultry farm by homeowners surrounding the farm who complained of flies emanating to their homes. Our data indicate that flies from in and around the neighbors' homes were genetically similar to those collected from the farm, and that it is not possible to conclude that the flies did not emanate from the farm. In a broader perspective, our results suggest that the house fly populations in this region of Arizona, USA largely intermix with one another and experience significant differences in genetic structure during different seasons. This data could be utilized to create more effective pest control measures in this region. Additionally, to our knowledge, this work marks the first use of AFLP on *M. domestica,* which may guide future research possibilities.

Overall, this work provides a better understanding of genetic differences in two forensically important insects in different regions of the United States. Our results build upon knowledge obtained from previous studies and may guide the direction of future related studies.

This research can be applied to forensic investigations, development of DNA-based identification methods, taxonomy and identification of subspecies, and pest management.

REFERENCES

1. Byrd, J.H. and J.L. Castner, *Forensic entomology : the utility of arthropods in legal investigations*. 2nd ed. ed. 2009: Taylor & Francis.

2. Anderson, G. and S.L. VanLaerhoven, *Initial studies on insect succession on carrion in southwestern British Columbia.* Journal of Forensic Science, 1996. **41**(4): p. 617-625.
3. Byrd, J.H. and J.K. Tomberlin, *Forensic entomology: the utility of arthropods in legal investigations*. 2019: CRC press.

4. Amendt, J., et al., *Best practice in forensic entomology--standards and guidelines.* Int J Legal Med, 2007. **121**(2): p. 90-104.

5. Barton, P.S., et al., *The role of carrion in maintaining biodiversity and ecological processes in terrestrial ecosystems.* Oecologia, 2013. **171**(4): p. 761-772.

6. Payne, J.A., *A Summer Carrion Study of the Baby Pig Sus Scrofa Linnaeus.* Ecology, 1965. **46**(5): p. 592-602.

7. Templeton, A.R. and I. Wiley, *Population Genetics and Microevolutionary Theory*. 2006, Hoboken, N.J.: Wiley-Liss.

8. Hartl, D.L. and A.G. Clark, *Principles of population genetics*. 4th ed. ed. 2007: Sinauer Associates.

9. Smith, K.G., *A manual of forensic entomology*. 1986.

10. Wells, J.D., R. Wall, and J.R. Stevens, *Phylogenetic analysis of forensically important Lucilia flies based on cytochrome oxidase I sequence: a cautionary tale for forensic species determination.* Int J Legal Med, 2007. **121**(3): p. 229-33.

11. Sperling, F.A., G.S. Anderson, and D.A. Hickey, *A DNA-based approach to the identification of insect species used for postmortem interval estimation.* J Forensic Sci, 1994. **39**(2): p. 418-27.

12. Wells, J.D. and J.R. Stevens, *Application of DNA-based methods in forensic entomology.* Annu Rev Entomol, 2008. **53**: p. 103-20.

13. Simon, C., et al., *Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers.* Annals of the Entomological Society of America, 1994. **87**(6): p. 651-701.

14. Hebert, P.D., et al., *Biological identifications through DNA barcodes.* Proc Biol Sci, 2003. **270**(1512): p. 313-21.

15. Williams, K. and M.H. Villet, *Ancient and modern hybridization between Lucilia sericata and L. cuprina (Diptera: Calliphoridae).* European Journal of Entomology, 2013. **110**(2): p. 187-196.

16. Williams, K.A., J. Lamb, and M.H. Villet, *Phylogenetic radiation of the greenbottle flies (Diptera, Calliphoridae, Luciliinae).* Zookeys, 2016(568): p. 59-86.

17. Tourle, R., D.A. Downie, and M.H. Villet, *Flies in the ointment: a morphological and molecular comparison of Lucilia cuprina and Lucilia sericata (Diptera: Calliphoridae) in South Africa.* Medical & Veterinary Entomology, 2009. **23**(1): p. 6-14.

18. Stevens, J.R., R. Wall, and J.D. Wells, *Paraphyly in Hawaiian hybrid blowfly populations and the evolutionary history of anthropophilic species.* Insect Mol Biol, 2002. **11**(2): p. 141-8.

19. Debry, R.W., et al., *mtDNA-based identification of Lucilia cuprina (Wiedemann) and Lucilia sericata (Meigen) (Diptera: Calliphoridae) in the continental United States.* Forensic Sci Int, 2010. **202**(1-3): p. 102-9.

20. Owings, C.G., et al., *Developmental variation among Cochliomyia macellaria Fabricius (Diptera: Calliphoridae) populations from three ecoregions of Texas, USA.* Int J Legal Med, 2014. **128**(4): p. 709-17.

21. Tarone, A.M. and D.R. Foran, *Components of developmental plasticity in a Michigan population of Lucilia sericata (Diptera: Calliphoridae).* J Med Entomol, 2006. **43**(5): p. 1023-33.

22. Picard, C. and J. Wells, *Survey of the Genetic Diversity of Phormia regina (Diptera: Calliphoridae) Using Amplified Fragment Length Polymorphisms.* Journal of medical entomology, 2009. **46**: p. 664-70.

23. Picard, C.J. and J.D. Wells, *The population genetic structure of North American Lucilia sericata (Diptera: Calliphoridae), and the utility of genetic assignment methods for reconstruction of postmortem corpse relocation.* Forensic Science International, 2010. **195**(1): p. 63-67.

24. Wiley, E.O. and B.S. Lieberman, *Phylogenetics: theory and practice of phylogenetics systematics*. 2011: Wiley-Blackwell.

25. Stevens, J. and R. Wall, *Genetic relationships between blowflies (Calliphoridae) of forensic importance.* Forensic Science International, 2001. **120**(1-2): p. 116-123.

26. Stevens, J.R., *The evolution of myiasis in blowflies (Calliphoridae).* International Journal for Parasitology, 2003. **33**(10): p. 1105.

27. Stevens, J. and R. Wall, *The evolution of ectoparasitism in the genus Lucilia (Diptera: Calliphoridae).* International Journal for Parasitology, 1997. **27**(1): p. 51-59.

28. Stevens, J. and R. Wall, *Species, Sub-Species and Hybrid Populations of the Blowflies Lucilia cuprina and Lucilia sericata (Diptera: Calliphoridae).* Proceedings: Biological Sciences, 1996. **263**(1375): p. 1335-1341.

29. Stevens, J. and R. Wall, *Genetic variation in populations of the blowflies Lucilia cuprina and Lucilia sericata (Diptera: Calliphoridae). Random amplified polymorphic DNA analysis and mitochondrial DNA sequences.* Biochemical Systematics and Ecology, 1997. **25**(2): p. 81-97.

30. Broadmeadow, M., et al., *The pathognesis of flystrike in sheep.* 1984.

31. Levot, G. and G. Casburn, *Primefact 485: Sheep blowflies.* NSW Department of Primary Industries, 2016.

32. Lane, J., et al., *Priority list of endemic diseases for the red meat industries.* 2015.

33. Anstead, C.A., et al., *A blow to the fly—Lucilia cuprina draft genome and transcriptome to support advances in biology and biotechnology.* Biotechnology advances, 2016. **34**(5): p. 605-620.

34. Williams, R.E., et al., *Livestock Entomology*. 1985, New York: Wiley.

35. Phillips, C.J., *A review of mulesing and other methods to control flystrike (cutaneous myiasis) in sheep.* Animal Welfare, 2009. **18**(2): p. 113-121.

36. French, N.P., et al., *Prevalence, regional distribution and control of blowfly strike in England and Wales.* Vet Rec, 1992. **131**(15): p. 337-42.

37. Owings, C.G. and C.J. Picard, *New Distribution Record for Lucilia cuprina (Diptera: Calliphoridae) in Indiana, United States.* Journal of insect science (Online), 2018. **18**(4): p. 8.

38. Waterhouse, D.F. and S.J.A.J.o.B.S. Paramonovo, *The Status of the Two Species of Lucilia (Diptera, Calliphoridae) Attacking Sheep in Austhalia.* 1950. **3**(3): p. 310-336.

39. Whitworth, T., *Keys to the genera and species of blow flies (Diptera: Calliphoridae) of America north of Mexico*. 2006. p. 689.

40. Stoganovich, C., H. Pratt, and E. Bennington, *Fly larvae: key to some species of public health importance*. 1962, CDC US Department of Health, Education and Welfare: Atlanta, GA.

41. Williams, K.A. and M.H. Villet, *Morphological identification of Lucilia sericata, Lucilia cuprina and their hybrids (Diptera, Calliphoridae).* ZooKeys, 2014(420): p. 69-85.

42. McDonagh, L.M. and J.R. Stevens, *The molecular systematics of blowflies and screwworm flies (Diptera: Calliphoridae) using 28S rRNA, COX1 and EF-1α: insights into the evolution of dipteran parasitism.* Parasitology, 2011. **138**(13): p. 1760-1777.

43. Simon, C., et al., *Incorporating Molecular Evolution into Phylogenetic Analysis, and a New Compilation of Conserved Polymerase Chain Reaction Primers for Animal Mitochondrial DNA.* 2006. **37**(1): p. 545-579.

44. Madeira, F., et al., *The EMBL-EBI search and sequence analysis tools APIs in 2019.* Nucleic Acids Res, 2019. **47**(W1): p. W636-W641.

45. Hall, T.A., *BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.* 1999, Nucl. Acids. Symp. Ser. 41:95-98.

46. Andere, A.A., et al., *Genome sequence of Phormia regina Meigen (Diptera: Calliphoridae): implications for medical, veterinary and forensic research.* BMC genomics, 2016. **17**(1): p. 842-842.

47. Swofford, D., *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10*. Vol. Version 4.0. 2002.

48. Dereeper, A., et al., *Phylogeny.fr: robust phylogenetic analysis for the non-specialist.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W465-9.

49. Dereeper, A., et al., *BLAST-EXPLORER helps you building datasets for phylogenetic analysis.* BMC Evol Biol, 2010. **10**: p. 8.

50. Bunchu, N., et al., *Morphology and developmental rate of the blow fly, Hemipyrellia ligurriens (Diptera: Calliphoridae): Forensic entomology applications.* Journal of parasitology research, 2012. **2012**.

51. White, R.S., D. Aubertin, and J. Smart, *The fauna of British India, including the remainder of the oriental region : Diptera. Vol.6, Vol.6*. 1940, London.

52. Klong-Klaew, T., et al., *Spatial Distribution of Forensically Significant Blow Flies in Subfamily Luciliinae (Diptera: Calliphoridae), Chiang Mai Province, Northern Thailand: Observations and Modeling Using GIS.* Insects, 2018. **9**(4): p. 181.

53. Kurahashi, H., *The Tribe Luciliini from Australian and Oriental Regions: I. Genus Hypopygiopsis TOWNSEND (Diptera, Calliphoridae).* 昆蟲, 1977. **45**(4): p. 553-562.

54. Kurahashi, H., *Blow flies of New Guinea, Bismarck Archipelago, and Bougainville Island (Diptera, Calliphoridae).* 1987.

55. Sukontason, K., et al., *Survey of forensically-relevant fly species in Chiang Mai, Northern Thailand.* Journal of vector ecology : journal of the Society for Vector Ecology, 2003. **28**(1): p. 135-138.

56. DeBry, R.W., et al., *DNA-Based Identification of Forensically Important Lucilia (Diptera: Calliphoridae) in the Continental United States* DNA-Based Identification of Forensically Important Lucilia (Diptera: Calliphoridae) in the Continental United States.* Journal of Forensic Sciences, 2013. **58**(1): p. 73-78.

57. Sonet, G., et al., *Why is the molecular identification of the forensically important blowfly species Lucilia caesar and L. illustris (family Calliphoridae) so problematic?* Forensic Science International, 2012. **223**(1-3): p. 153-159.

58. Stevens, J. and R. Wall, *Classification of the genus Lucilia (Diptera: Calliphoridae): a preliminary parsimony analysis.* Journal of Natural History, 1996. **30**(7): p. 1087-1094.

59. Park, S.H., et al., *Using the developmental gene bicoid to identify species of forensically important blowflies (Diptera: Calliphoridae).* BioMed Research International, 2013. **2013**.

60. Hall, M. and R. Wall, *Myiasis of humans and domestic animals*, in *Advances in parasitology*. 1995, Elsevier. p. 257-334.

61. Warman, G., et al., *Analysis of the circadian clock gene period in the sheep blow fly Lucilia cuprina.* Genetics Research, 2000. **75**(3): p. 257-267.

62. Szalanski, A.L., et al., *Detection of Campylobacter and Escherichia coli O157:H7 from filth flies by polymerase chain reaction.* Medical and Veterinary Entomology, 2004. **18**(3): p. 241-246.

63. Organization, W.H., *Vector control series: the housefly: training and information guide*. 1986, World Health Organization.

64. Awache, I. and A. Farouk, *Bacteria and fungi associated with houseflies collected from cafeteria and food Centres in Sokoto.* FUW Trends Scie Technol J, 2016. **1**(1): p. 123-5.

65. Khamesipour, F., et al., *A systematic review of human pathogens carried by the housefly (Musca domestica L.).* BMC Public Health, 2018. **18**(1): p. 1049.

66. Iqbal, W., et al., *Role of housefly (Musca domestica, Diptera; Muscidae) as a disease vector; a review.* Journal Entomology and Zoology Studies, 2014. **2**(2): p. 159-163.

67. Tsagaan, A. and K. Okado, *Study of pathogenic bacteria detected in fly samples using universal primer-multiplex PCR.* Mongolian Journal of Agricultural Sciences, 2015. **15**(2): p. 27-32.

68. Lysyk, T.J. and R.C. Axtell, *Movement and Distribution of House Flies (Diptera: Muscidae) Between Habitats in Two Livestock Farms.* Journal of Economic Entomology, 1986. **79**(4): p. 993-998.

69. Schoof, H.F. and R.E. Siverly, *Multiple Release Studies on the Dispersion of Musca domestica at Phoenix, Arizona.* Journal of Economic Entomology, 1954. **47**(5): p. 830-838.

70. Nazni, W.A., et al., *Determination of the flight range and dispersal of the house fly, Musca domestica (L.) using mark release recapture technique.* Trop Biomed, 2005. **22**(1): p. 53-61.

71. Chakrabarti, S., S. Kambhampati, and L. Zurek, *Assessment of House Fly Dispersal between Rural and Urban Habitats in Kansas, USA.* Journal of the Kansas Entomological Society, 2010. **83**(2): p. 172-188, 17.

72. Ahmad, A., et al., *Insects in confined swine operations carry a large antibiotic resistant and potentially virulent enterococcal community.* BMC Microbiology, 2011. **11**(1): p. 23.

73. Marquez, J.G. and E.S. Krafsur, *Mitochondrial diversity evaluated by the single strand conformation polymorphism method in African and North American house flies (Musca domestica L.).* Insect Mol Biol, 2003. **12**(2): p. 99-106.

74. Krafsur, E.S., J.M. Helm, and W.C. Black, *Genetic diversity at electrophoretic loci in the house fly,Musca domestica L.* Biochemical Genetics, 1992. **30**(5): p. 317-328.

75. Wright, J.W. and R. Pal, *Genetics of insect vectors of disease*. 1967, Amsterdam, &c., Elsevier Publ. Co. p. xix + 794 pp.

76. Black, W.C. and E.S. Krafsur, *Seasonal breeding structure in house fly, Musca domestica L., populations.* Heredity, 1986. **56**(3): p. 289-298.

77. Black, W.C., IV and E.S. Krafsur, *Population Biology and Genetics of Winter House Fly (Diptera: Muscidae) Populations.* Annals of the Entomological Society of America, 1986. **79**(4): p. 636-644.

78. Marquez, J.G. and E.S. Krafsur, *Gene flow among geographically diverse housefly populations (Musca domestica L.): a worldwide survey of mitochondrial diversity.* J Hered, 2002. **93**(4): p. 254-9.

79. Chakrabarti, S., et al., *Characterization of microsatellite loci in the house fly, Musca domestica L. (Diptera: Muscidae).* Molecular Ecology Notes, 2004. **4**(4): p. 728-730

.

80. Endsley, M.A., M.D. Baker, and E.S. Krafsur, *Microsatellite loci in the house fly, Musca domestica L. (Diptera: Muscidae).* Molecular Ecology Notes, 2002. **2**(1): p. 72-74.

81. Feldmeyer, B., I. Pen, and L.W. Beukeboom, *A microsatellite marker linkage map of the housefly, Musca domestica: evidence for male recombination.* Insect Molecular Biology, 2010. **19**(4): p. 575-581.

82. Krafsur, E.S., et al., *Geographic Differentiation in the House Fly Estimated by Microsatellite and Mitochondrial Variation.* Journal of Heredity, 2005. **96**(5): p. 502-512.

83. Picard, C.J., M.H. Villet, and J.D. Wells, *Amplified fragment length polymorphism confirms reciprocal monophyly in Chrysomya putoria and Chrysomya chloropyga: a correction of reported shared mtDNA haplotypes.* Medical and Veterinary Entomology, 2012. **26**(1): p. 116-119.

84. Blears, M.J., et al., *Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications.* Journal of Industrial Microbiology and Biotechnology, 1998. **21**(3): p. 99-114.

85. Reineke, A., P. Karlovsky, and C. Zebitz, *Amplified fragment length polymorphism analysis of different geographic populations of the gypsy moth, Lymantria dispar (Lepidoptera: Lymantriidae).* Bulletin of Entomological Research, 1999. **89**(1): p. 79-88.

86. Cervera, M.T., et al., *Genetic relationships among biotypes of Bemisia tabaci (Hemiptera: Aleyrodidae) based on AFLP analysis.* Bulletin of Entomological Research, 2000. **90**(5): p. 391-396.

87. Timmermans, M., et al., *Genetic structure in Orchesella cincta (Collembola): strong subdivision of European populations inferred from mtDNA and AFLP markers.* Molecular Ecology, 2005. **14**(7): p. 2017-2024.

88. Núñez, Y., et al., *Analysis of population structure of Erysiphe necator using AFLP markers.* Plant Pathology, 2006. **55**(5): p. 650-656.

89. Green, M.R. and J. Sambrook, *Isolation of High-Molecular-Weight DNA Using Organic Solvents.* Cold Spring Harbor Protocols, 2017. **2017**(4): p. pdb.prot093450.

90. Greenberg, B., *Flies and Disease: I. Ecology, Classification, and Biotic Associations*. Vol. 1. 1971, Princeton, NJ: Princeton University Press.

91. Dodge, H., *Diptera: pictorial key to the principal families of public health importance.* US Department of Health, Education, and Welfare, Communicable Disease Center (ed.). Pictorial Keys to Arthropods, Reptiles, Birds and Mammals of Public Health Significance. CDC Training Branch, Atlanta, Georgia, 1966: p. 120-133.

92. Dodge, H., *Domestic Flies: Pictorial Key to Common Species in the US.* Public Health Service. Communicable Diesase Center Training Branch. Atlanta, Georgia, 1958.

93. Dodge, H., *Domestic flies: pictorial key to common species in southern US In.* CDC Pictorial Keys: Arthropods, Reptiles, Birds and Mammals of Public Health Significance, 1953.

94. Scott, H. and M. Borom, *Domestic Flies: Pictorial Key to Common Species.* Atlanta, GA, USA: US Department of Health, Education and Welfare, 1962.

95. Contributors, Z.E.P. *The Beetle D N A Lab*. 2017; Available from: https://zsm-entomology.de/w/index.php?title=The_Beetle_D_N_A_Lab&oldid=1216.

96. Gompert, Z., et al., *Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly.* Molecular Ecology, 2006. **15**(7): p. 1759-1768.

97. Duchesne, P. and L. Bernatchez, *AFLPOP: A computer program for simulated and real population allocation, based on AFLP data.* Molecular Ecology Notes, 2002. **2**(3): p. 380-383.

98. Hardy, O.J. and X. Vekemans, *SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels.* Molecular ecology notes, 2002. **2**(4): p. 618-620.

99. Ritland, K., *Estimators for pairwise relatedness and individual inbreeding coefficients.* Genetical Research, 1996. **67**(2): p. 175-185.

100. Pritchard, J.K., M. Stephens, and P. Donnelly, *Inference of population structure using multilocus genotype data.* Genetics, 2000. **155**(2): p. 945-959.

101. Earl, D.A., *STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.* Conservation genetics resources, 2012. **4**(2): p. 359-361.

102. Peakall, R. and P.E. Smouse, *GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research.* Molecular ecology notes, 2006. **6**(1): p. 288-295.

103. Peakall, R. and P.E. Smouse, *GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update.* Bioinformatics, 2012. **28**(19): p. 2537.

104. Excoffier, L., P.E. Smouse, and J.M. Quattro, *Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data.* Genetics, 1992. **131**(2): p. 479-91.

105. Diniz-Filho, J.A.F., et al., *Mantel test in population genetics.* Genetics and Molecular Biology, 2013. **36**(4): p. 475-485.

106. Picard, C.J. and J.D. Wells, *A test for carrion fly full siblings: a tool for detecting postmortem relocation of a corpse.* J Forensic Sci, 2012. **57**(2): p. 535-8.

107. Underground, W. *Phoeniz, AZ Weather History*. Available from: https://www.wunderground.com/history/daily/us/az/tonopah/KPHX.

108. Elvin, M.K. and E.S. Krafsur, *Relationship between Temperature and Rate of Ovarian Development in the House Fly, Musca domestica L. (Diptera: Muscidae).* Annals of the Entomological Society of America, 1984. **77**(1): p. 50-55.

109. Malik, A., N. Singh, and S. Satya, *House fly (Musca domestica): a review of control strategies for a challenging pest.* J Environ Sci Health B, 2007. **42**(4): p. 453-69.

110. Mueller, U.G. and L.L. Wolfenbarger, *AFLP genotyping and fingerprinting.* Trends in ecology & evolution, 1999. **14**(10): p. 389-394.

111. Benecke, M., *A brief history of forensic entomology.* Forensic Science International, 2001. **120**(1): p. 2-14.

112. Bergeret, M., *Homicide of a new-born child found in a chimney and its natural mummification. Determination of a postmortem interval by the use of insect larvae and their metamorphasis.* Ann. Hyg. Legal Med, 1855. **4**: p. 442-452.