University of Mississippi

eGrove

Honors Theses

Honors College (Sally McDonnell Barksdale Honors College)

Spring 5-8-2022

Invasion of the Northeastern United States by the Southern Pine Beetle, Dendroctonus Frontalis, and the Impacts on Signatures of Isolation by Distance Assessed using Mitochondrial DNA

Lora Grace Holman

Follow this and additional works at: https://egrove.olemiss.edu/hon_thesis

 Part of the Biology Commons, and the Genetics and Genomics Commons

Recommended Citation

Holman, Lora Grace, "Invasion of the Northeastern United States by the Southern Pine Beetle, Dendroctonus Frontalis, and the Impacts on Signatures of Isolation by Distance Assessed using Mitochondrial DNA" (2022). *Honors Theses*. 2607.
https://egrove.olemiss.edu/hon_thesis/2607

This Undergraduate Thesis is brought to you for free and open access by the Honors College (Sally McDonnell Barksdale Honors College) at eGrove. It has been accepted for inclusion in Honors Theses by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.

Invasion of the Northeastern United States by the Southern Pine Beetle, *Dendroctonus frontalis*, and the Impacts on Signatures of Isolation by Distance Assessed Using Mitochondrial DNA

by
Lora Grace Holman

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of the requirements of the Sally McDonnell Barksdale Honors College

Oxford, MS
April 2022

Approved by

Advisor: Doctor Ryan Garrick

Reader: Doctor Susan Balenger

Reader: Doctor Ísis C. Arantes

© 2022
Lora Grace Holman
ALL RIGHTS RESERVED

ACKNOWLEDGEMENTS

I would like to thank the USDA Forest Service for funding for this project, their collaborators for collection of specimens, and Drs. Nathan Havill and Ísis Arantes for data generation. I would also like to thank Dr. Garrick and the members of the Garrick lab for all help that I received on this project. Finally, I would like to thank the SMBHC for their support and guidance during this time.

ABSTRACT

The southern pine beetle (SPB) is a small, black beetle that parasitizes pine trees across the United States and Central America. Recently it has been recognized that the range of SPB is not limited to the southern United States, but rather extends northward across the eastern United States, which raises concerns for pine trees that have never encountered the pest before. Using mitochondrial DNA sequences from a section of the cytochrome oxidase I gene and four different measures of genetic distance compared to geographic distance, we tested for evidence of isolation by distance (IBD) among sampled SPB to see whether this range expansion was recent, and thus showed no evidence of IBD, or relatively old, such that there would be a positive relationship between genetic and geographic distance. We found that there was only one, possibly erroneous, statistically significant IBD relationship, and surprisingly, it was negative, which suggested a lack of IBD that could possibly be explained by human interference. When comparing two measures of genetic distance that are informative over different time scales (F_{ST} and Φ_{ST}), it was discovered that the population genetic structure, though weak, is likely quite old.

TABLE OF CONTENTS

List of Tables & Figures	6
List of Abbreviations	7
Introduction	8
Methods	13
Results	19
Discussion	25
Bibliography	30

LIST OF TABLES & FIGURES

Figure 1	An image of the southern pine beetle	8
Figure 2	Map of the Eastern United States, including sample sites	13
Figure 3	Breakdown of the workflow of this study	15
Table 1	General breakdown of the raw data used in this study	20
Table 2	Significance and relationship of data sets versus distance measure	22
Figure 4	An example isolation by distance plot using the significant result	23
Figure 5	Bar graph comparing global F_{ST} and Φ_{ST}	24

LIST OF ABBREVIATIONS

COI	Cytochrome Oxidase Subunit I gene
IBD	Isolation by Distance
ML	Maximum Likelihood
mtDNA	Mitochondrial DNA
PCR	Polymerase Chain Reaction
SPB	Southern Pine Beetle(s)
USDA	United States Department of Agriculture

Introduction

The southern pine beetle (SPB), *Dendroctonus frontalis* (Coleoptera: Scolytidae; Figure 1) is a small, black, herbivorous insect that parasitizes and kills pine trees (*Pinus* spp.) in North America (Havill et al. 2019). Outbreaks of SPB infestations cause negative impacts on both the economy and ecology, including multi-million dollar disruptions of the commercial timber industry and devastation of pine forests across the United States and Central America (Havill et al. 2019). Originally, the range of SPB was restricted primarily to the southeastern United States, from Texas to Florida (Dodds et al. 2018). It is understood now that the reach of SPB has expanded beyond that range relatively recently, spreading northward across the eastern United States (Dodds et al. 2018).

Figure 1: An image of the southern pine beetle (Dendroctonus frontalis), courtesy of USDA Forest Service, Bugwood.org.

The detection of this apparent range expansion has raised questions and concerns. Have SPB always been native to these northern areas, or has there been a more recent, rapid expansion? If the expansion is recent, how will these parasitic beetles affect pine trees that have never been attacked by SPB before? Through the course of this study, we aimed to help uncover the truth about the northern expansion of SPB using mitochondrial DNA (mtDNA) analysis to assess evidence of isolation by distance (i.e., a positive correlation between genetic and geographic distance), as exemplified by species such as the *Eurycea* salamanders in Texas (Lucas et al. 2008), to discern whether or not the aforementioned expansion was recent and rapidly occurring or gradual.

Global Warming and Range Expansion of SPB

Over the past decades climate change, colloquially referred to as global warming, has become a matter of ever increasing concern. Trends show that average temperatures in the United States alone have risen by 1.5°F since the end of the 19th century, with approximately 80% of that increase occurring since 1980 (Brian Kahn, climatecentral.org). The rising average temperatures indirectly result in a cascade of ecological changes. Populations of arthropods like SPB are naturally controlled by low winter temperatures, with SPB facing mortality at -16°C specifically (Monro et al. 2020). As average temperatures rise, new geographic areas become available to SPB that previously were uninhabitable due to their low extreme temperatures. In a study by Munro et al. (2020), it was found that the number of SPB captured per day increased with higher temperatures. It follows, then, that if SPB rates increase with temperature inclines,

the spread of SPB into the northern United States could reasonably be linked to trends of global warming, but the method and speed of dispersal by which the beetles spread is still unknown.

Expansion of species is a balancing act between selection, which can drive expansion, and chance events like drift (Williams et al. 2019). Bark beetles like SPB face obstacles to dispersal, the most important of which include their body mass, lipid content, external temperature, wind conditions, and various landscape conditions (Jones et al. 2019). However, other methods of dispersal are indirectly available to these arthropods in our modern age, such as accidental transport through the commercial distribution of timber. Whether the spread of SPB was achieved through natural or human-mediated means, the range of SPB has expanded past what was previously known (Dodds et al. 2018). In the past, populations of SPB were confined to the south, primarily from Texas to Florida to New Jersey (Dodds et al. 2018).

Mitochondrial DNA and the Cytochrome Oxidase Subunit I Gene

This study makes use of mtDNA sequence data, specifically, a section of the cytochrome oxidase subunit I (COI) gene, as a genetic marker, to assess whether there is evidence for isolation by distance (IBD) in SPB. Mitochondrial DNA has been used as a tool for understanding the biology of natural populations, from dispersal and gene flow to the connectivity between local populations, because this gene region has several advantageous features.

On average, mtDNA has a faster mutation rate than nuclear DNA (Ballard & Whitlock, 2004). This is likely due to mtDNA having fewer DNA repair systems in place for maintaining the integrity of genetic material, or, simply put, reduced levels of “proofreading” during replication (Zinovkina, 2018). While the reason for higher mutation levels is disputed, the

relatively fast mutation rate of the mitochondrial genome and the subsequent persistence of these mutations within organisms and their offspring makes it an "information-rich" marker for population genetics because differences among individuals and populations are prerequisites for downstream analysis of details like population size, future growth or decay of populations, or demographic history of populations (Putman & Carbone, 2014).

Mitochondrial DNA sequence data is relatively easy to collect, given that eukaryotic organisms usually have many mitochondria per cell, each with multiple copies of mtDNA chromosomes (Ballard & Whitlock, 2004). Kocher et al. (1989) published a set of highly conserved primers that could amplify mtDNA from a large range of taxa by polymerase chain reaction (PCR). This database of known primers and ease of amplification compounded with other known traits of mtDNA–such as the maternal mode of inheritance, lack of recombination, and the haploid nature that allows for simplicity during analysis–drew many researchers over time to study systems using mtDNA (Ballard & Whitlock, 2004). Subsequently, many mtDNA sequences have been recorded, giving a breadth of available and reliable data for future researchers to access.

Gene Flow and how Isolation by Distance Occurs

The term "gene flow" refers to the collection of mechanisms that move alleles from one population to another within a species (Slatkin, 1985). This sharing of genetic material blurs the line between distinct populations within a species, while also allowing for the possible expansion of available genetic material within a population via swapping of alleles that originally arose in just one population, or the creation of entirely new mutations through the recombination of different alleles. The transfer of genetic material from one population to another when

populations are geographically close to one another and without physical barriers separating them is a factor that needs to be taken into consideration, especially when discussing possible IBD within a species.

Many species exhibit a genetic population structure that is characterized by IBD, which suggests that due to dispersal, individuals or subpopulations that are geographically close to one another are more likely to be genetically similar than those that are further apart (Meirmans, 2012). Time, distance, and a lack of sufficient dispersal allow for populations to accumulate genetic differences due to mutation and drift. Evidence for IBD would provide support for the hypothesis that SPB gradually spread throughout the eastern United States over a long period of time and have simply remained in nearly undetectable levels in the northern populations until recent observation. Conversely, if there is no strong evidence for IBD, there are two possible explanations–the spread of SPB was simply too rapid for IBD to occur, or there is a very high level of gene flow among the populations across North America despite geographic distance separating them.

Methods

Population Sampling and Genetic Data Acquisition

To begin this project, 382 SPB individuals were collected from 17 sites (referred to as "populations" herein) across the eastern United States, from Louisiana to Rhode Island (Figure 2).

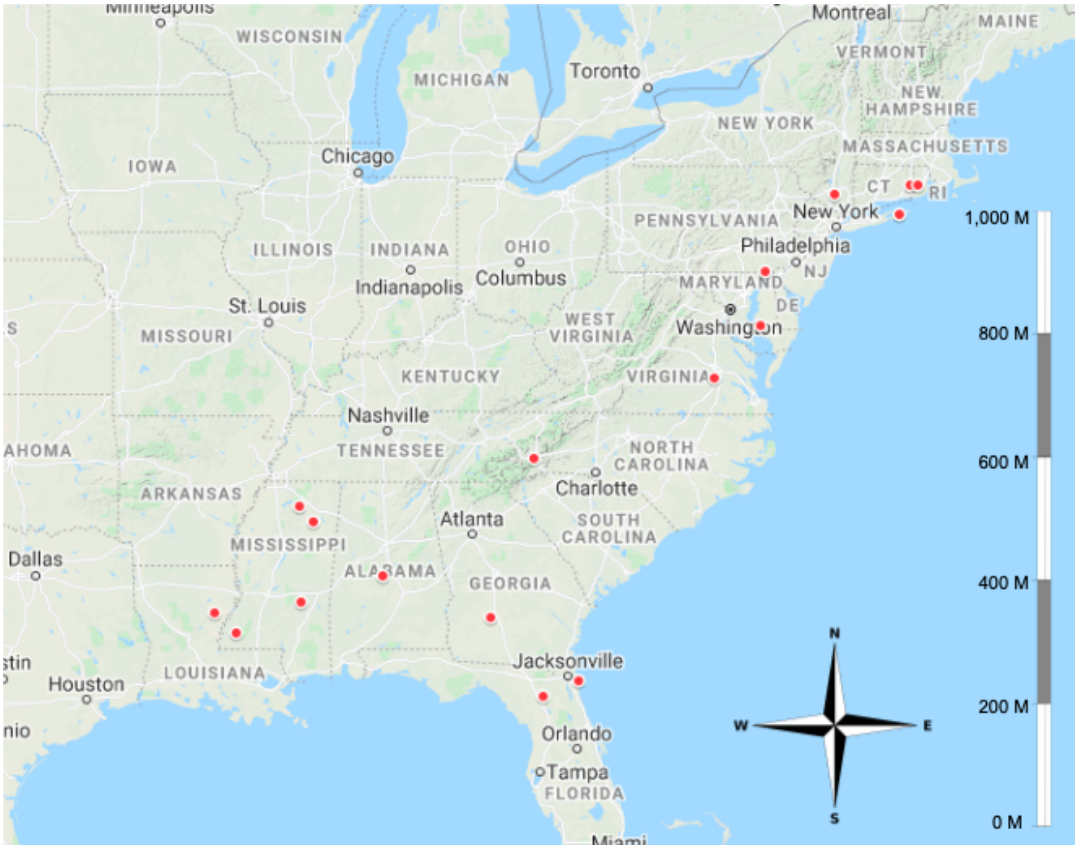

Figure 2: A map showing the sampling sites that were included in this study. Each red dot represents a location from which beetles were sampled across the eastern United States.

Beetles were collected by USDA Forest Service collaborators using pheromone traps. After capture, the sex of each individual beetle was determined by Dr. Nathan Havill through microscopic examination. DNA sequence data from a section of the mitochondrial COI gene region from members of the eastern North American SPB populations previously published by Havill et al. (2019) were combined with newly generated COI sequences gathered by Dr. Ísis C. Arantes using the same primers and PCR amplification conditions, and this pooled data was analyzed here. Altogether, the data set for this study included mitochondrial COI sequences of 658 base pairs in length.

Data Analysis

As stated above, the goals of this study were to examine the evidence for IBD. As part of this investigation, the data were partitioned in several ways, based on the sex of SPB individuals, and reanalyzed using several alternative measures of genetic distance, in order to determine what impact these might have on the conclusions. To do this, multiple steps were taken and multiple different computer programs were used to analyze or organize the data. As such, for the purpose of this paper, the process of the methods will be broken down into four distinct steps, illustrated in the workflow diagram below (Figure 3).

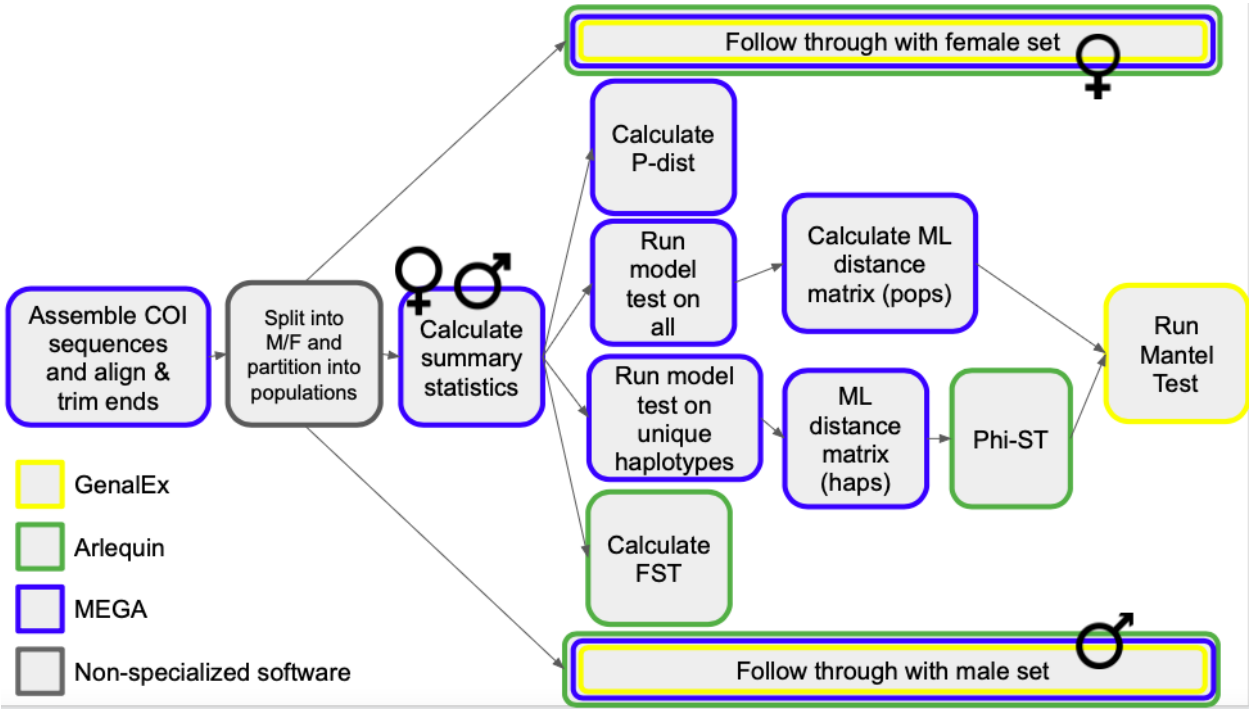

Figure 3: A workflow diagram of the analytical pipeline used in this study. The individual steps are color-coded relating to which computer program was used during each step.

DNA Sequence Data Assembly

Using the computer program MEGA v.6.0 (Tamura et al. 2013), all of the mitochondrial COI sequences were aligned and trimmed to a uniform length. Trimming involved getting rid of long strings of unknown bases on the ends of samples that were not sequenced as clearly as the others. This trimming formed a measure of quality control, as it reduced the amount of missing data included in the final data matrix. Then, based on the information that Dr. Nathan Havill collected, the individual sequences were divided into three sets for later analysis: one set that consisted of all of the specimens, an all-female set, and an all-male set.

Partitioning data into Populations, and Characterizing Levels of Diversity

After parsing the data into three groups based on the sex of SPB individuals, they were then further organized and partitioned into populations based on the sampling sites that they were collected from (Figure 2). For the full dataset and male-only dataset, sequences were parsed into 17 populations, whereas the female-only dataset was only divided into 16 populations, as no female beetles were gathered from the Holly Springs, Mississippi sampling site. There were noticeable differences in the sample sizes of males and females within many of the populations, including one female specimen out of the total 19 in the North Carolina population, or the one-to-17, female-to-male specimen ratio in the Connecticut population.

Two measures of genetic diversity within populations were calculated for this study. The first was nucleotide diversity. This measures the number of individual nucleotides that differ between individuals in order to determine the degree of differentiation between said individuals. The second measure of diversity was a count of the number of different haplotypes found within each population. Haplotype diversity calculates the probability of two randomly chosen alleles being different, whereas nucleotide diversity is the proportion of different nucleotides per individual. For every population, the number of distinct haplotypes, as well as the number of haplotypes for both the male-only set and the female-only set, was calculated using MEGA.

Calculating Pairwise Genetic Distances

For all pairs of local populations, genetic distance was calculated using four different metrics. Each of these has different strengths and weaknesses, so a secondary goal of this study was to compare the consistency of outcomes across different genetic distance metrics.

1. Mean uncorrected *p*-distance: the proportion (*p*) of nucleotide sites at which two aligned DNA sequences are different, averaged across all pairs of individuals that come from different populations. This is the simplest measure of overall sequence dissimilarity, and for this study, it was calculated in the computer program MEGA.
2. Mean maximum likelihood (ML)-corrected distance: the proportion (*p*) of nucleotide sites at which two aligned DNA sequences are different, after correcting for multiple substitutions at the same site (i.e., "double hits") and substitution rate biases (e.g., transition/transversion ratio that differs from standard ts/tv = 2.0). It was averaged across all pairs of individuals that come from different populations. In this study, the most appropriate nucleotide substitution model for a given individual-based DNA sequence dataset was estimated using the "Find Best DNA/Protein Model" function in MEGA.
3. Wright's (1950) Fixation index (F_{ST}): a measure of how different two or more population gene pools are from one another based on similarity in frequencies of DNA haplotypes, and their variance. Unlike *p*-distance and ML-distance, no weight is given to how many mutational differences exist between two DNA haplotypes. Also, here, the population gene pool is the basic unit of analysis. F_{ST} was calculated using Weir and Cockerham's (1984) estimate of F_{ST}, in ARLEQUIN v.3.5.2.2 (Excoffier & Lischer, 2010).
4. Excoffier et al.'s (1992) Φ_{ST}: a measure of how different two or more population gene pools are from one another based on similarity in frequencies of DNA haplotypes and

their variance, after weighing how many mutational differences exist between two DNA haplotypes, accounting for multiple substitutions at the same site and substitution rate biases. The most appropriate nucleotide substitution model for a given haplotype-based DNA sequence dataset was estimated using the "Find Best DNA/Protein Model" function in MEGA, and Φ_{ST} was then calculated in ARLEQUIN.

Testing the Correlation between Genetic and Geographic Distance

Using the different levels of organization and the geographic distance data from the original collection data, the computer program GenAlEx v.6.5 (Peakall & Smouse, 2012) was used to calculate the correlation with a Mantel (1967) test. Before entering the data into the program, it had to be gathered and reorganized using a simple text editor. With the geographic distance between populations on one axis and the ML-distance among the haplotypes or the Φ_{ST} values from the ML-distance among the populations on the second axis, the data were organized into a triangular matrix that was then sent through GenAlEx to get the final statistical data on how distance physically between populations of SPB is related to genetic distance.

Assessing the Antiquity of Genetic Structure Among All Populations

To understand whether genetic differentiation among the sampled populations had arisen over relatively short versus longer timescales, two of these genetic distances were directly compared: F_{ST} (which can change quickly due to genetic drift alone) and Φ_{ST} (which changes more slowly, over timescales during which new mutations arise and accumulate within gene pools). This F_{ST} versus Φ_{ST} comparison was made for the full data set, males only, and females only.

Results

Population Sampling and Genetic Data Acquisition

Sampled from 17 different sites across the Eastern United States, from Florida to Louisiana to Connecticut, this study included 382 individual SPB with a breakdown of 278 male specimens and 104 female specimens. The male beetles were taken from all 17 sites, while the female beetles were collected from only 16 sites, excluding the Holly Springs, Mississippi collection site (Table 1). With these samples, 118 distinct haplotypes of the COI gene were identified. It was found that most polymorphic sites in the COI gene were in third positions relative to the codon reading frame, which is consistent with expectations for protein coding genes, as third position polymorphisms are far less likely to generate functional mutations during translation. It is important to note that although the samples used in this study came from two separate sources, there was no noticeable difference between the new data collected by Dr. Isis Arantes and the previously analyzed data taken from Dr. Nathan Havill's work.

Table 1: A general breakdown of the raw data, including sample sites, abbreviations used for each, the ratio of males to females within each population, the number of haplotypes, and a measure for nucleotide diversity. In a few populations, the nucleotide diversity value for only females is labeled "n/a", this is because there are either no female samples from that population, or only one, rendering nucleotide diversity data irrelevant for those female populations.

Locality	**State**	**Alias**	**No. of individuals (m; f)**	**No. of COI haplotypes (m; f)**	**Nucleotide diversity (m; f)**
Talladega	Alabama	AL	26 (18; 8)	15 (11; 6)	0.0032 (0.0029; 0.0034)
Hopeville Pond State Park	Connecticut	CT	18 (17; 1)	12 (11; 1)	0.0037 (0.0036; n/a)
LaCrosse	Florida	FL_LC	2 (1; 1)	2 (1; 1)	0.0015 (n/a; n/a)
Ponte Vedra	Florida	FL_PV	23 (17; 6)	7 (4; 6)	0.0026 (0.0023; 0.0036)
Warwick	Georgia	GA	24 (13; 11)	7 (5; 5)	0.0026 (0.0031; 0.0020)
Sicily Island	Louisiana	LA	27 (20; 7)	18 (15; 6)	0.0053 (0.0052; 0.0055)
Woolford	Maryland	MD	30 (25; 5)	8 (7; 4)	0.0027 (0.0025; 0.0040)
Bienville	Mississippi	MS_Bie	20 (11; 9)	17 (11; 8)	0.0048 (0.0051; 0.0046)
Holly Springs	Mississippi	MS_Hol	23 (23; 0)	12 (12; 0)	0.0042 (0.0042; n/a)
Homochitto	Mississippi	MS_Hom	28 (13; 15)	17 (9; 11)	0.0046 (0.0053; 0.0040)
Tombigbee	Mississippi	MS_Tom	28 (24; 4)	8 (7; 4)	0.0036 (0.0030; 0.0079)
Pisgah	North Carolina	NC	19 (18; 1)	17 (16; 1)	0.0059 (0.0061; n/a)

Sag Harbor	New York	NY_Sag	15 (8; 7)	9 (4; 7)	0.0026 (0.0023; 0.0032)
Schunemunk Mountain State Park	New York	NY_Sch	21 (16; 5)	16 (13; 5)	0.0043 (0.0046; 0.0036)
Goat Hill	Pennsylvania	PA	41 (30; 11)	14 (10; 8)	0.0035 (0.0034; 0.0037)
Arcadia Management Area	Rhode Island	RI	20 (13; 7)	12 (8; 5)	0.0035 (0.0029; 0.0046)
Pocahontas State Park	Virginia	VA	17 (11; 6)	14 (8; 6)	0.0044 (0.0030; 0.0071)
All populations combined	-	-	382 (278; 104)	118 (88; 54)	-

Characterizing Levels of Diversity and Calculating Pairwise Genetic Distances

There was no discernable trend in the difference in male versus female nucleotide diversity values (Table 1). Model selection analyses run both on all individual beetles and the set of distinct haplotypes identified the Tamura 3-parameter model (Tamura, 1992) as the best fit for the data. This model was then used when calculating all ML-distance and Φ_{ST} distance values (Table 1).

Testing the Correlation between Genetic and Geographic Distance

The IBD test was run 12 different times, for each of the four measurements of genetic distance (*p*-distance, ML-distance, F_{ST}, and Φ_{ST}) and each of the three sets (the set of all individuals, the all male set, and the all female set). Only one result was statistically significant, which we recognized by a *P*-value less than 0.05 (Table 2). This result was calculated from the full set and the F_{ST} measure of genetic distance, and in fact the *P*-value for this result was highly

significant, at 0.005 (Table 2). Surprisingly, the slope of this relationship was negative, insinuating that the relationship between genetic distance and geographic distance is negatively correlated in a statistically significant way (Figure 4).

Table 2: P-value, slope and R^2 values for each data set and distance measure, including mean uncorrected p distance (p-distance), maximum likelihood-corrected distance (ML-distance), F statistics (F_{ST}), and Φ_{ST}. The bolded and italicized result is the only statistically significant result.

Data Set and Distance Measure	***P*-value**	**Slope**	**R^2**
Full Set, P-distance	0.365	-5E-08	0.0004
Full Set, ML-distance	0.260	-7E-08	0.0008
Full Set, F_{ST}	***0.005***	***-1E-05***	***0.0241***
Full Set, Φ_{ST}	0.078	-2E-05	0.0108
Male Set, P-distance	0.988	4E-07	0.0270
Male Set, ML-distance	0.979	4E-07	0.0279
Male Set, F_{ST}	0.282	-5E-06	0.0023
Male Set, Φ_{ST}	0.186	-1E-05	0.0045
Female Set, P-distance	0.935	4E-07	0.0162
Female Set, ML-distance	0.939	4E-07	0.0150
Female Set, F_{ST}	0.354	7E-06	0.0007
Female Set, Φ_{ST}	0.716	1E-05	0.0023

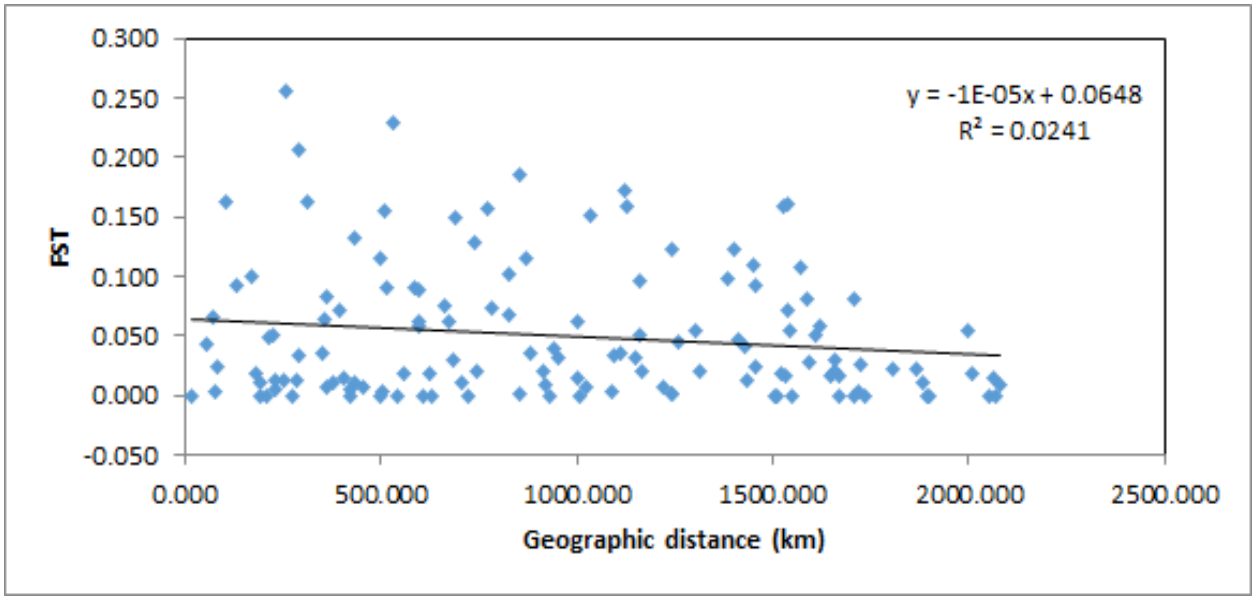

Figure 4: An example isolation by distance plot. One of these was developed for each measure of genetic distance for each partition. This particular plot depicts the statistically significant F_{ST} analysis for the full set of beetles. The slope of the line of best fit is slightly negative.

Assessing the Antiquity of Genetic Structure Among all Populations

In all three partitions (full data set, male-only and female-only), it was found that global Φ_{ST} values (i.e., Φ_{ST} values averaged across all pairs of populations) were greater than corresponding global F_{ST} values (Figure 5). Given that no statistical test for a difference between these two genetic distance measures is available, this is a qualitative assessment, but the outcome is consistent with a relatively old origin of the genetic differences among populations.

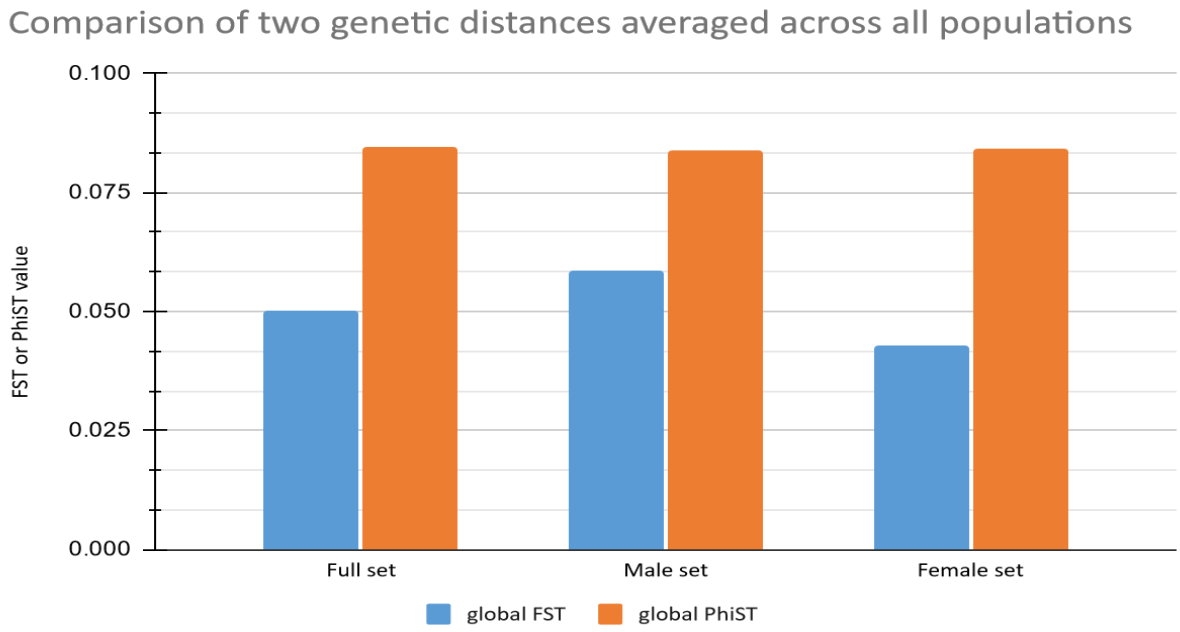

Figure 5: A bar graph comparing the global F_{ST} and Φ_{ST} values for each of the three data sets, a full set, male only, and female only.

Discussion

This study aimed to answer the question of how the range expansion of SPB into the northeastern United States occurred. Using mtDNA and four measures of genetic distance on three partitions of beetle samples, our goal was to determine whether the evidence pointed towards a rapid and recent range expansion of SPB, or an older range expansion, meaning that the observed northern beetles were from previously unnoticed but established populations. From the data collected and analyzed within this study, there are three important results that need to be discussed: 1) the unexpectedly old population structure in SPB suggested by F_{ST} versus Φ_{ST} analyses, 2) the overall lack of evidence for IBD barring one significant outcome, and 3) the apparent lack of difference in male dispersal compared to female dispersal.

Unexpectedly Enduring Population Structure in SPB

The comparison between the F_{ST} and Φ_{ST} genetic distances averaged across all populations (Figure 5) showed that although the level of differentiation between populations is overall quite weak, it is indicated to have originated a long time ago. This suggests in turn that SPB have a long history of occupation of the areas that were sampled in this study, which is somewhat consistent with the hypothesis that they were always native to said areas but at a low density. However, during this study we did not explicitly assess the relative contribution of northern versus southern populations to this outcome. As Dodds et al. (2018) suggested, it is

possible that SPB originated in the south, in a range from Texas to Florida to New Jersey, with the northern populations from states like New York and Rhode Island originating at a later date. It is possible that if the populations sampled in this study were divided into two or more demes, such as a northern and a southern set. If so, rerunning the F_{ST} versus Φ_{ST} comparison separated on each might yield different results. Identifying exactly which subset of local populations contribute most to the F_{ST} versus Φ_{ST} differences would be a valuable follow-up to this study.

Lack of Evidence for IBD

As seen in Table 2, only one statistically significant IBD test result was found during this study. It is possible that this was a false positive, as one erroneous result out of 12 different IBD tests (the three dataset partitions by the four genetic distance measures) approximately matches with the confidence level of this test (i.e., the *P*-value alpha was set at 0.05, suggesting that approximately one in 20 outcomes may be erroneous). However this result cannot be completely discarded, as it was highly significant, with a *P*-value of 0.005. Overall, the implication of most of the outcomes of the tests, which showed a lack of evidence for IBD, is consistent with the hypothesis of recent rapid range expansion. Notably, this contrasts with the inference of long-standing population structure across the sampled range, based on the F_{ST} vs. Φ_{ST} comparison, described above.

The relationship between a lack of IBD and recent rapid range expansion can be seen in many invasive species across the globe. The codling moth, *Cydia pomonella*, is one such species, first reported in China in the 1950s, followed by an observed rapid expansion of range (Men et al. 2012). Knowing the expansion pattern of the moths allowed Men et al. (2012) to study the genetic diversity of the species to shed light on micro-evolutionary processes related to invasion.

Using eight microsatellite loci, Men et al. (2012) found that *C. pomonella* populations from twelve regions in China, when compared to two external regions from Germany and Switzerland, showed that among the northwestern populations, there was no significant correlation between genetic and geographic distance. There was also no evidence of a bottleneck event that might have otherwise explained away this finding (Men et al. 2012). Like the codling moth, the lack of IBD that was found for SPB could be correlated with a relatively fast invasion of the northeastern United States.

If the significant IBD test result using the full set of SPB and the F_{ST} genetic distance metric is taken at face value as a true indicator of population processes, it is interesting to note that the relationship observed between F_{ST} values and geographic distance was negative, such that genetic distance decreases as geographic distance increases. This relationship could be explained by the unintentional human influence on the range expansion of SPB (e.g., repeated long-distance translocations from a common source population), as was concluded after observing a similar slightly negative correlation between genetic and geographic distance in coqui frogs in Hawaii (Everman & Klawinski, 2013). This explanation is plausible for SPB, as timber shipments could inadvertently spread SPB from region to region. Another possible explanation for the negative relationship, if it is considered a biologically meaningful result, could be that dispersal of SPB is limited by habitat similarity. Tonnis et al. (2005) provided evidence for the hypothesis that a negative IBD relationship observed in Warbler finches could be attributed to female finches instinctively seeking out habitats in which to breed that closely match their own natal habitat. This could be further assessed by looking for a positive correlation of genetic distances between populations and the similarity between the habitats that those populations occur within (Tonnis et al., 2005). The same would have to be done with the

populations of SPB sampled to determine if the habitat similarity explanation could be true for this species as well.

No Indication of Male/Female Differences in Dispersal

As the only significant IBD test result found by this study was for the full set of SPB sampled, it is simple to conclude that there is no difference between male and female SPB in regards to IBD. However, Garrick et al. (2021) repeatedly found significant IBD results in tests that contained male SPB using nuclear microsatellites for a measure of relatedness, but not in sets that included only female SPB. This result was consistent with broad-scale dispersal of female pioneers.

This dissonance between studies involving the same species could be attributed to the genetic information used in the studies. For this study, we used mtDNA, which is maternally inherited. As such, information gathered from mtDNA data would not show dispersal or IBD differences based on sex as clearly, as all male SPB would have inherited their mtDNA strictly from their mother, while data procured from nuclear DNA microsatellites would be more axiomatic in regard to sex differences. It is also important to note that there was an obvious difference in sample size of male versus female SPB in this study, specifically, over double male SPB in regard to female SPB (278/104, male/female). This also could have impacted the appearance of possible differences between male and female SPB dispersal.

Limitations, Caveats, and Future Directions

Some limitations of this study include the male-biased sampling method that resulted in an imbalance of sex of the SPB collected, and even the complete lack of female specimens from

one sampling site, which could have produced results that are unrepresentative of SPB populations in the United States as a whole, as well as uncertainty about how dividing the subpopulations of SPB into northern and southern demes could affect the overall results in regards to IBD and F_{ST} versus Φ_{ST} comparisons. It may also be beneficial to measure geographic distance in ways other than Euclidean straight-line distance, such as distance as measured using only major roadways, or distance using a path completely avoiding roadways. In the future, designing experiments that address these limitations could elucidate the relationships between SPB subpopulations, and whether the conclusions reached in this study are supported. Other factors that may have contributed to erroneous or unclear results include establishing whether the sampling sites that we used as populations have evidence of improper pooling, such that sets of SPB that we treated as separate populations were incorrectly pooled together or separated from another population that could actually be pooled into one. For future studies, these issues could be rectified using analysis tools that can assess the connectivity between sampling sites to search for true boundaries of subpopulations, before running IBD tests again.

Bibliography

Ballard, J. William, and Michael C. Whitlock. “The Incomplete Natural History of Mitochondria.” *Molecular Ecology*, vol. 13, no. 4, 2004, pp. 729–744., https://doi.org/10.1046/j.1365-294x.2003.02063.x.

Brian Kahn Follow @blkahn. “Climate Change in the U.S. in 8 Compelling Charts.” *Climate Central*, 6 May 2014, https://www.climatecentral.org/news/climate-change-in-8-compelling-charts-17406.

Dodds, Kevin J, et al. “Expansion of Southern Pine Beetle into Northeastern FORESTS: Management and Impact of a Primary Bark Beetle in a New Region.” *Journal of Forestry*, vol. 116, no. 2, 2018, pp. 178–191., doi:10.1093/jofore/fvx009.

Everman, Elizabeth, and Paul Klawinski. “Human-Facilitated Jump Dispersal of a Non-Native Frog Species on Hawai'i Island.” *Journal of Biogeography*, vol. 40, no. 10, 2013, pp. 1961–1970., https://doi.org/10.1111/jbi.12146.

Excoffier, Laurent, et al. “Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data.” *Genetics*, vol. 131, no. 2, 1992, pp. 479–491., https://doi.org/10.1093/genetics/131.2.479.

Excoffier, Laurent, and Heidi E. L. Lischer "Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows." *Molecular Ecology Resources*, vol. 10, no. 3, 2010, pp 564-567., https://doi.org/10.1111/j.1755-0998.2010.02847.x.

Garrick, Ryan C., et al. “Weak Spatial-Genetic Structure in a Native Invasive, the Southern Pine Beetle (Dendroctonus Frontalis), across the Eastern United States.” *PeerJ*, vol. 9, 2021, https://doi.org/10.7717/peerj.11947.

Havill, Nathan P, et al. “New Molecular Tools for Dendroctonus Frontalis (Coleoptera: Curculionidae: Scolytinae) Reveal an East–West Genetic Subdivision of Early Pleistocene Origin.” *Insect Systematics and Diversity*, vol. 3, no. 2, 2019, https://doi.org/10.1093/isd/ixz002.

Jones, Kelsey L., et al. “Factors Influencing Dispersal by Flight in Bark Beetles (Coleoptera: Curculionidae: Scolytinae): From Genes to Landscapes.” *Canadian Journal of Forest Research*, vol. 49, no. 9, 2019, pp. 1024–1041., https://doi.org/10.1139/cjfr-2018-0304.

Kocher, T D, et al. “Dynamics of Mitochondrial DNA Evolution in Animals: Amplification and Sequencing with Conserved Primers.” *Proceedings of the National Academy of Sciences*, vol. 86, no. 16, 1989, pp. 6196–6200., https://doi.org/10.1073/pnas.86.16.6196.

Lucas, Lauren K., et al. “Geographic and Genetic Isolation in Spring-Associated Eurycea Salamanders Endemic to the Edwards Plateau Region of Texas.” *Conservation Genetics*, vol. 10, no. 5, 2008, pp. 1309–1319., https://doi.org/10.1007/s10592-008-9710-2.

Mantel, Nathan. “The detection of disease clustering and a generalized regression approach.” Cancer Res. 1967 Feb;27(2):209-20. PMID: 6018555.

Meirmans, Patrick G. “The Trouble with Isolation by Distance.” *Molecular Ecology*, vol. 21, no. 12, 2012, pp. 2839–2846., https://doi.org/10.1111/j.1365-294x.2012.05578.x.

Men, Qiu-Lei, et al. “Genetic Structure and Diversity of a Newly Invasive Species, the Codling Moth, Cydia Pomonella (L.) (Lepidoptera: Tortricidae) in China.” *Biological Invasions*, vol. 15, no. 2, 2012, pp. 447–458., https://doi.org/10.1007/s10530-012-0299-5.

Munro, Holly L., et al. “Through Space and Time: Predicting Numbers of an Eruptive Pine Tree Pest and Its Predator under Changing Climate Conditions.” *Forest Ecology and Management*, vol. 483, 2021, p. 118770., https://doi.org/10.1016/j.foreco.2020.118770.

Peakall, R., and P. E. Smouse. “GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research--an Update.” *Bioinformatics*, vol. 28, no. 19, 2012, pp. 2537–2539., https://doi.org/10.1093/bioinformatics/bts460.

Putman, Alexander I., and Ignazio Carbone. “Challenges in Analysis and Interpretation of Microsatellite Data for Population Genetic Studies.” *Ecology and Evolution*, vol. 4, no. 22, 2014, pp. 4399–4428., https://doi.org/10.1002/ece3.1305.

Slatkin, M. “Gene Flow in Natural Populations.” *Annual Review of Ecology and Systematics*, vol. 16, no. 1, 1985, pp. 393–430., https://doi.org/10.1146/annurev.es.16.110185.002141.

Tamura, Koichiro, et al. “Mega6: Molecular Evolutionary Genetics Analysis Version 6.0.” *Molecular Biology and Evolution*, vol. 30, no. 12, 2013, pp. 2725–2729., https://doi.org/10.1093/molbev/mst197.

Tamura, Koichiro. “Estimation of the Number of Nucleotide Substitutions When There Are Strong Transition-Transversion and G+C-Content Biases.” *Molecular Biology and Evolution*, 1992, https://doi.org/10.1093/oxfordjournals.molbev.a040752.

Tonnis, Brandon, et al. “Habitat Selection and Ecological Speciation in Galápagos Warbler Finches (Certhidea Olivacea and Certhidea Fusca).” *Proceedings of the Royal Society B: Biological Sciences*, vol. 272, no. 1565, 2005, pp. 819–826., https://doi.org/10.1098/rspb.2004.3030.

Weir, B. S., and C. Clark Cockerham. “Estimating F-Statistics for the Analysis of Population Structure.” *Evolution*, vol. 38, no. 6, 1984, p. 1358., https://doi.org/10.2307/2408641.

Williams, Jennifer L., et al. “How Evolution Modifies the Variability of Range Expansion.” *Trends in Ecology & Evolution*, vol. 34, no. 10, 2019, pp. 903–913., https://doi.org/10.1016/j.tree.2019.05.012.

Wright, Sewall. “Genetical Structure of Populations.” *Nature*, vol. 166, no. 4215, 1950, pp. 247–249., https://doi.org/10.1038/166247a0.

Zinovkina, L. A. “Mechanisms of Mitochondrial DNA Repair in Mammals.” *Biochemistry (Moscow)*, vol. 83, no. 3, 2018, pp. 233–249., https://doi.org/10.1134/s0006297918030045.