Spring 5-9-2020

Population Genetics of a Recent Range Expansion by the Southern Pine Beetle, Dendroctonus Frontalis, into the Northeastern United States

Megan Stubbs
University of Mississippi

Population genetics of a recent range expansion by the southern pine beetle, *Dendroctonus frontalis*, into the Northeastern United States

by Megan Stubbs

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 May 2020

Approved By

Advisor: Dr. Ryan Garrick

Reader: Dr. Rebecca Symula

Reader: Dr. Sarah Liljegren

ACKNOWLEDGEMENTS

I would like to extend my greatest gratitude to Dr. Ryan Garrick for giving me the opportunity join his lab and take on my own research project with the southern pine beetle. He has been a memorable mentor over the past one and a half years and opened my eyes to the different perspectives of biology and of life. I credit my desire to pursue biological research as a career to him and the impression he has had on me during my time in his lab. I would also like to thank Dr. Nathan Havill for providing DNA samples and microsatellite primers to test, and Dr. Ísis Arantes for scoring genotypes and providing data from six additional loci to include in my analyses. Both of their contributions were pertinent to completing this study. Lastly, I would like to thank all members of the Garrick lab for their meaningful input and support, and for always celebrating my successes.

ABSTRACT

MEGAN STUBBS: Population genetics of a recent range expansion by the southern pine beetle, *Dendroctonus frontalis*, into the Northeastern United States (Under the direction of Dr. Ryan Garrick)

Population genetics as a field of study aims to determine the genetic variation among individuals in a population, and differences among populations. Certain population genetic analyses can provide such information and be used to better understand the biological aspects to a species' expansion beyond its native range. The southern pine beetle (SPB), *Dendroctonus frontalis*, has become an invasive pest to pine forests in northeastern United States with its recent range expansion. Nine microsatellite loci were first developed and then used in analyses. To determine what the genetic variation is among individuals in SPB populations across its entire range, including ones at the leading edge, Hardy Weinberg Equilibrium (HWE) tests were performed. To assess the differences among populations, Fst and "exact" tests were performed. It was found that of 27 geographical populations, 6 were inconsistent with HWE and majority with significant ($P < 0.05$) F_{IS} values indicative of inbreeding. The study also found 4 populations with significant pairwise F_{ST} values t ($P < 0.05$, based on exact tests) indicating that these populations are genetically different and isolated from most others. The analyses performed in this study can be expanded with additional SPB samples and in conjunction with other SPB microsatellite loci. These findings can be used to better understand the biological aspects to SPB recent range expansion and further applied to species management strategies.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

Population genetics and microsatellite markers

Population genetics is a subfield of evolutionary biology that focuses on the genetic variation among individuals within a population, and differences among populations (Aitken et al., 2012). This field of study considers a multitude of biological processes that influence the genetic composition of populations such as the type of mating that typically occurs (e.g., random, inbreeding, or extreme outbreeding), natural selection, and genetic drift. Research in population genetics can also provide information about many other aspects of the biology of a species that are difficult to directly observe, such as how individuals disperse across the landscape. As such, this type of research enables connections to macroevolution by a better understanding of processes that lead to change over time below the species level.

Nuclear microsatellite markers are repetitive sections of DNA that vary in length. In general, microsatellite genetic data are like DNA fingerprints that tell us something unique about each individual, and this can help to draw conclusions about dispersal into new areas. Notably, microsatellite DNA regions have fast mutation rates and therefore have the potential to reflect recent and on-going dispersal events. Because of their rapid evolution, as Barker (2002) described, these markers are a useful tool for understanding the population genetic consequences of recent range expansions.

Southern pine beetle range expansion

In the present work, population genetic analyses of microsatellite markers were used to study a beetle that is native to the southeastern United States but has recently been moving into new areas in the northeast (Havill et al., 2019). The southern pine beetle (SPB), *Dendroctonus frontalis*, is a major native pest of pine trees in the southeastern United States, Mexico, and parts of Central America. Invasive species cause economic and ecological damage, particularly when they compete with or attack native species (Dodds et al., 2018). SPB has become invasive owing to expansion beyond its natural geographic range, particularly into the northeastern United States. Given that SPB is one of the most destructive pests of native pine trees, there is concern about the consequences of the recent and rapid expansion from their native regions to northeastern states such as New Jersey, New York, Connecticut and Massachusetts (Havill et al., 2019). Some of these states harbor ecologically unique pine barrens habitats, which are unlikely to have been exposed to SPB before, and so there is concern that these regions may be particularly susceptible to attack. Consequences of uncontrolled SPB range expansion and invasion include not only the infestation of pine forests, but also subsequent habitat loss (Lesk et al., 2017). Population genetics methods can help determine genetic relationships within and among populations of these beetles by developing genetic markers and incorporating population-level analyses of the resulting genotype data. Improving our understanding of the population-level range expansion of SPB may help explain how this beetle species has reached its northeastern frontiers.

Goals

The goals of this project were to (1) develop new microsatellite markers for SPB, (2) determine the most common type of mating among individuals within populations (i.e., random vs. non-random), (3) assess the level of genetic differences (i.e., genetically isolated or not isolated) among populations across the species' current range (i.e., native, plus recent expansion), and (4) based on these measures, draw some conclusions about the range expansion of this species. By knowing how and to what degree these populations are genetically connected to one another, it may become possible to predict and manage for further expansion. Outcomes from this project will be shared with collaborators at USDA Forest Service, to help develop strategies for effective management of SPB and conservation of biodiversity that may be under threat from this species.

METHODS

Sampling, DNA extraction, and development of new microsatellite markers

Beetle sampling was performed as described by Havill et al. (2019). Briefly, USDA Forest Service personnel used funnel traps with pheromone baits to catch SPB individuals, and then preserved specimens in ethanol. Following sampling, procedures for handling beetles and extracting genomic DNA from abdomen and leg tissue were performed as detailed in Havill et al. (2019). Descriptions of locations of sampling sites plus the number of individuals sampled in each site are listed in Table 1, and the geographic distribution of the 27 sampling sites is shown In Figure 1. Microsatellite regions that had previously been identified from genome sequencing data were selected for further assessment of their use as population genetic markers. Specifically, this study focused on 15 previously untested primer pairs, each intended to amplify a single unique locus. These were assessed using a test panel of 7-8 SPB individuals that were representative of the geographic sampling. Informative loci were identified as those that amplified reliably and varied in length. Then, informative loci and were subsequently used for large-scale population sampling. Microsatellite loci that were identified as informative were subsequently used for large-scale population screening.

Table 1: Beetle sampling sites locations, and number of individuals collected per site.

Sampling site	**State; region**	**Longitude**	**Latitude**	**No. of individuals**
Site_01	Florida; Ponte Vedra	-81.388256	30.236439	24
Site_02	Florida; LaCrosse	-82.399920	29.839006	2
Site_03	Arizona; Flagstaff	-111.651300	35.198300	7

Site_04	Alabama; Talladega	-86.984001	32.774676	30
Site_05	Mississippi; Bienville	-89.313320	32.132130	20
Site_06	Mississippi; Homochitto	-91.152320	31.381830	29
Site_07	Pennsylvania; Goat Hill	-76.076320	39.726630	31
Site_08	Pennsylvania; Susquehannock	-76.277927	39.807098	3
Site_09	Pennsylvania; Codorus	-76.883951	39.798588	2
Site_10	Pennsylvania; French Creek	-75.807750	40.213980	6
Site_11	New York; Sag Harbor	-72.264500	40.994400	27
Site_12	Maryland; Woolford	-76.215650	38.531270	30
Site_13	Louisiana; Sicily	-91.739810	31.864900	27
Site_14	Mississippi; Tombigbee	-88.931820	34.037570	28
Site_15	Mississippi; Holly Springs	-89.349660	34.412950	26
Site_16	Connecticut; Naugatuck	-72.949100	41.448750	1
Site_17	Connecticut; Wharton Brook	-72.833900	41.426760	1
Site_18	Connecticut; Hopeville Pond	-71.925600	41.608400	10
Site_19	Connecticut; Tilcon	-72.017023	41.377748	5
Site_20	Connecticut; Oswegatchie	-72.196003	41.335621	2
Site_21	Georgia; Warwick	-83.889400	31.771300	30
Site_22	Mexico; Michoacan	-102.156116	19.443702	6
Site_23	New York; Minnewaska	-74.217826	41.738276	2
Site_24	New York; Bear Mountain	-73.991845	41.300368	2
Site_25	Rhode Island; Arcadia	-71.697008	41.598270	31
Site_26	Illinois; Shawnee	-88.668891	37.439150	1

Site_27	North Carolina; Pisgah	-82.642380	35.519440	25

Figure 1: Map showing locations of 27 sampling sites from which SPB were collected.

Goal 1: Microsatellite marker development and screening for genetic variation

To develop new microsatellite markers for SPB, I performed a three step process that 1) identified primer pairs that reliably amplified a product of the expected size and that showed length variation in a test panel of nine SPB, 2) created combinations of primer pairs that could be combined into a single, "multiplexed" polymerase chain reaction (PCR), and 3) screened samples from several populations in the geographic range of SPB.

For the first step of the process, microsatellite loci were amplified from beetle DNA via polymerase chain reaction (PCR). For each locus, a forward and reverse primer was used to amplify the target DNA region. Many reactions were conducted in which PCR conditions were

varied to determine which of the 15 potentially useful primer pairs reliably amplified a product of the expected size and showed length variation. Specifically, we varied reagents within the reaction and thermal cycler conditions.

For the second step of the process, additional PCR trials were performed with the promising loci identified in step one to determine which loci could be combined in groups of three (i.e., in multiplex PCRs), without compromising successful amplification of each locus. During these trials, outcomes from three different sets of PCR reagents (i.e., master mixes) were also compared to determine which produced the best outcome. Based on the test panel of beetles, nine informative microsatellite loci were identified, and these could be amplified in three multiplex PCRs (three loci in each; see Results).

In the third step of the process, I amplified three microsatellite loci from 96 samples in one multiplex PCR from a total of 401 SPB samples. The names of loci included in this particular multiplex PCR, and their respective dyes and sequence lengths in base pairs (bp), are listed in Table 2. For each multiplex PCR amplification, individual beetle DNA samples were added to separate 0.2 mL PCR tubes with a master mix containing all ingredients needed for PCR. Components of the master mix used for this multiplex are given in Table 3. After adding the master mix to diluted beetle DNA, the PCR reagents were mixed by vortexing and then briefly centrifuged. Thermal cycling conditions used for the amplification of the three microsatellite loci in a single multiplex PCR are listed in Table 4.

Table 2: Multiplex polymerase chain reaction, with locus-specific fluorescent dye at the 5’ end of the forward primer of each locus, and product length.

Locus name	**Fluorescent dye**	**Length (bp)**
SPB265317	FAM	398-404

SPB4155	PET	112-121
SPB979494	NED	208-229

Table 3: Standard polymerase chain reaction Master Mix components, and volumes, per 15 µL reaction.

Component (concentration)	**Volume (µL)**
dH_2O	4.5
Qiagen Type-It microsatellite PCR mix	7.5
Forward-primer (10 µM)	0.75
Reverse-primer (10 µM)	0.75
DNA (1:19 dilution with dH_2O)	1.5

Table 4: Polymerase chain reaction thermal cycling conditions for a "touchdown" amplification profile.

Step in cycle	**Temperature (°C)**	**Duration**	**Number of Cycles**
Denaturation	95	2 min	1
Denaturation	95	45 sec	5
Annealing	61	-2°C/cycle 30 sec	
Extension	72	45 sec	
Denaturation	95	45 sec	30
Annealing	51	30 sec	
Extension	72	45 sec	
Final extension	60	30 min	1
Hold	12	Indefinite	

Agarose gel electrophoresis

After PCR, agarose gel electrophoresis was used to estimate the length of the amplified products via comparison to a DNA ladder. The approximate lengths were used to confirm successful PCR amplification of each targeted microsatellite region, and also to confirm there was no contamination. Gels were poured and set using 1.5% agarose and 1.5 µL of Gel Red, into which samples were loaded (see Figure 2). When the gel was set, 4 µL of each PCR sample plus 2 µL of loading dye were loaded into the wells. A 100-bp ladder was added to the far-left well (Lane 1) and a negative control in the last well (Lane 25) for comparison. The gels were run for 90 minutes at 80 volts. After running the gels, dark bands (see Figure 2), which referenced the fragment size of the amplified microsatellite loci, were expected to be approximately 400-bp, 215-bp, or 115-bp long, if all three loci had successfully co-amplified in a single reaction.

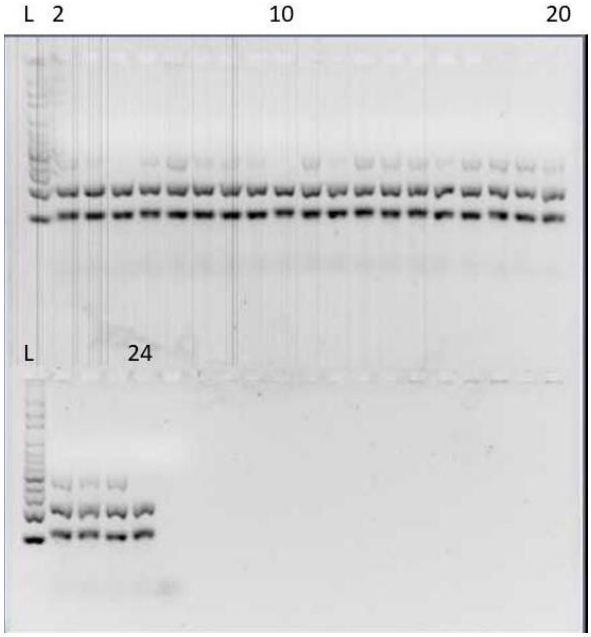

Figure 2: Agarose gel showing multiplex amplification of three microsatellite loci from a set of 22 SPB samples. Lane 1 contains a 100-bp ladder. PCR products amplificated from beetle DNA were loaded in Lanes 2-24. The approximate sizes of dark bands are expected to be 400-bp (locus SPB265317), 215-bp (SPB979494), and 115-bp (SPB4155), from top to bottom. Lane 25 was a negative control (i.e., a PCR that contained no DNA) and indicated no contamination.

Fragment analysis via high-resolution electrophoresis

PCR products were sent to Yale University's DNA Analysis Facility on Science Hill for fragment analysis using an Applied Biosystems 3730xl 96-capillary electrophoresis genetic analyzer. Briefly, this platform separates fragments by size, and runs each sample together with a Rox-500 DNA ladder to enable size estimates to the nearest 1-bp. The software Geneious v.6.1.8 (Kearse et al., 2012) was used to edit and make final determinations of allele lengths at each locus, for each individual beetle, and this work was done by Dr. Ísis Arantes to insure consistent fragment size identification. The data from three microsatellite loci described above were combined with data generated by generated by Dr. Ísis Arantes for the other six new microsatellite loci. The final data set used for analysis consisted of nine loci and 401 individuals.

Goals 2 and 3: Data analysis

Following successful development of new microsatellite loci (Goal 1, above), the resulting data were analyzed to examine the type of mating that occurred within populations (Goal 2), and to assess levels of genetic differentiation among populations (Goal 3). After first assigning individuals into populations based on the geography of sampling 27 different locations show in Figure 1 (i.e., beetles collected from the same place were considered to be members of the same population), the software GenePop 4.7.5 (Rousset, 2017) was used to conduct analyses regarding (a) Hardy-Weinberg Equilibrium (HWE) to assess departures from the null hypothesis of random mating within populations (Goal 2), and (b) Fst to measure the amount of genetic

differentiation among populations (Goal 3). Briefly, whenever a population showed significant ($P < 0.05$) departure from HWE based on all nine loci analyzed together, an attempt was made to determine if departure from HWE was due to inbreeding vs. extreme outbreeding by looking at F_{IS} values for each locus separately (where significantly positive F_{IS} indicates inbreeding, and significantly negative F_{IS} indicates extreme outbreeding). Also, to help interpret the importance of F_{ST} values calculated for all possible pairs of populations (which range from 0 to 1, where 1 indicates the maximum genetic differentiation), we also assessed their significance ($P < 0.05$) using the "exact tests" option in GenePop.

RESULTS

Goal 1: Development of new microsatellite loci

Of the 15 potentially useful microsatellite loci assessed with a test panel of SPB, after a lot of trial and error using different PCR reagents (e.g., 5x buffer, $MgCl_2$, dNTPs, BSA, taq polymerase), nine were found to amplify reliably, and were variable. The Qiagen Type-it microsatellite PCR mix was always the most successful Master Mix. The loci that were discarded either did not amplify reliably or showed very low fragment size variation. The loci that were retained were combined into multiplex PCRs comprised of three loci each. Within each multiplex, the three loci had non-overlapping sizes, and also had different fluorescent dyes. Overall, this made it easy to score each locus, and saved time and reagents.

Goal 2: Understanding the most prevalent type of mating within populations

Of the 27 geographical populations, Populations 6, 7, 11, 12, 14, and 15 showed significant ($P < 0.05$) deviation from random mating, as measured by multi-locus tests of HWE (Table 5). For each of these six non-randomly mating populations (i.e., those significantly out of HWE), the deviations from random mating generally appeared to be due to inbreeding. For each of these populations, the number of loci with significantly positive F_{IS} (consistent with inbreeding) outnumbered those with significantly negative F_{IS} (outbreeding; see Table 5). There was one exception: for population 15, it was difficult to interpret why it was out of HWE, since none of the loci individually showed significantly positive or negative F_{IS}. Interestingly, there

was no obvious geographic pattern in terms of where the inbreeding populations were found, as they spanned the full current geographic range of SPB (e.g., populations from Mississippi in the southeastern native range, plus populations in Maryland, Pennsylvania and New York in the northeastern invaded range).

Table 5: Assessment of departures from the null hypotheses of random mating within populations, using tests of Hardy Weinberg Equilibrium (HWE) in GenePop. Populations that are significantly out of HWE have a bolded *P*-value. The boxes containing dashes represent the populations for which it was not possible to run the tests.

Local population	Multilocus HWE test *P*-value (null: random mating)	No. of loci with significantly positive F_{IS} (inbreeding)	No. of loci with significantly negative F_{IS} (outbreeding)
Pop 1 - Florida; Ponte Vedra	0.2636	1	0
Pop 2 - Florida; LaCrosse	-	-	-
Pop 3 - Arizona; Flagstaff	0.0889	1	0
Pop 4 - Alabama; Talladega	0.1107	3	0
Pop 5 - Mississippi; Bienville	0.226	1	0
Pop 6 - Mississippi; Homochitto	< **0.0001**	2	0
Pop 7 - Pennsylvania; Goat Hill	**0.0022**	3	1
Pop 8 - Pennsylvania; Susquehannock	0.9422	0	0
Pop 9 - Pennsylvania; Codorus	1	0	0
Pop 10 - Pennsylvania; French Creek	0.3505	0	0
Pop 11 - New York; Sag Harbor	**0.0019**	3	0
Pop 12 - Maryland; Woolford	**0.0406**	2	0
Pop 13 - Louisiana; Sicily	0.4242	0	0
Pop 14 - Mississippi; Tombigbee	**0.0193**	3	0
Pop 15 - Mississippi; Holly Springs	**0.0471**	0	0

Pop 16 - Connecticut; Naugatuck	-	-	-
Pop 17 - Connecticut; Wharton Brook	-	-	-
Pop 18 - Connecticut; Hopeville Pond	0.8722	0	1
Pop 19 - Connecticut; Tilcon	0.9526	0	0
Pop 20 - Connecticut; Oswegatchie	0.9277	0	0
Pop 21 - Georgia; Warwick	0.0832	1	0
Pop 22 - Mexico; Michoacan	0.6926	0	0
Pop 23 - New York; Minnewaska	0.9751	0	0
Pop 24 - New York; Bear Mountain	-	-	-
Pop 25 - Rhode Island; Arcadia	0.8615	1	0
Pop 26 - Illinois; Shawnee	-	-	-
Pop 27 - North Carolina; Pisgah	0.6568	4	0

Goal 3: Levels of genetic differentiation among populations

From the matrix of all possible population pairs shown in Table 6, the resulting F_{ST} values ranged 0.00-0.75, with 106 significant ($P < 0.05$, based on exact tests) pairwise F_{ST} values. The patterns reflected in Table 6 show that most of the large F_{ST} values (colored boxes) and most of the significant F_{ST} values are associated with pairwise comparisons involving any one of Populations 3, 5, 17 and 22. This suggests that these four populations are genetically different and isolated from most others. Populations 3, 5, 17, and 22 are located in Flagstaff, Arizona; Bienville, Mississippi; Wharton Brook, Connecticut; and Michoacan, Mexico (Table 1; Figure 1).

Table 6: Assessment of genetic differences among populations, using F_{ST} and "exact tests" in GenePop. All possible pairs of populations are represented as a matrix. The diagonal represents comparison of each population to itself, and so no values are reported. The off-diagonal (lower left) contains all of the pairwise F_{ST} values. Values that are significantly larger than zero ($P < 0.05$, based on exact tests) are in bold italics and underlined. To help visually identify patterns, color coding was used to indicate different ranges of F_{ST} values as a "heat map", as follows: dark red: $F_{ST} > 0.5$; red, $F_{ST} > 0.4$; orange: $F_{ST} > 0.3$; yellow: $F_{ST} > 0.2$; blue: $F_{ST} > 0.1$.

Pop	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1																											
2	0.00																										
3	***0.26***	***0.16***																									
4	***0.02***	0.00	***0.26***																								
5	***0.11***	***0.08***	***0.30***	***0.05***																							
6	0.00	0.00	***0.27***	0.01	***0.08***																						
7	0.00	0.00	***0.30***	***0.02***	***0.10***	0.00																					
8	0.02	0.00	***0.21***	0.06	***0.13***	0.02	0.01																				
9	0.04	0.00	***0.13***	0.10	***0.14***	0.01	0.02	0.02																			
10	0.01	0.02	***0.29***	0.04	***0.10***	0.00	0.00	0.00	0.00																		
11	0.00	0.00	***0.27***	***0.01***	***0.11***	0.00	0.00	0.00	0.02	0.01																	
12	***0.02***	0.00	***0.31***	***0.02***	***0.12***	0.01	***0.01***	0.01	0.10	0.03	0.00																
13	0.00	0.00	***0.30***	***0.03***	***0.11***	0.00	0.00	0.05	0.06	0.00	***0.01***	***0.03***															
14	0.00	0.00	***0.25***	***0.02***	***0.09***	0.00	0.00	0.00	0.00	0.00	0.00	***0.01***	***0.00***														
15	0.01	0.00	***0.26***	***0.02***	***0.10***	0.00	0.01	0.07	0.09	0.03	***0.01***	***0.02***	***0.01***	0.01													
16	0.00	0.06	0.08	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00												
17	0.12	0.24	0.11	0.10	0.19	0.05	0.11	0.20	0.15	0.21	0.13	0.15	0.11	0.08	0.09	0.61											
18	0.00	0.03	***0.30***	0.01	***0.11***	0.00	0.00	0.09	0.10	0.00	0.00	0.03	0.00	0.00	0.00	0.04	0.22										
19	***0.07***	0.05	***0.27***	***0.06***	***0.11***	0.01	0.00	0.08	0.04	0.02	0.03	***0.07***	0.05	***0.03***	0.03	0.00	0.21	0.04									
20	0.00	0.00	***0.11***	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00								
21	0.00	0.00	***0.28***	***0.01***	***0.11***	0.00	0.00	0.02	0.06	0.01	***0.00***	0.00	0.00	***0.00***	0.01	0.00	0.06	0.00	***0.07***	0.00							
22	***0.37***	***0.24***	***0.08***	***0.36***	***0.39***	***0.37***	***0.41***	***0.37***	***0.22***	***0.41***	***0.37***	***0.44***	***0.40***	***0.34***	***0.35***	***0.21***	***0.30***	***0.40***	***0.37***	***0.24***	***0.41***						
23	0.00	0.00	***0.17***	0.01	***0.08***	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.03	0.00	0.00	0.01	0.01	0.00	0.00	***0.31***					
24	0.04	0.06	0.23	0.00	***0.13***	0.00	0.00	0.15	0.08	0.05	0.00	0.03	0.03	0.02	0.00	0.08	0.38	0.04	0.00	0.00	0.03	***0.33***	0.12				
25	***0.02***	0.00	***0.28***	***0.02***	***0.11***	0.00	0.01	0.06	0.09	***0.06***	0.00	***0.01***	***0.03***	***0.01***	0.01	0.00	0.13	0.01	0.03	0.00	0.01	***0.39***	0.02	0.00			
26	0.01	0.15	0.17	0.00	0.07	0.00	0.00	0.22	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.45	0.75	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00		
27	***0.04***	0.00	***0.21***	***0.04***	***0.12***	***0.03***	***0.03***	0.02	0.02	0.03	***0.04***	***0.05***	***0.05***	***0.03***	***0.05***	0.00	0.01	***0.04***	***0.04***	0.00	***0.04***	***0.35***	0.00	0.03	***0.04***	0.00	

DISCUSSION

The four goals addressed in this study were to (1) develop new microsatellite markers for SPB, (2) determine the most common type of mating among individuals within populations (i.e., random vs. non-random), (3) assess the level of genetic differences among populations across the species' current range (i.e., native, plus recent expansion), and (4) based on these measures, draw some conclusions about the range expansion of this species. It was determined that nine out of 15 microsatellite markers were most reliable for amplification, and they showed genetic variability. Of the 27 geographic populations sampled, six showed significant ($P < 0.05$) deviation from random mating. Most notably, though, was the significant deviation from random mating for Population 15 in Holly Springs, MS, whose multi-locus test is inconsistent with HWE, yet none of the single locus tests are significant. This population may truly be non-randomly mating, and the reason is probably due to inbreeding. This rationale was determined because many of the locus-by-locus F_{IS} values for this population were positive (7 out of 8), although not significantly so, and the other populations that were inconsistent with HWE were in the direction of inbreeding. However, this inference about why the Holly Springs population was out of HWE is speculative since the tests do not provide a substantial amount of confidence. Finally, four populations stood out from the rest due to their consistently large and significant pairwise F_{ST} values, indicating that they are genetically isolated from each other, and from the other populations.

The study by Havill et al. (2019) developed 24 loci for SPB from the eastern United States, however, but only 18 were suitable for application to all populations in that study (i.e., including western samples from Arizona, and southwestern samples from Mexico). Nine more loci have been developed in this study and can be contributed toward a total of 33 loci. Application of these new nine loci to the Havill et al. (2019) dataset may find more than 18 suitable for genotyping individuals across the wide range of SPB. As described by Tixier et al. (1997), the greater the number of independent microsatellite loci, the more informative the dataset is in return. In the application to the sampled populations in Havill et al. (2019), the new 33 loci can improve the dataset by taking more loci into consideration when performing analyses like ones from this study—tests of Hardy-Weinberg Equilibrium, and F_{ST} and "exact tests".

There were six inbreeding populations found that had significant deviations from HWE ($P < 0.05$). However, there were no instances of extreme outbreeding found, and the geographical locations of the inbreeding populations were mainly concentrated in two regions. Populations 6, 14, and 15 were all from Mississippi, within the native range. Conversely, Populations 7, 11, and 12 were all from the leading edge of SPB range expansion in northeastern United States. For this study, we did not predict how populations at the leading edge of the range might differ from other SPB populations. However, the results follow the pattern found in other invasive insects—that is, the populations in the invaded ranges tend to be more inbred that in the native range. For example, Vargo and Husseneder (2009) described the colonies of subterranean termites, *Reticulitermes flavipe*s, recently introduced into France, as differing radically from the termites studied in the native range in Southeast United States. They reported that studies of populations in Paris and northwestern France to have only highly inbred colonies. Another

invasive insect example that showed great tendencies for inbreeding in invaded regions is the ant, *Brachyponera chinensis,* described by Eyer et al. (2018). Interestingly, this ant was found to have inbreeding tolerance as a pre-adapted trait for invasion success such that inbreeding in the invaded regions is not a consequence of the founder effect following introduction, but it is due to their mating behavior (Eyer et al., 2018). The reporting about the termites and the ants are examples of similar findings in other insects, however it is not intended to suggest that inbreeding at the leading edge is either common or expected, in large suite of insect species.

The most notable findings for Goal 3 was that only a small handful of populations are genetically isolated from all others. These were Populations 3, 5, 17, and 22. The differences in the F_{ST} values can be due to a couple of potential reasons and Figure 1 helps illustrate these. For instance, the sampling site in Flagstaff, Arizona (Population 3) and in Michoacan, Mexico (Population 22) are both at a greater geographic distance away from each other and the remaining sampling sites. Between Population 3 and 22 it is approximately a 2,400-kilometer distance, and the next closest sampled population to Population 3 is 2,140 km away (Population 13) and to Population 22 is 2,125km away (Population 6). Comparatively, the distance between Population 13 and 6 is only 130 km. The large geographic distances between some sites, particularly Population 3 in Arizona and Population 22 in Mexico, may limit the genetic connectivity to other sites primarily in southern and eastern United States. Another possibility for very large F_{ST} values is that within the sampling range of the four outlier populations, there may be other factors that potentially play a role in affecting genetic connectivity. Specifically, in Bienville, Mississippi (Population 5), the population may appear genetically isolated because it recently became established following human-mediated transport of wood that contained SPB

from distant locations. It is possible that visitors to the Bienville National Forest or commercial forest operations are transporting wood, and inadvertently bringing SPB with them. Besides these explanations, there also resides the potential for small sample sizes to lead to spurious estimates of F_{ST}, which might be the case for Wharton Brook, Connecticut (Population 17), since only one individual was sampled there. With numerous sites (Table 1; Populations 2, 9, 16, 17, 20, 23, 24, and 26) having only 1-2 numbers of individuals sampled, there is reason to believe that resulting significant F_{ST} values are not truly significant.

Some limitations of this study can be directed toward each goal. For Goal 1, the numerous trial and error with different master mixes eventually led to a determination that the Qiagen Type-It microsatellite PCR mix was the most successful. When assessing even more loci, this master mix may be best to be used in the first place to save time and resources. Regarding Goal 2, the sampling sites with very few samples were not ideal, because they limit the ability to detect significant departure from random mating. Thus, resampling these sites and improving the sample size would allow for more confidence in the results HWE tests. Concerning Goal 3, the same solution presented to address limitations for Goal 2 applies here. Resampling sites and improving the currently small sample sizes would decrease the potential for spurious results in determining F_{ST} values. Another consideration is to perform follow-up research on the outlier population in Bienville, Mississippi (Population 5) to gather more information on whether human-mediated translocations of wood and SPB is a plausible reason for explaining the large F_{ST} values associated with it. Lastly, Goal 4 was limited with the geographic distribution and overall and number of sampling sites. This can be improved with more samples from the native range in addition to from the populations at the leading edge of the range expansion.

Furthermore, expanding the geographic coverage overall for SPB can improve the understanding of SPB range expansion.

REFERENCES

Aitken S.N., Luikart G.H., Allendorf F.W. (2012). Conservation and the Genetics of Populations. *Wiley-Blackwell*.

Barker, G.C. (2002). Microsatellite DNA: A Tool for Population Genetic Analysis, *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 96:21-24.

Dodds, K.J., Aoki, C.F., Arango-Velez, A., Cancelliere, J., D'Amato, A.W., DiGirolomo, M.F., Rabaglia, R.J. (2018). Expansion of Southern Pine Beetle into Northeastern Forests: Management and Impact of a Primary Bark Beetle in a New Region. *Journal of Forestry*, 116:178-191.

Eyer, P., Matsurra, K., Vargo, E.L., Kobayashi, K., Yashiro, T., Suehiro, W., Himuro, C., Yokoi, T., Guenard, B., Dunn, R.R., Tsuji, K. (2018). Inbreeding Tolerance as a Pre-adapted Trait for Invasion Success in the Invasive Ant *Brachyponera chinensis*, *Molecular Ecology*, 27:4711-4724.

Havill, N.P., Cognato, A.I., del-Val, E., Rabaglia, R.J., Garrick, R.C. (2019). New Molecular Tools for *Dendroctonus frontalis* (Coleoptera: Curculionidae: Scolytinae) Reveal an East–West Genetic Subdivision of Early Pleistocene Origin. *Insect Systematics and Diversity*, 2:1-14.

Lesk, C., Coffel, E., D'Amato, A.W., Dodds, K., Horton, R. (2017). Threats to North American

Forests from Southern Pine Beetle with Warming Winters. *Nature Climate Change*, 7:713-717.

Rousset, F. (2008). Genepop'007: A Complete Re-implementation of the Genepop Software for Windows and Linux. *Molecular Ecology Resources*, 8:103-106.

Tixier, M.H., Sourdille, P., Röder, M., Leroy, P., Bernard, M. (1997). Detection of Wheat Microsatellites using a Non Radioactive Silver-nitrate Staining Method. *Journal of Genetics and Breeding*, 51:175-177.

Vargo, E.L., Husseneder, C. (2009). Biology of Subterranean Termites: Insights from Molecular Studies of *Reticulitermes* and *Coptotermes*. *Annual Review of Entomology*, 54:379-403.