Population Genetic Analysis of the Wood Turtle from Maine to Virginia

Final Report submitted to the Massachusetts Division of Fisheries and Wildlife, the U.S. Fish and Wildlife Service, Wildlife Management Institute, and the Northeast Association of Fish and Wildlife Agencies

"Population Genetic Analysis of the Wood Turtle from Maine to Virginia" was supported in part by State Wildlife Grant funding awarded through the Northeast Regional Conservation Needs (RCN) Program. The RCN Program joins thirteen northeast states, the District of Columbia, and the U.S. Fish and Wildlife Service in partnership to address landscape-scale, regional wildlife conservation issues. Progress on these regional issues is achieved through combining resources, leveraging funds, and prioritizing conservation actions identified in the State Wildlife Action Plan. See RCNGrants.org for more information. This project was also supported by a Competitive State Wildlife Grant, awarded to Massachusetts Division of Fisheries and Wildlife and its partners, "Conservation Planning and Implementation for the Wood Turtle (*Glyptemys insculpta*) and Associated Riparian Species of Greatest Conservation Need from Maine to Virginia"

Dana Weigel, PhD[1] and Andrew Whiteley, PhD[2]

January 2018

[1]University of Massachusetts Amherst (former address), University of Idaho, Moscow, Idaho (current address); dweigel@uidaho.edu

[2]University of Massachusetts Amherst (former address), University of Montana, Missoula, Montana (current address); Andrew.whiteley@mso.umt.edu

Abstract

In 2015, eight Northeastern States began a cooperative project for Conservation Planning for the Wood Turtle (*Glyptemys insculpta*) under a Competitive State Wildlife Grant (CSWG). This portion of the study uses genetic data to identify genetic diversity across the study area (Maine to Virginia), identify the number of populations in the study area, and determine the success of genetic assignment of individuals to sites of origin. Tissue samples were collected as blood, tail tips, toenails and shell shavings or scutes from 1,895 Wood Turtles. Most tissue samples were collected in 2015 and 2016; however, some collectors submitted tissue samples from tissue archived from previous collections with the earliest collection dated 2005. Tissue samples were genotyped at 16 microsatellite markers for 1,244 individuals. Genetic data were analyzed for genetic diversity (using HP-RARE, GENEPOP and GENALEX), allele frequency exact test (using GENEPOP), genetic clustering (using STRUCTURE), full siblings (using COLONY), and genetic assignment (using GENECLASS). Samples sizes ranged from 5 to 50 individuals (average n=17.4) collected from 62 sites. Unbiased allelic richness ranged from 3.4 to 6.2 (average 5.1), private alleles ranged from 0 to 0.3 (average 0.05), unbiased expected heterozygosity ranged from 0.5 to 0.7 (average = 0.6) and F_{IS} ranged from -0.21 to 0.14 (average =0). F_{ST} ranged from 0 to 0.23 (average 0.07). Allele frequency exact tests identified significant pairwise differences between 91% of the sites. The Bayesian genetic clustering analysis indicated that there are likely 3 to 5 clusters with 4 clusters providing the most optimal clustering pattern in the data set. The major population groups identified were northern ME, Potomac, coastal MA and NJ/NY. Sites in PA and NH showed admixture with the neighboring clusters. The results indicate that clear genetic differences among populations (or subpopulations) are detectable across the study area. The Bayesian clustering analysis indicate that an island stepping-stone model describes the population genetic structure where sites are exchanging individuals with neighboring sites creating a gradation of genetic structure over the study area. Isolation by distance was significant for 2 of 3 clusters tested in Potomac and Maine/NH ($p<0.01$). The northern Maine cluster showed a similar pattern but was not significant for isolation by distance ($p=0.17$). Tests for full sibling families indicated a maximum distance between family members of 50 km. Genetic assignments indicated that 52% of individuals in the data set assigned correctly to the collection site. The genetic assignment was moderately successful with some sites providing relatively high (>75%) correct genetic assignment; however, assignment success using these markers varied across the sites/populations and, at some sites, correct assignment was relatively low (<50%) limiting the application of this method for management and enforcement for Wood Turtles confiscated from illegal harvest.

Introduction

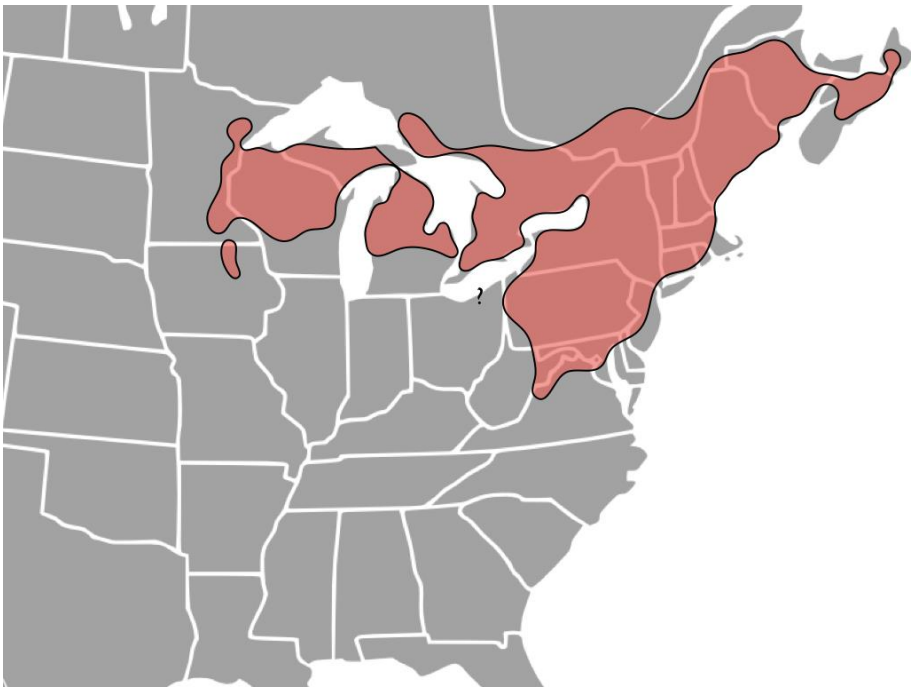
In 2015, eight northeast states began a cooperative project for Conservation Planning for the Wood Turtle (*Glyptemys insculpta*) under a Competitive State Wildlife Grant (CSWG). The lead state agency for this project was Massachusetts Division of Fisheries and Wildlife, and the project was coordinated and managed by the Cooperative Fish and Wildlife Research Unit at the University of Massachusetts-Amherst and the American Turtle Observatory. The participating state agencies included Maine, New Hampshire, Connecticut, Pennsylvania, New Jersey, Maryland and Virginia. Other cooperating entities included the State University of New York (SUNY) Potsdam, the Smithsonian Conservation Biology Institute, and numerous volunteers. The project aimed to conduct standardized surveys of known and representative Wood Turtle populations, survey data-deficient basins, test various sampling methods, describe optimal Wood Turtle habitat, describe the population genetics of Wood Turtle, and where possible, quantify demographic information. The information from this project will be used to develop a cooperative Conservation Plan and long-term implementation framework across the eastern range of this species. In this report, we summarize the population genetic analyses and findings from this study.

Population genetic analyses can be used to support management assessments and conservation planning. Specifically, these analyses can identify genetic diversity, low population size, fragmentation, population structure or designation, gene flow and migration rates – all of which assist the management of populations and associated habitat (Paetkau et al. 2004; Manel et al. 2005). Management units are defined as demographically independent units based on genetic divergence (Pasboll et al. 2007). Understanding the genetic and demographic interactions is important for predicting how populations will respond to environmental and anthropogenic disturbances.

In addition to the identifying populations, genetic data can be used to assign individuals (or parts thereof, such as shells, horns or teeth) to populations of origin (Paetkau et al. 2004; Manel et al. 2005). This method is commonly applied in the illegal animal trade and can be useful to identify illegal poaching activity. In some circumstances, genetic assignment may be used to release confiscated animals to their population of origin (such as Gaur et al. 2006). The ability to identify the site of origin of confiscated animals may assist species' conservation efforts as the threat from illegal harvest continues to increase while the population abundances and habitat quality are continuing to decline.

The Wood Turtle is native to eastern and central North America including the southern and eastern portion of Canada and the northern and eastern portion of the United States (Figure 1). The Wood Turtle was listed as Vulnerable on the International Union for the Conservation of Nature (IUCN) Red List in 1996 and subsequently up-listed in 2011 to Endangered (www.iucn.org). The species is also listed in Appendix II of the Conservation on International Trade in Endangered Species of Wild Fauna and Flora (CITES). The species is listed as threatened under the Species at Risk Act (SARA) in Canada. The species is listed in 13 northeastern states as a Species of Greatest Conservation Need (NEPARC 2010), and currently proposed for listing under the Endangered Species Act in the United States. The species is documented as declining in most areas of the native range. Threats to the species include loss of habitat, population fragmentation, predation by human-subsidized meso-predators, mortality from roads and other anthropogenic disturbances (such as farming and logging), and collection and removal for the pet trade. This effort represents the first attempt to develop a multi-state, regional conservation plan for the species in the United States.

Figure 1. The native range of Wood Turtle in North America.



Currently, there is little population genetic information on the Wood Turtle across its range. One study found little genetic variation and structuring across the range of the Wood Turtle examining mitochondrial DNA (Amato et al. 2008). Several other studies have used nuclear microsatellite DNA markers to examine patterns of population genetic structure at smaller geographic scales, within or across adjacent major basins (Tessier et al. 2005; Castellano et al. 2009; Spradling et al. 2010; Fridgen et al. 2013; Willoughby et al. 2013). Although these studies provide some information about the genetic status of the Wood Turtle, the limited geographic scope precludes the identification of species-wide genetic diversity. Therefore, the CSWG proposed genetic sampling to support the broad conservation planning efforts. The participants of the Wood Turtle CSWG collected tissue samples from across the eastern portion of the species' range to guide conservation planning efforts.

The objectives of this portion of the study were to: 1) describe population genetic diversity (heterozygosity, allelic richness, private alleles); 2) identify the most likely number of population groups in the study area; 3) measure relative isolation by distance comparing genetic and geographic distances; 4) estimate contemporary migration rates; and 5) test population genetic assignment methods to identify the origin of confiscations from the illegal animal trade.

Methods

Tissues were collected from participating states in the Northeast including: Maine, New Hampshire, Massachusetts, Connecticut, New Jersey, Pennsylvania, Maryland, Virginia and West Virginia. Samples were also collected by cooperators in Vermont, New York and Rhode Island. Additional samples were submitted from other studies from the Midwestern U.S. as out-groups for this study. Analysis is not completed for these samples, but preliminary data and population groups are shown in Appendix F to group confiscated, captive and unknown samples from this study.

Tissue was collected as blood, tail tissue, toenail, and shell. Other soft body parts were occasionally collected from recent mortalities (such as toes or foot). Blood was preserved in 95% ethanol, lysis buffer (e.g., Queens lysis) and PBS, depending on the collector. Other tissue types were preserved in 95% ethanol. Samples were stored at -20°C until processed in the lab.

Laboratory Methods - DNA extraction varied for tissue types. DNA was extracted using a MoBio Ultra Clean Tissue and Cells DNA isolation kit ™ (MoBio, Inc., Calsbad, CA) (blood, tail) or a Qiagen DNeasy Blood and Tissue Kit ™ (Qiagen, Inc., Germantown, MD) (blood, tail, nail, shell) according to manufacturer's protocols. Blood and tail and other soft tissue were incubated overnight at 55°C, and nail and shell samples were incubated for 2 days on a shaking incubator (Henry Troemner LLC, Thorofare, NJ). The concentration of DNA was measured in each sample using a Nanodrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE). Samples with DNA concentrations greater than 40 ng/ul were diluted with DNA/RNA-free water to 20-25 ng/ul.

Seventeen microsatellite markers were selected based on the performance in other published studies of Wood Turtle. Most markers were described for Bog Turtle (*Glyptemys muhlenbergii*), but were tested on the Wood Turtle (King and Julian 2004). Additional markers were included that were described for Blanding's Turtle (*Emydoidea blandingii*) and Painted Turtle (*Chrysemys picta*) (Table 1). Four multiplexes were performed with 3 to 5 markers each. *GmuD51* was isolated for the PCR reaction and then added to multiplex 1 prior to electrophoresis.

Table 1. Microsatellite loci, citations and multiplex assignment used for this study.

Locus	Citation	GenBank no.	Wood Turtle Genetics Studies by location						SWG study multiplex
			Castellano et al. 2009	Chinnici and Huffman 2016	Spradling et al. 2010	Willoughby et al. 2013	Fridgen et al. 2013	Tessier et al. 2005	
GmuA19	King and Julian 2004	AF517227					X		4
GmuA32	King and Julian 2004	AF517228				X			2
GmuB21	King and Julian 2004	AF517231		X	X	X	X	X	4
GmuD16	King and Julian 2004	AF517235	X	X	X	X		X	1
GmuD28	King and Julian 2004	AF517237			X	X	X		3
GmuD40	King and Julian 2004	AF517238	X	X	X	X	X	X	4
GmuD51	King and Julian 2004	AF517239	X	X					1
GmuD55	King and Julian 2004	AF517240			X	X			2
GmuD62	King and Julian 2004	AF517241		X					
GmuD70	King and Julian 2004	AF517242							4
GmuD79	King and Julian 2004	AF517243			X				3
GmuD87	King and Julian 2004	AF517244	X	X	X	X	X	X	2
GmuD88	King and Julian 2004	AF517245	X	X	X	X			1
GmuD90	King and	AF517247			X[a]				

GmuD93	Julian 2004 King and	AF517248	X[a]	X		X	X	
GmuD95	Julian 2004 King and	AF517249	X[a]	X				
GmuD114	Julian 2004 King and	AF517251		X	X[a]			
GmuD121	Julian 2004 King and	AF517252			X			3
Eb17	Julian 2004 Osentoski et al. 2002	AF416295						1
Eb19	Osentoski et al. 2002	AF416296						3
BTCA9	Libants et al. 2004	AY335790						2
Cp2	Pearse et al. 2000							2

[a] indicates the marker was dropped from analysis

We conducted 10 μl PCR reactions in a 96-well plate using a thermal cycler (MJ Research, PTC-200). Each reaction consisted of 5 μl of Qiagen Multiplex PCR Master Mix, 1 μl template DNA, 1 μl of primer (6-FAM primers were 0.15 uM, PET and VIC primers were 0.2 μM and NED primers were 0.25 μM), and 3 μl of PCR grade water. The PCR reactions were adjusted for nails to include 1 μl BSA and 2 μl of PCR grade water. Forward primers were fluorescently labeled and acquired from Applied Biosystems (colors NED and PET, Foster City, CA) and Integrated DNA Technologies (colors 6-FAM and VIC, Coralville, IA). Thermocycling conditions included an initial 15 mins at 95°C followed by 35 cycles of 94°C for 30 sec, 57°C for 90 sec (or 51°C for *GmuD51*), 72°C for 90 sec and a final cycle of 72°C for 10 mins. One negative control was included on each plate. PCR products were diluted to 1:50 with PCR grade water. The PCR products were run on an Applied Biosystems 3130xl Genetic Analyzer with a LIZ600 ladder for size standard. Peaks were scored using Geneious version 9 (Biomatters Ltd, Auckland, New Zealand). Peaks were visually checked for conformity to expected profiles. Duplicate samples for the quantification of error rates ranged for the multiplex and the locus. The number of duplicate samples ranged from 33 for multiplex 1 to 48 for multiplex 2. The number of re-run samples by locus ranged from 23 for *GmuD51* to 48 for *GmuD87*. The percent error was estimated as the percent of alleles from the total duplicated samples that were not equal. This estimate would include scoring error, binning error, variation in runs and null alleles.

Statistical Analysis - Individuals without location information or with fewer than 8 successfully genotyped loci were removed from the data set prior to statistical analysis. Sites with fewer than 5 individuals were removed prior to statistical analysis. One individual identified from a pair of full siblings with a 95% confidence using 100 randomizations in ML Relate (Kalinowski 2006) was removed from the data set; this was done to avoid bias in site-based population genetic measures.

Exact tests for deviation from Hardy Weinberg proportions and linkage disequilibrium were performed using GENEPOP version 4.5 (Raymond and Rousset 1995). Heterozygosity and unbiased estimates of allelic richness and private alleles were calculated using HP Rare (Kalinowski 2005). F_{IS} was calculated using GENALEX (Peakall and Smouse 2006, 2012). A log likelihood G test from Goudet et al. (1996) in GENEPOP version 4.5 was used to test for genetic differences among sites. A Bonferroni correction was applied to all significance tests with multiple comparisons (Rice 1989).

We used STRUCTURE version 2.3.4 (Prichard et al. 2000) to estimate the number of populations. STRUCTURE is a Bayesian-based model that clusters individuals according to allelic frequencies while

minimizing linkage disequilibrium and deviation from Hardy–Weinberg equilibrium. The model allows for admixture between population groups. The admixture model with correlated allele frequencies in STRUCTURE was run by using 10,000 iterations for burn-in and 100,000 iterations with a Markov-chain Monte Carlo resampling algorithm as described by Pritchard et al. (2000). Ten runs were performed for each K value tested (K=1 to 20). Data from sites with more than 15 genotyped individuals were used in this analysis except NY where sites had 11 individuals. We conducted an initial analysis on all the genotyped individuals included in the complete data set, hereafter referred to as uneven data set. Due to bias inherent in structure-based analyses (see Kalinowski 2011; Puechmaille 2016), we performed a secondary analysis on a subset of the data by reducing sites to a sample size of 16-18 individuals, hereafter referred to as the even data set. Individuals with incomplete data were removed from the data set first, followed by a random selection as needed. Finally, a last set of runs was performed with the location prior option, which uses the capture location as a prior in the model. STRUCTURE output was compiled and visualized using STRUCTURE HARVESTER (Earl and vonHoldt 2012). After identifying the K value, a final run (hereafter called full data set) with all sites with n>6 and individuals with 14 or more loci were tested using this optimal K value.

A K-means test was performed using the even sample in GENODIVE version 2.0b27 using a Bayesian Information Criterion (BIC) (Meirmans and Van Tienderen 2004) to verify the number of clusters we identified using STRUCTURE. The K-means clustering identifies the optimal clustering as the K value with the smallest amount of variation within clusters, which is calculated using the within-clusters sum of squares. The value of K with the lowest BIC value is identified as the best fit for the data. Finally, a principal components analysis (PCA) was performed in GENODIVE using allele frequency data and co-variance matrix, and graphed in R version 3.4.1 (R Corps Team 2013).

Isolation by distance was tested using a Mantel test on pairwise F_{ST} values and geographic distance (Euclidean and stream distance) for sites within the clusters identified using STRUCTURE. The test was performed using IBDWS version 3.23 using 1000 randomizations (Jensen et al. 2005). Several sites are not connected by stream or river corridor in the clusters, such as the Allegheny River and the Potomac River sites. The stream distance tests were done using 10,000,000 km as a pairwise distance value for these unconnected sites. The stream distance analysis was also performed without including the unconnected sites.

Full siblings and parent-offspring pairs were identified among the sites within the clusters identified in the structure analysis. Colony version 1.2 (Wang 2004) was used to identify full sibling groups. Simulations for similar analyses found that full sibling groups with 3 or more individuals was 97% accurate using 16 loci (Whiteley et al. 2014). This method has also been found to out-perform STRUCTURE for identifying recent migrants (Whiteley et al. 2014). Sites were grouped according to the major population groups identified via the Bayesian clustering analysis were used for this test. Population groups in northern Maine, ME/NH, western MA, and Potomac were tested for family groups. The specific sites used for this analysis are listed in Appendix E, Table E.2. Some overlap was considered between northern ME and ME/NH in order to test movement between sites and across major population groups.

Samples from the known collections (sites) were tested in GENECLASS version 2.0h (Piry et al. 2004) using a leave-one-out method where each sample is sequentially removed from the data set and assigned to a population. This test provides an estimate of accuracy for assignment success. Additionally, unknown samples from captive populations or from the illegal pet trade were assigned to reference populations. The frequentist analysis described in Petkau et al. (1995) was used for this test as it slightly out-performed the other options. Individual samples with more than one missing locus were removed from the data set prior to performing this analysis.

Results

The Wood Turtle CSWG participants collected more than 1,895 tissue samples of various tissue types (Fig. 2). Samples were prioritized for genotyping based on site location, sample size and success by tissue type. Blood and soft tissue (toes, tail tips) were selectively chosen for genotyping due to ease of extraction and higher success rates. Toenails were highly successful, but only when an adequate amount of nail and associated soft tissue was sampled (see Lutterschmidt et al. 2010 for details on toenail tissue success). Shell shavings and scutes also had sufficient success rates. Nails and shells were used to increase sample sizes when other tissue types were not available. To estimate the success rates of the tissue types, we examined a sub-set of samples where the collection and treatment of the samples provided a fair estimation of the sample success. A tissue was considered failed if 3 or more loci were missing. Tail tissue was the most successful (97%), followed by blood (87%). We did not examine blood by preservation method, but ethanol and PBS provided higher success rates and more ability to manipulate the amount of tissue used in the extraction. Toenails were successful when the nails provided sufficient soft tissue, and the more successful samples ranged from 70% to 94.5% but certain sites provided high failure rates (90-100%) when the sample collection did not provide adequate tissue. Shell samples were the smallest portion of our data set and were about 60% to 80% successful depending on the collector.

Figure 2. Tissue types included in the study (blood, tail or soft tissue, toenail, shell shavings) collected and genotyped for this study.

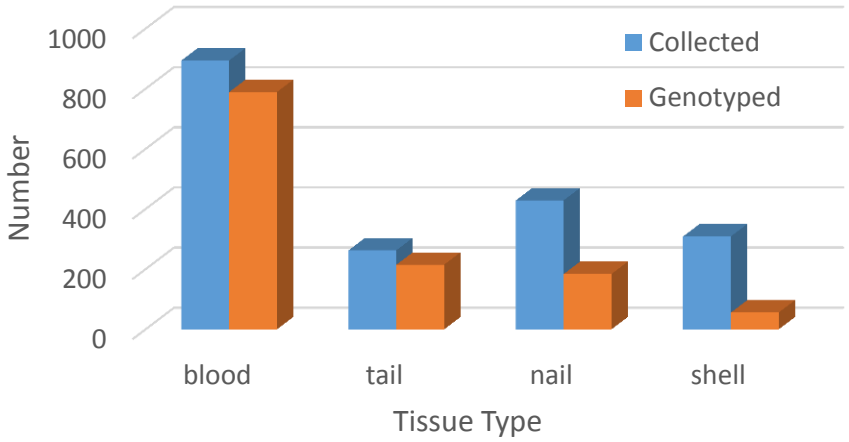
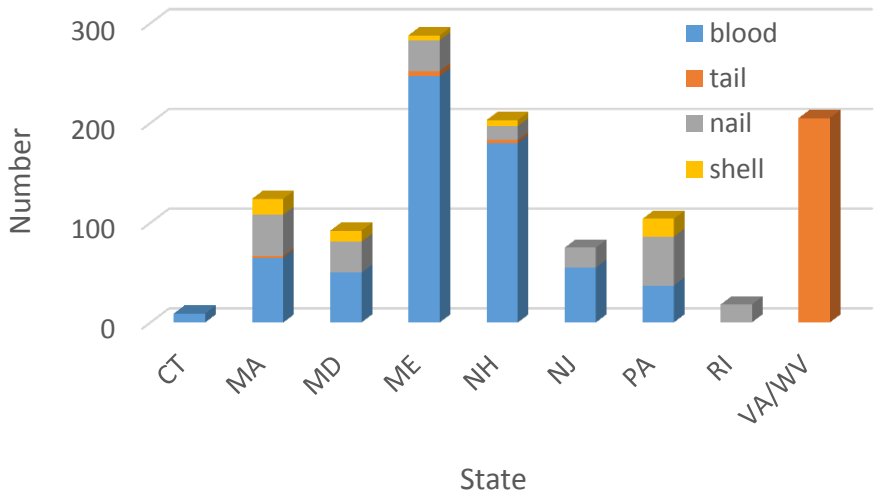


Figure 3. Number of samples successfully genotyped by state and tissue type (blood, tail or soft tissue, toenail, shell). The sample must have >7 loci amplified to be considered successful and included in data analysis.



Tests for Assumptions and Genetic Diversity

Samples sizes ranged from 5 to 50 individuals (average n=17.4) collected from 62 sites. One locus, *Gmu A19*, was removed due to scoring difficulties. For the remaining 16 loci, genotyping error ranged from 0 to 3.4% (Table 2). Exact tests for deviations from Hardy Weinberg proportions identified significant deviations for *GmuA32* at three populations (MA Worcester, NH Turpentine, NH ArBar), *GmuD51* at two populations (MA Wildcat, NJ Potato) and *GmuD21* at one population (PA Coral). Significant linkage disequilibrium was detected at 6 pairs of loci, but there was no pattern to the loci or populations. Based on these results, we kept all these loci in the analysis due to potential for a significant test based on random chance, uneven sample sizes (which we address later), and the robustness of many statistical tests to deviations from Hardy Weinberg proportions.

Unbiased allelic richness ranged from 3.4 to 6.2 (average 5.1), private alleles ranged from 0 to 0.3 (average 0.05), unbiased expected heterozygosity ranged from 0.5 to 0.7 (average = 0.6) and F_{IS} ranged from -0.21 to 0.14 (average =0). F_{ST} ranged from 0 to 0.23 (average 0.07). The overall genetic diversity is within the range documented in other studies of Wood Turtle (Table 3).

Table 2. Loci, size ranges (bp), number of alleles, percent failed amplification and genotyping error for this study.

Locus	Size Range min	Size Range max	No. alleles	% fail amplification	Genotype error	Comments
GmuA19						Removed due to difficulties scoring
GmuA32	147	208	29	6.7	0	Stutters; high failure in USFS Midwest samples
GmuB21	193	204	21	3.2	0	
GmuD16	151	237	27	18.6	3.4	

GmuD28	185	258	25	3.6	0
GmuD40	136	197	24	3.9	0
GmuD51	220	396	51	8.7	0
GmuD55	182	204	17	10.6	0
GmuD70	151	193	9	12.4	0
GmuD79	149	265	6	10.2	0
GmuD87	226	303	25	4.5	1.0
GmuD88	102	185	29	3.2	0
GmuD121	124	174	12	1.8	2.3
Eb17	88	104	11	3.0	1.7
Eb19	92	97	4	1.7	0
BTCA9	136	152	8	4.0	1.1
Cp2	188	263	22	8.7	0

Table 3. Summary of genetic diversity in studies of Wood Turtle by citation, location, number of loci, unbiased allelic richness (asterisk indicates number of alleles as only value reported), expected heterozygosity and F_{IS}.

Citation	Location	No. loci	AR	He	F_{IS}
Castellano et al. 2009	Delaware Water Gap	7	10.3-13.8	0.88-0.95	-0.20-0.019
Fridgen et al. 2013	Southern Ontario, Canada	5	3.6-5.4	0.48-0.87	-0.07-0.33
Spradling et al. 2010	Iowa, Minnesota, WV	11	1.0-16.4	0-0.9	0-0.007
Tessier et al. 2005	Quebec, Canada	5	13-36*	0.8-0.89	
Willoughby et al. 2013	Michigan	9	7.11-10.7*	0.37-0.91	-0.10-0.48
This study	**NE – mid Atlantic states**	**16**	**3.4-6.2**	**0.53-0.70**	**-0.21-0.14**

Three sites in Maine (Tananger, Arroyo Frijoles, and Arroyo Colorado) were tested for genetic differences by age class. These data do not indicate any differences across the ages within a site (Table 4), and all of the pairwise allele frequency exact tests were not significant following correction for multiple tests. The sample sizes for juveniles are low which is likely due to collectors avoiding natal areas and young turtles during sampling. Similarly, Fridgen et al. (2013) did not find statistically significant differences among age classes at a site in Ontario, Canada. Although these results indicate little genetic changes among the broad age classes at these sites, this type of test could be improved with greater sample size targeting as large an age range as possible and implementing this test in areas where populations are fragmented and/or declining. It should also be noted that this test was only possible in these relatively intact sites due to sampling limitations.

Table 4. Genetic diversity (observed heterozygosity, unbiased expected heterozygosity, unbiased allelic richness, unbiased private alleles) of age classes at three sites. Age categories are juvenile (J=<15 years old), middle (M=15-25 years old) and oldest (O=>25 years old). Ages were provided by M. Jones, unpublished data.

Site	Age	N	Ho	He	AR	PA
Tan	J	4	0.63	0.57	3.3	0.5
	M	10	0.52	0.59	4.7	0.8

	O	4	0.61	0.63	3.9	0.6
ArF	J	4	0.55	0.60	3.4	0.3
	M	9	0.53	0.58	4.9	1.3
	O	13	0.56	0.61	4.7	0.9
ArCol	J	5	0.61	0.57	3.5	0.2
	M	8	0.50	0.60	4.6	0.8
	O	8	0.60	0.60	4.7	1.0

Population Differentiation

The allele frequency (genic) exact tests indicate that 91% of the pairwise comparisons were statistically significant after correcting for multiple tests. F_{ST} ranged from 0.01 to 0.23, and generally lower F_{ST} values will correspond with insignificant allele frequency differences. Both tests indicate the amount of pairwise differences among sites. F_{ST} values among the sites are shown in Appendix C.

Within Maine, Camel Hut and Big Cypress were not significantly different, while all of the other pairwise comparisons were significantly different. F_{ST} values ranged from 0.03 to 0.13. Many of the sites in NH were not significantly different. The NH Fortification site showed the greatest divergence from the other sample sites, but all of the sites in NH had lower F_{ST} values (0.01 to 0.08). In general, most of the pairwise tests across sites within the Merrimack basin were not significant. Dead Lizard and Arroyo del Cuervo were not significantly different from several of the Merrimack sites. Bullhead was not significantly different from Sourdough and Crow but was significantly different from other sites in the Connecticut basin. Overall, the sites sampled in NH appear to have some migration among the sites and low genetic drift.

Within Massachusetts, Bumblebee, Charcoal House, Wildcat and Little Bearskin were not significantly different and could be considered one subpopulation. All of the MA sites had F_{ST} values ranging from 0.01 to 0.08 with the higher values (0.05-0.08) associated with MA Worchester. The other pairwise F_{ST} values at MA sites ranged from 0 to 0.03. MA Worcester was significantly different from the other Massachusetts sites, but not significantly different from the Rhode Island site. Connecticut Wheeler was not significantly different from NH Pickle and NH Millstone.

Maryland sites (Mary Davis, Wolfpen, Moose Meadow and Tomahawk) were not significantly different from each other, but were significantly different from Pumpkin Field. MD Pumpkin Field also had the highest F_{ST} values within the MD sites with the F_{ST} values ranging from 0.02 to 0.04 among the sites. Among the New Jersey sites, Potato, Barney, and Jackie were not significantly different and are likely one subpopulation. Barney and Sucker and Barney and Bulldozer were not significantly different. Williamson was the most divergent of the sites sampled in the state. F_{ST} values for NJ sites ranges from 0.02 to 0.11. In New York, Barrel Ranch and Yankee were not significantly different. In Pennsylvania, Snow was not significantly different from Nancy and Nancy was not significantly different from Coral. F_{ST} values in PA ranged from 0.02 to 0.04.

In Virginia/West Virginia, many sites sampled were not significantly different. St. Sebastian, Silvertip, Box Canyon, Waterfall, Hidden, Diversion and Lone Tule were not significantly different from each other and should be considered one subpopulation. Chicken Run was not significantly different from Waterfall Wash indicating some genetic exchange or lack of genetic drift. Chicken Run and August were the most divergent sites included in the state, and Chicken Run was the most divergent site in the complete, northeast sample showing the largest F_{ST} differences from the sites further north. F_{ST} values from the VA/WV sites ranged from 0.03 to 0.06.

STRUCTURE indicated that there are likely 3 to 5 clusters. The likelihood plot of the number of clusters (K) did not change substantially between the uneven and even runs without location prior and the even with location prior runs (Fig. 4). The clusters identified with and without the location prior provided similar clustering results (data not shown). Our presentation is focused on the results without using the location prior.

The most distinct clusters were the differentiation of the northern sites from the southern sites and sites in coastal MA/RI. For example, at K=4, the clusters are: coastal MA (MA Wo/RI); Potomac/Allegheny sites (MD, VA and WV); northern ME (Ar. Coyote, Ar. Frijoles, Ar. Tio Lino, Ar. Colorado, Ar. Yupa, Camel Hut, Tanager, Baxter); and NJ/NY sites. The Connecticut (MA Cr, MA Bu, MA LB), Merrimac (NH Tu, NH ABa, NH Fl, NH Cy), and Kennebec (ME Sm, ME CaH) basins indicate mixed ancestry between the coastal MA/RI cluster and the northern ME cluster, and should be considered a genetically similar group. Some locations in PA and NY showing mixed ancestry (Fig. 5). PA has one site that groups with the Potomac and two sites in the Susquehanna basin that cluster with NJ/NY. The sites in PA show admixture between NJ, MA and VA. The NH and coastal MA/RI sites separate from the ME sites when increasing from K=3 to K=4, while increasing from K=4 to K=5 separates the NH cluster from the coastal MA/RI cluster (Fig. 5). The K-means test also indicated 4 clusters as the best fit for the data, providing additional support for the STRUCTURE inferences.

Figure 4. Log likelihood for the number of clusters (K) in the data sets with (a) uneven sample sizes without location prior; (b) even sample sizes without location prior and (c) even sample sizes with location prior. Note different scale on x-axis for (c).

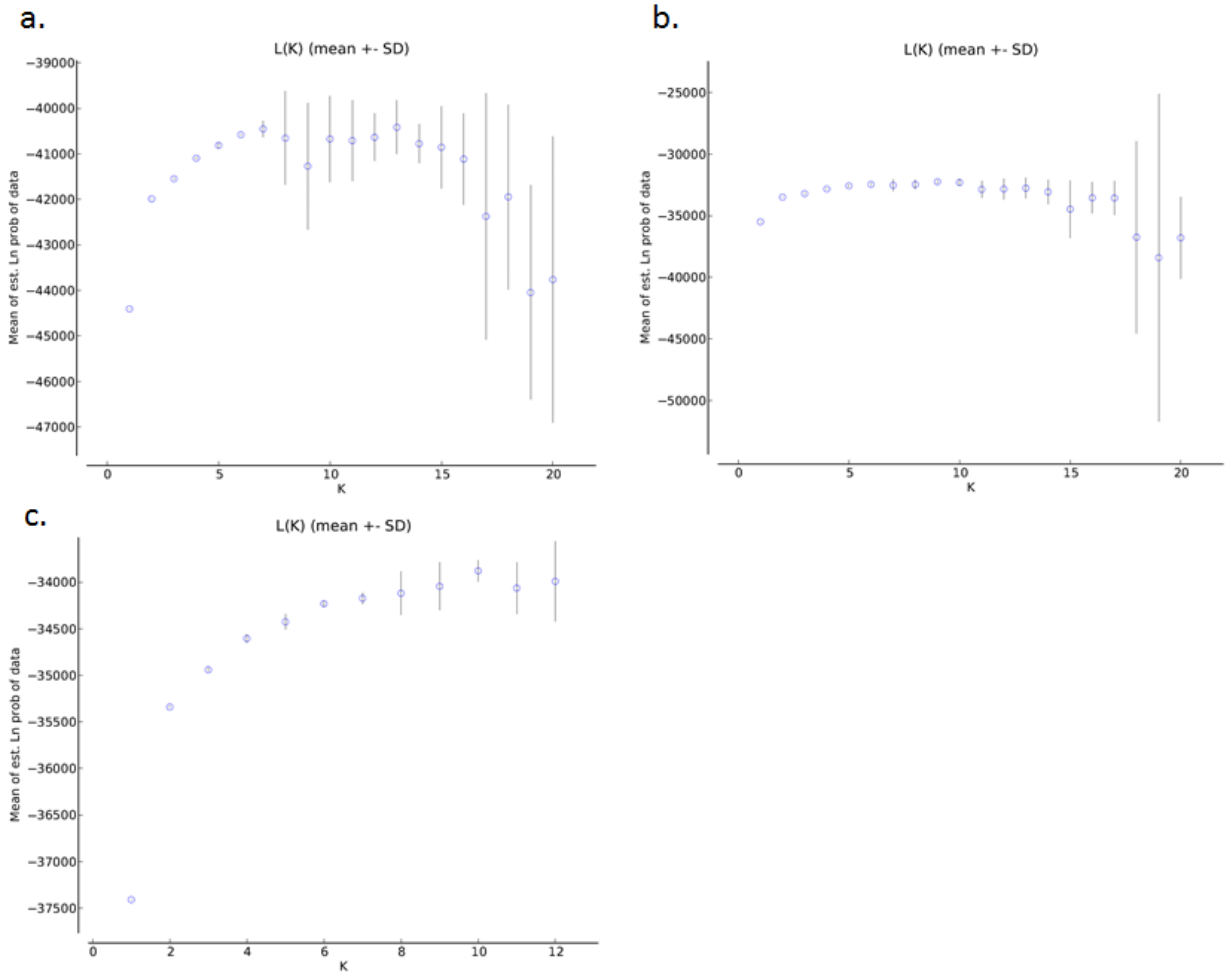
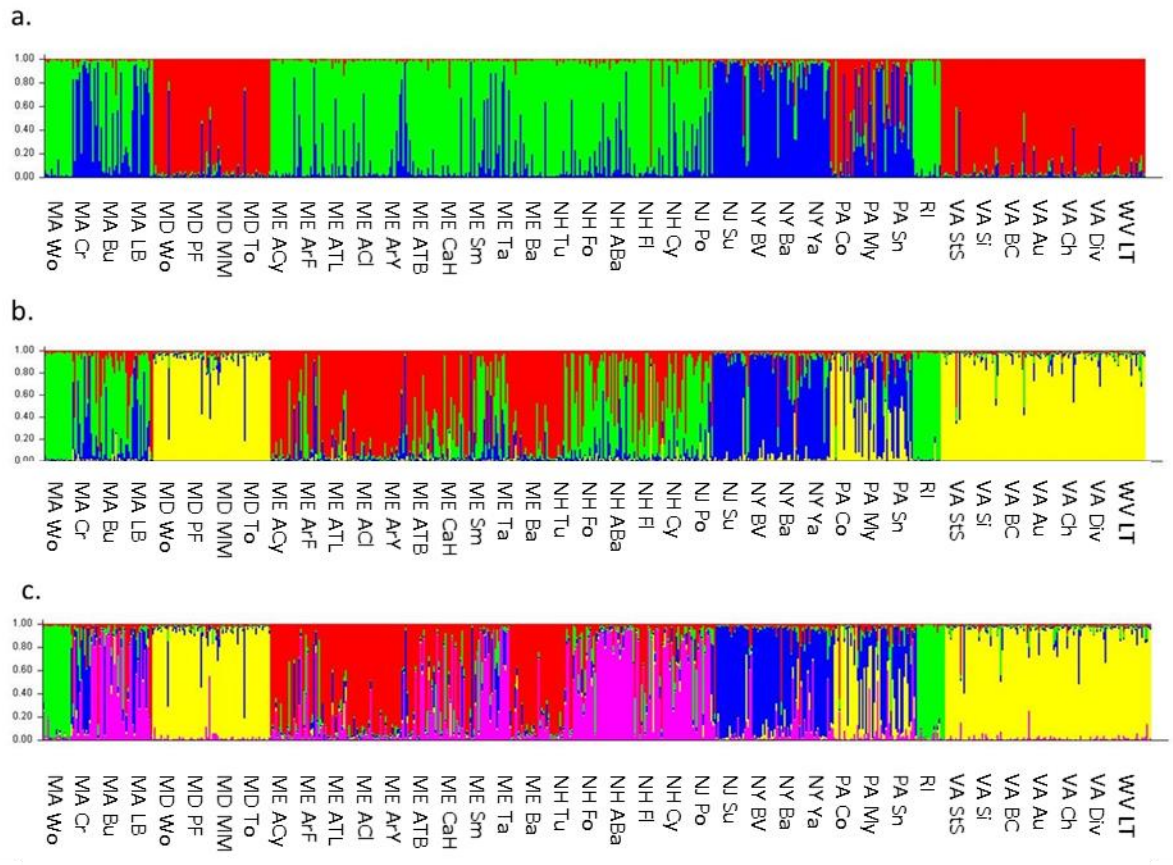


Figure 5. STRUCTURE plots for the even sample size without location prior runs for (a) K=3, (b) K=4, and (c) K=5 clusters. The *y*-axis shows the admixture coefficient (*Q*-value) and each bar or column in the figure represents one individual Wood Turtle. Site abbreviations are shown below the *x*-axis.



Principal components analysis revealed similar major groups as those identified with the STRUCTURE and the K-means analyses. The northern and southern sites were the most distinct clusters, the coastal MA/RI cluster was separate from both of these clusters and the sites geographically between these showed a gradation primarily along the cluster with the northern sites. PA, NJ and NY fall in between the northern and southern clusters (Fig. 6). One site (MA Crosby) clustered with the NY/NJ group, but locates just to the right of the ME/NH group. The other western MA sites group with the ME/NH group but toward the bottom of the cluster in the direction of the eastern MA/RI group.

We used analyses of isolation by distance within the major STRUCTURE-defined clusters of populations to test for nearest-neighbor patterns of gene flow (over land or via waterways). The isolation by distance tests for sites in the NH/ME group and the Potomac group were significant for Euclidean distance tests; however, the test in the northern Maine cluster was not significant (Table 5, Fig. 7). The northern Maine sites were significantly correlated with stream distance, but 13 out of 28 pairwise comparisons were not connected by stream corridor. Eliminating these unconnected data points reduced the correlation among sites. The NH/ME group was significantly correlated with stream and Euclidean distance, but the correlation between Euclidean distance and genetic distance was stronger. These results suggest that gene flow occurs both over land and by waterways. This group had too few sites connected by stream corridor to perform this test excluding the unconnected sites.

Table 5. Summary of Mantel test by major population group, pairwise distance calculation, unconnected sites out of total pairwise comparisons (unconnected/total), correlation coefficient (r) and p-value (p) for the isolation by distance tests. Sites that were not connected by a stream corridor were given a maximum value of 10,000,000 km pairwise distance in the stream corridor distance test. NA indicates there were too few populations to perform the analysis.

Population Group	Distance	Unconnected / Total	r	p
North Maine	Euclidean	0/28	0.21	0.170
	Stream	13/28	0.67	0.001
	Stream connected only	0/15	0.47	0.045
NH-ME	Euclidean	0/21	0.74	0.002
	Stream	16/21	0.60	0.011
	Stream connected only	0/5	na	na
Potomac	Euclidean	0/66	0.74	0.001
	Stream	21/66	0.34	0.070
	Stream connected only	0/45	0.15	0.160

The average site Q values from STRUCTURE run on the full data set with K=4 showed similar patterns as Fig. 5 (Fig. 8). Specifically, the CT and western MA sites showed mixed ancestry. CT has more of the NJ/NY ancestry and western MA sites show more of the coastal MA influence. ME Monroe and ME Smiley show ancestry similar to the NH sites and NY and PA show mixed ancestry with the major influence from the NJ/NY group.

Dispersal and Relatedness Tests

Full siblings were detected among the ME Aroyo Coyote and Tanager sites (individuals 237, 245 and 493) and Camel Hut and Baxter sites (individuals 433, 444, 214). Euclidean distances between these sites are 50.6 km and 30.5 km, respectively. Detection of related individuals is sample size dependent, so these results should be interpreted similarly to presence/absence data, where absence of detection does not indicate a lack of connectivity. Other clusters tested for full sibling groups did not detect full siblings across sites (Potomac, NH/ME, and western MA). Full siblings detected within sites are listed in Appendix E, Table E.2.

Figure 6. Principal components analysis showing average allele frequency by site. Symbols represent the major clusters identified in the K-Means analysis. The values on the axes represent the percent variation explained by PCA 1 and PCA 2.

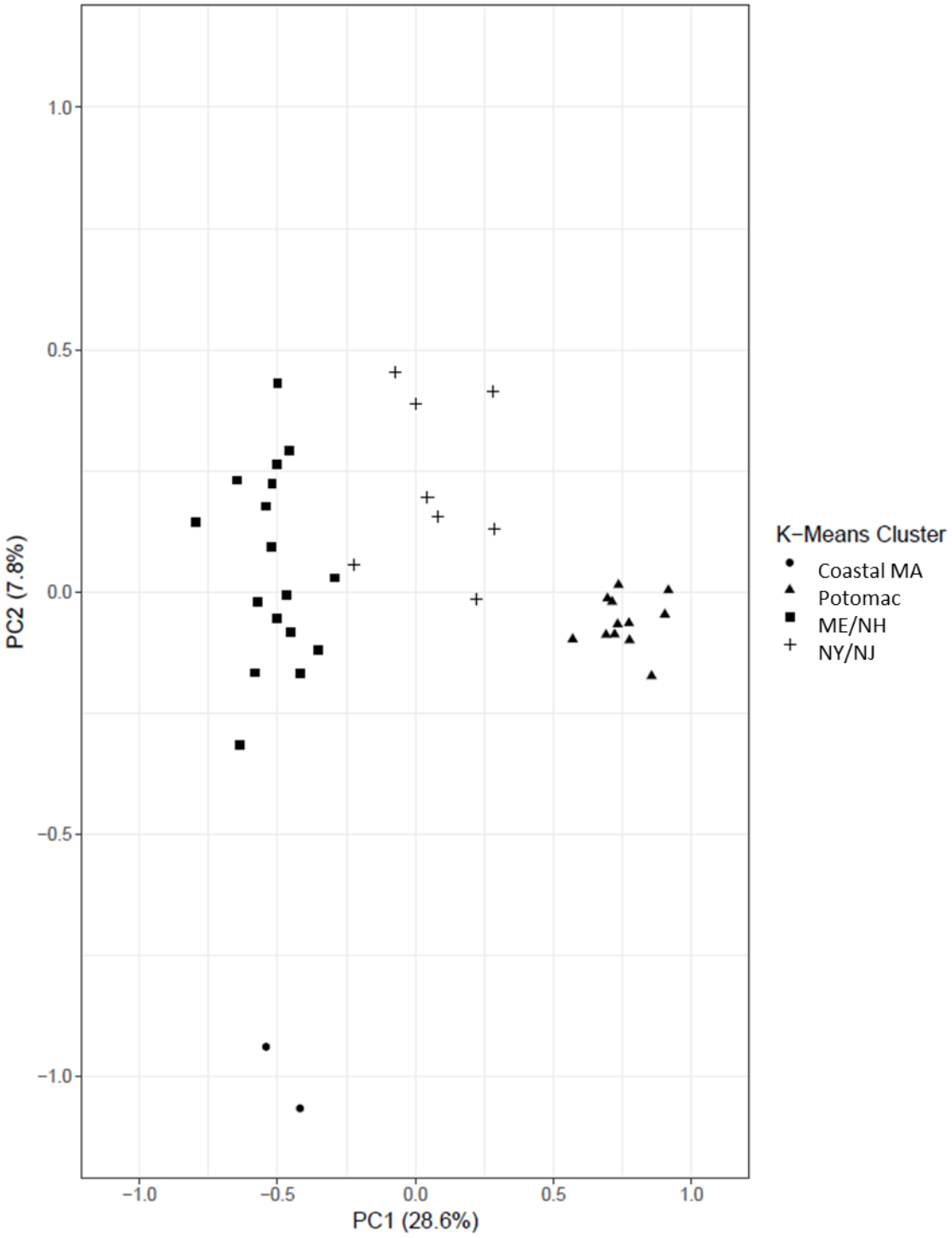


Figure 7. Isolation by distance tests for Euclidean geographic distance and pairwise F_{ST} among clusters identified by the STRUCTURE analysis (shown in Fig. 5b): a. NH/ME (red and green); b. Potomac (yellow); and c. northern Maine (red). The NH/ME and Potomac groups were significant, the northern Maine cluster was not significant (see Table 5).

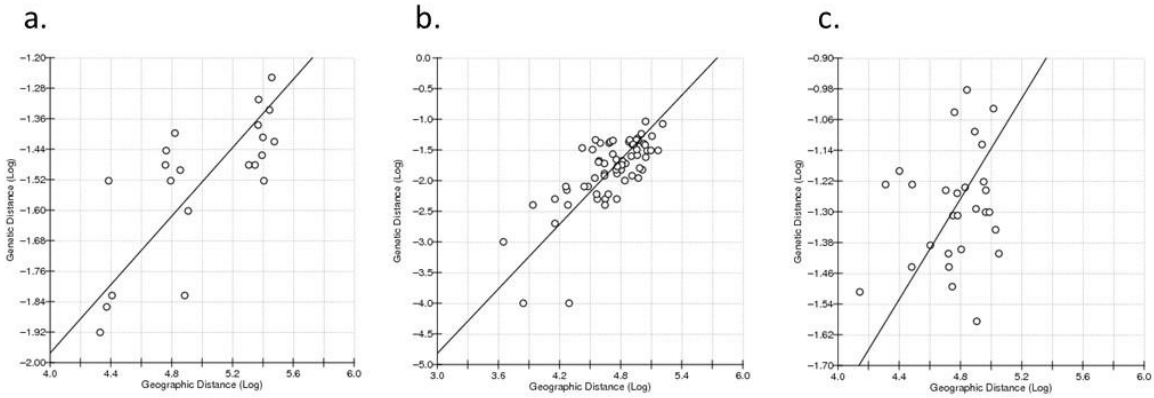
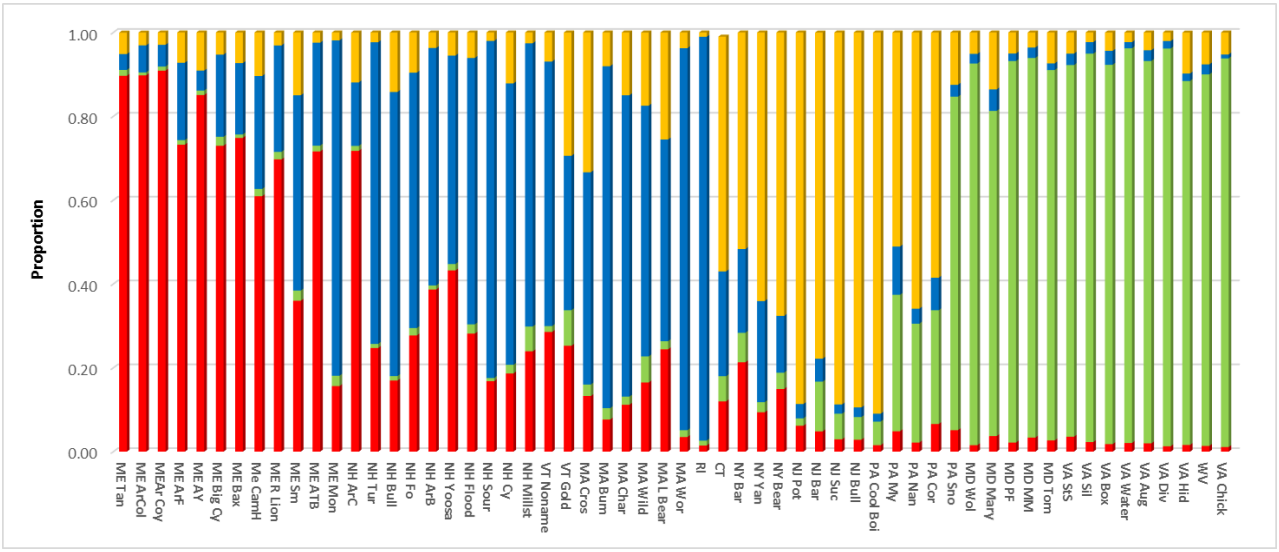


Figure 8. The average ancestry value (*Q*-value) for each major group identified in the Bayesian clustering analysis. Sites are ordered from north (left) to south (right) on the x-axis. This figure is color coded to match Figure 5. Each bar in this figure represents one site.



Individual Genetic Assignments

All sites with n≥7 were used in the genetic assignment tests. There was a weak relationship between sample size and the proportion of correct assignments (Fig. 9); however, genetic distinctness was also related to the proportion of correct assignments. In other words, the more genetic uniqueness from others (such as coastal MA) the greater the success in genetic assignment. Genetic assignment to the correct site was fairly low (51.9%). However, when allowing the individual to assign to any site where pairwise allele frequencies were not significantly different, assignment success increased and ranged from 12 to 100% by site (average 73%) (Fig. 10). Overall, genetic assignment varied geographically. Assignment success was generally high in the Maine and Potomac sites and low in the New York and Pennsylvania sites (Fig. 11). Therefore, genetic assignment will work better for some locations than others, and in most cases it is unlikely that the exact site of origin can be identified using these markers. However, in many cases, the correct major population group can be identified for more than 70% of the samples. Generally, the sites with higher admixture (such as PA and NY) had lower assignment success.

The unknown samples provided by the U.S. Fish and Wildlife Service from the southern part of the study area mostly assign to a Potomac site (56%). The remainder of the samples assigned to various locations in Maine, Pennsylvania and New Hampshire. These individuals were presumably from a site in the study area. However, these results could arise from samples from the Potomac cluster as some individuals show admixture from these other clusters (Fig. 5). One confiscated sample submitted from Massachusetts assigned to the Potomac cluster.

The unknown samples provided from captive populations in New Jersey may or may not originate from samples included in the study area. The NJA samples assigned to all states included in the analysis with the most samples assigning to the Potomac basin sites (33%), and the western MA sites (24%). Other individuals assigned to sites in New York, New Jersey, New Hampshire, Maine and Pennsylvania. One sample did not assign strongly to any sampled site indicating that the site of origin may not be included in the reference sites tested. The NJB samples similarly assigned to many different sites with 33% not assigning strongly to any particular sampled site, followed by 28% assigning to the Potomac cluster. Other sites included in this sample were New Hampshire, Pennsylvania, New Jersey and Maine.

Figure 9. Sample size and proportion of correct assignments for sites included in the genetic assignment test.

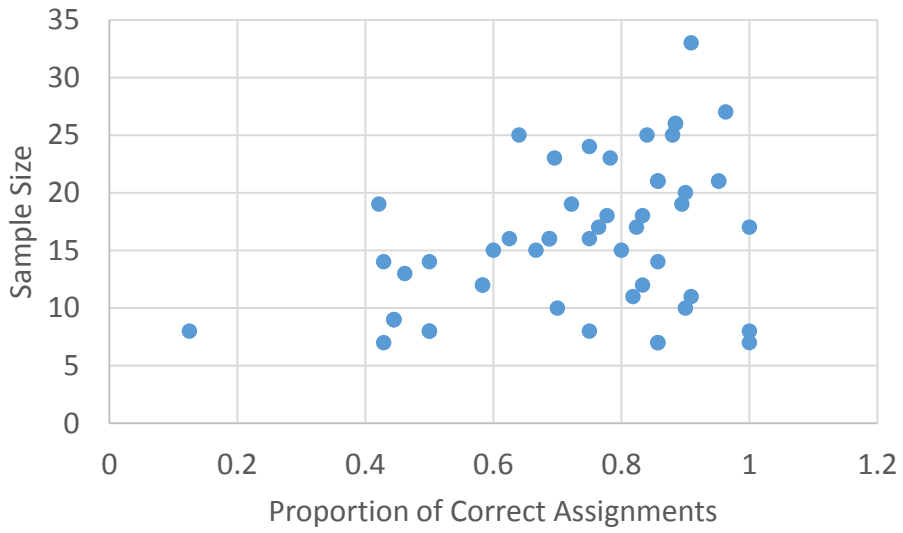


Figure 10. Proportion of correct genetic assignment by site. Black bars indicate the proportion of correct assignments to the site where the sample was collected. Gray bars indicate correct assignment to the sample site and any site where no significant allele frequency differences were detected.

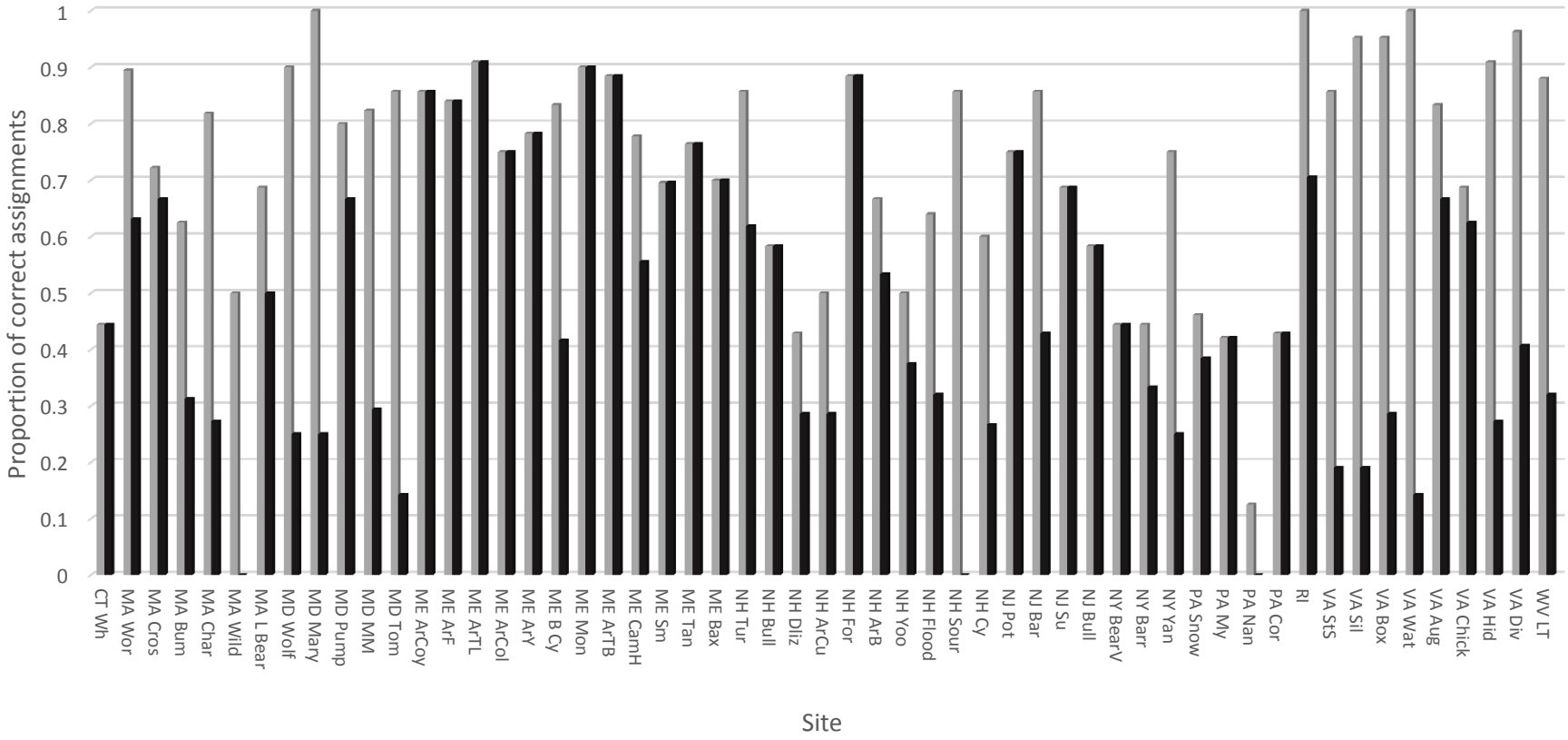
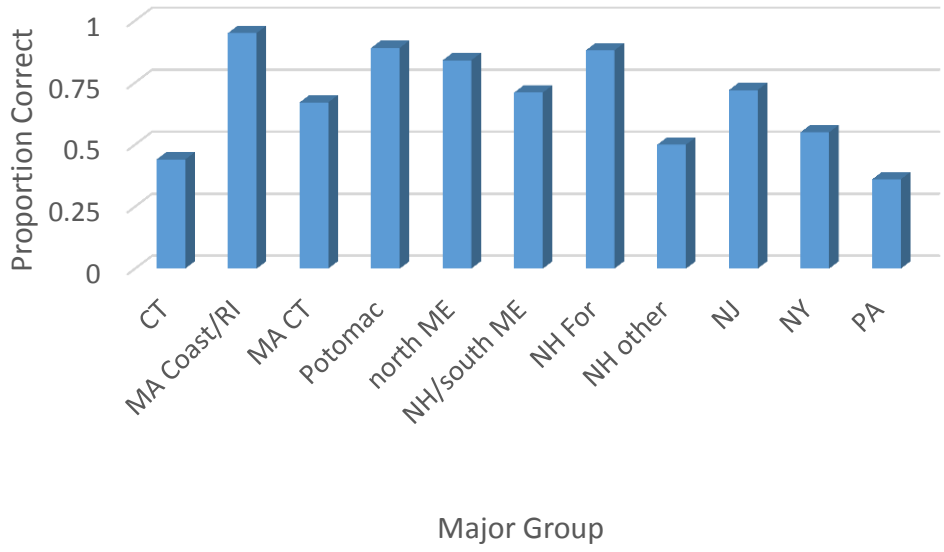


Figure 11. Proportion of correct genetic assignments by major group (or cluster).



Discussion

The regional, CSWG-funded project had several objectives for the population genetic analysis and implications for the conservation planning for Wood Turtles across the 11 northeastern states in the study area. The study objectives center on the conservation plan and management units for this planning effort that include an evaluation of genetic diversity and the identification of major population groups. This study makes inferences about the migration of turtles among the sampled sites. Lastly, this study quantifies the accuracy of population assignment for the potential application for releasing turtles confiscated in the illegal animal trade at or near their site of origin. Results from these analyses should be interpreted cautiously as Wood Turtle longevity and biology may not indicate current patterns and processes in the landscape. Specifically, contemporary genetic signals would be two generations which is equivalent to 100 years. Landscapes have changed substantially in the study area during the last 100 years, and processes such as fragmentation may not be detected for more than seven generations depending on the dispersal ability and analytical methods (Blair et al. 2012) and abundances.

The population genetic differences and admixture detected in this analysis may reflect historic genetic-demographic signals, current population interactions or a combination of historic and current effects. For example, the large population genetic differences can arise from a founder effect, population bottleneck or genetic drift, which is accelerated with smaller population sizes and isolation. Post-glacial colonization will also influence the genetic structure and admixture can result from colonization patterns or inter-population migration patterns. Demographic studies can assist in determining the level of population inter-action and dispersal abilities of a species; however, when migration rates are low but genetically influential, such as one individual per generation, identifying this signal from demographic studies can be

challenging (Lowe and Allendorf 2010). Additionally, demographic studies often cannot identify the difference between dispersal (movement) and successful migration or mating in a new population (genetic exchange). Due to the longevity of the Wood Turtle (~50 year generation), current population genetic data can reflect conditions as long as 100 years ago. Although this can make determination of current levels of connectivity challenging, it can describe gene flow and movement patterns in less developed and impacted landscapes before modern development and fragmentation, and can be used as a guide for conservation planning.

Genetic Diversity

Genetic diversity measured across the northeastern states was similar to other studies reported for Wood Turtles in the literature. Heterozygosity and allelic richness did not indicate loss of genetic diversity in the samples. The age-based test also did not indicate any genetic diversity differences across generations, but power to detect this trend is limited. Based on these tests, there is no indication of the detrimental effects of fragmentation or inbreeding in these samples; however, these results should be interpreted cautiously and with consideration of current demographic information as the longevity of the species and other behavioral attributes could potentially mask the genetic effects until the population has reached very low population sizes.

Population Differentiation

Evolutionary Significant Units (ESU) and Management Units are used in the conservation and management of threatened species. ESUs are populations or groups of populations that merit separate management or priority for conservation because of high genetic or ecological distinctiveness. Management units are generally smaller than ESUs and define demographic units for monitoring and management (Allendorf and Luikart 2007). Importantly, management units are demographically independent populations or meta-populations. Stepping stone models make the designation of management units ambiguous because admixture between the more divergent groups can impede the identification of clear boundaries (Palsboll et al. 2007). In this model, sites between major population groups are a combination, or rather, gradation of the groups. A trade-off exists between management units that are too large and do not provide adequate protection to the species and associated critical habitats, and those that are too small and may provide over-protection and undue costs of management or associated economic impacts. Genetic data are useful to quantify genetic distinctness among major population groups and subsequent management units, but should be considered a guide in the identification of distinct groups that are demographically independent.

Our results revealed a hierarchical genetic structure, with larger cohesive assemblages that exhibit stronger genetic differentiation. Within these cohesive assemblages, genetic differentiation was weaker, suggesting that there is more gene flow and possible metapopulation structure. The pattern in the clustering data generally indicated that the most genetically unique clusters in the study area were northern ME, coastal MA, Potomac and NJ. Areas of admixture were located between these major groups, such as the Merrimack, Connecticut and Kennebec basins and areas in PA and NY. Although the genetic data indicate four major clusters, we have recommended five major population groups to guide management planning. It might be warranted to consider these larger assemblages ESUs. Definition of ESUs for a species generally draws from genetic, life history, ecological, geological, and socioeconomic sources of information (Allendorf et al. 2013). We provide one of these sources of information here. Genetic differentiation among the major assemblages of populations is on the scale observed with ESUs of other species, such as Pacific Salmonids (NMFS 2018, WDFW 2018).

These major groups could be further divided into sub-groups (or management units) based on demographic independence. Most (91%) of the sites were significantly genetically differentiated from each other, indicating that the Wood Turtle is finely genetic structured across the study area. Whether subsets of populations within the major clusters should be considered MUs depends on determining the degree of demographic independence among the populations under consideration. Demographic independence will rely on the maximum dispersal ability of the Wood Turtle. Other studies of Wood Turtle did not detect significant genetic differences among sites <50 km unless there was a barrier to movement such as a large water body (Tessier et al. 2005; Castellano et al. 2009; Spradling et al. 2010; Fridgen et al. 2013; Willoughby et al. 2013). Therefore, sites less than 50 km apart with functional pathways for connectivity are probably not demographically independent. The pairwise F_{ST} and allele frequency tests indicated that the Wood Turtle is demographically supporting populations across drainage boundaries, and connectivity should be maintained

The patterns in the clusters we observed indicate that although grouping by major basin will capture much of the genetic diversity, there is some indication that gene flow or colonization has occurred across the headwaters of adjacent basins, such as the Potomac and the Allegheny, and the Delaware and Susquehanna. Therefore, an island stepping stone model describes the patterns of genetic structure and connectivity between the geographic areas representing each cluster may be important to maintain genetic diversity and exchange.

Migration and Gene Flow

Significant isolation by distance was detected in all the population groups tested. Isolation by distance in freshwater turtles has been detected in other studies, but appears to be spatially scale dependent. Specifically, at smaller geographic distances, isolation by distance is not detected (Castellano et al. 2009; Howeth et al. 2008). But, at larger geographic distances significant isolation by distance can be detected (Howeth et al. 2008; Shoemaker and Gibbs 2013). Yet, Sethuraman et al. (2014) found a positive but not significant correlation for isolation by distance in the Blanding's Turtle from sites located across Iowa, southern Minnesota and northern Illinois. Isolation by distance indicates a stepping stone model where neighboring subpopulations have a higher probability of sharing migrants. In the case of a freshwater turtle, the stepping stone model would be a two dimensional network of sites with the neighboring sites surrounding a site sharing individuals (Kimura and Weiss 1964). With this movement model, the sites most distant will show greater genetic divergence. Collectively, these studies indicate that population or group boundaries are fairly large (~100 km) for freshwater turtles. As populations decline, it will be increasingly more important to maintain connectivity among adjacent sites, and ideally this connectivity would be maintained across the entire study area to support the movement of turtles from one site or population to the next. Euclidean distance provided a stronger correlation with F_{ST} than stream distances for the Potomac and the NH/ME groups, but provided a weaker correlation for the northern Maine group. These findings indicate that overland corridors are more likely connecting sites than pathways along the stream corridor – particularly for the Potomac sites. It is also possible that the turtles are utilizing both types of corridors and perhaps for different purposes. For example, turtles may make local movements along the stream corridor while making less frequent and longer distance migrations overland. The genetic data suggest that overland movements happen across basins as well as within basins and most likely in an overland pathway that is closer to Euclidean distance than travel restricted within the stream corridor.

Little is known about longer dispersal distances for Wood Turtles. Only a few observations of longer range movements exist for Wood Turtles, and these movements were observed overland and along stream corridors. Individual turtles moving among sites are documented based on individual identification or

number codes, and movements up to 50 km are known to occur (T. Akre, personal communication). An individual male turtle equipped with a GPS tag moved at least 16 km overland and over basin divides in 6 months (T. Akre, personal communication). Turtle migrations may be necessary to reach critical habitats for feeding and reproduction, and could also be made by individuals emigrating from sites. Longer distance movements may be infrequent or sporadic based on alterations in habitat or high water events. Jones and Sievert (2009) documented Wood Turtles in Massachusetts dispersing up to 16.8 km after flood events. Turtles tracked by Jones and Sievert (2009) confirm that Wood Turtles can move overland or in the stream corridor. Additionally, it appears that males may be more likely to disperse longer distances than females, which had a higher rate of attempting to return to their home site (Jones and Sievert 2009). Long term studies are needed to accumulate observations to understand these movements and hence connectivity among sites.

Relatedness tests of full sibling groups may provide some indication of dispersal abilities. Based on this test, dispersal distances for Wood Turtles were a maximum of 50 km. Although, we cannot rule out human transport as a possible mode of movement, this 50 km distance is similar to the distance where genetic differentiation has not been detected in this species (Tessier et al. 2005; Castellano et al. 2009; Spradling et al. 2010; Fridgen et al. 2013; Willoughby et al. 2013). Increasing sample sizes would improve the conclusions from these analyses, particularly when considering maximum migration distances which is attempting to detect the few individuals that successfully migrate these long distances. Certainly, more information about the dispersal of the species, including the landscape attributes and habitats where the turtles travel would provide valuable information about corridors for managing connectivity between sites.

Genetic Assignment

Genetic assignment was only moderately successful for Wood Turtles in the study area and the level of success varied across sites. Certain sites in the study area have high site-level success rates where as other sites only can identify individuals to a major population group. Some sites had low genetic assignment success, particularly those with admixture from neighboring populations. Our study found that only 52% of individual turtles assigned correctly to the sample site. This low success rate can be due to closely located sites (<40 km) and a lack of genetic distinctness. A study of a freshwater turtle from South America found similar genetic assignment results where 59% of individuals correctly assigned to their sample site (Escalona et al. 2009). Tessier et al. (2005) found assignment success ranged from 84 to 98% when assigning individuals to population groups; however, this study was limited in geographic scope and examined populations divided by the St. Lawrence River which showed high genetic divergence between the north and south shore.

Based on our results, genetic assignment using the microsatellite markers we used would have limited application for enforcement in the illegal animal trade, and results may not be reliable if desiring to identify the exact site of origin for unknown samples. Newer population genomic methods that use large numbers of single nucleotide polymorphisms (SNPs) should be investigated for the potential for finer-scale differentiation among sites or smaller groups of sites due to the potential to obtain and efficiently genotype high numbers of loci (>100 – 1000's). SNP data generated by this method are also more easily compared across different laboratories and may provide finer genetic differentiation than microsatellites (see Malenfort et al. 2015). Alternate methods, such as permanent tagging methods like passive integrated transponder tags, may provide more certainty in the identifications and also allow more detailed demographic data to accumulate over the life span of the turtles, while also providing site of origin for enforcement and repatriation.

Recommendations and Data Gaps

Major Population Groups and Genetic Assignment

This study identified significant isolation by distance and a stepping stone pattern of admixture. The study identified four major population groups or clusters: northern ME, Potomac, coastal MA/RI, and NJ/NY. The Connecticut, Merrimack and Kennebec basins showed admixture between the coastal MA and the northern ME group and could be managed as an additional group based on similar genetic attributes. The sites included from PA showed admixture among the NJ, coastal MA and Potomac groups and we recommend should be managed according to the genetic admixture reflected in the data. For example, the Susquehanna basin should be managed with the NY/NJ group where it predominantly clusters, whereas the site in the Potomac basin in PA should be managed with the other Potomac sites.

Updating to genomic sequencing methods can provide many loci that improves the resolution for analysis of population differentiation. Additionally, these techniques have numerous applications in evolution and ecology that can assist conservation planning (see Andrews et al. 2016).

Migration and Connectivity

The site-based genetic differentiation combined with the estimate of contemporary migration rates and relatedness indicates that Wood Turtles are capable of migrating 50 km and perhaps greater distances. Therefore, sites less than 50 km apart should be managed to maintain connectivity to support adjacent populations. More information on maximum dispersal distances and habitat attributes associated with the movement corridors is greatly needed to identify the preferred migration habitats and target them for habitat restoration and conservation. Acquiring these data and associated GIS based analyses should be a high priority to inform conservation planning efforts.

Landscape and Conservation Planning

Landscape and conservation planning should strive to maintain long term genetic diversity and stable or increasing population growth. Therefore, the genetic data and population designations need to be considered in terms of demographic data (abundances, age class diversity, reproduction, sex ratios, and dispersal). All of these factors will directly influence genetic diversity and the resilience of individual populations. Data analyses in this report considered collectively with other genetic studies in Wood Turtles indicate that migration distances are more than 50 km. Additionally, considering the stepping-stone model of migration, connectivity among sites less than 100 km apart should be a high priority. Connectivity among sites across basins and also across the major population groups should be maintained in any planning and restoration efforts. This will allow populations to exchange individuals in source-sink dynamics, reduce the risk of extinction, and promote the conservation of genetic diversity.

Genetic Assignment

The success of the genetic assignments indicate that the population (site) where an individual was sampled could be correctly identified for 52% of the individuals in the sample, only slightly higher than random chance. When considering assignment to major population groups within which we detected no significant allele frequency differences, correct assignment ranged from 12 to 100%. High assignment success (>75% correct) could be identified for several population sub-groups: coastal MA, northern ME, Potomac and NH Fortification. These groups are genetically distinct from other groups. The success of assignment to the exact site where an individual was captured was relatively low, but identifying the

subpopulation or cluster from where an individual originated may be possible with these markers depending on where the individual originated. Assignment success was low (less than 50% correct) for CT Wheeler, NH (except NH Fortification) and PA sites, which limits the application of these markers in the enforcement of the illegal harvest of Wood Turtles across the broad geographic area.

A transition to next generation genomics could also improve population genetic assignments. SNPs have a lower error rate than microsatellites and the data are comparable across labs without requiring a standardization process needed with microsatellites. Therefore, the application of genomics and identification of SNP panels for Wood Turtles could improve the genetic assignment success for forensic applications. If this route is pursued, expanding the reference collections to the entire range of the Wood Turtle would increase the assignment success and consider all potential sources for release of confiscated turtles.

Tissue Sampling Strategies for Future Genetic Collections and Monitoring

Tissue sampling for turtles is challenging and should consider the intrusion and stress to the turtle, genotyping success of the tissue type, experience of the collector and logistic difficulty in collecting the samples. Specifically, tissues that require the least handling with the highest success rates are desired when the study requires high numbers of samples with rapid processing in the lab (i.e. > 500 samples and < 12 months). Minimizing the different tissue types within a large study allows more streamlining in the lab for faster processing. Tail tips and toes were the most successful tissue type, followed by blood. Shell samples were reasonably successful and seemed to have higher consistency across samplers than toenails. In other words, shell samples seemed to require less information or experience by the collector whereas toenails had considerable variation across collectors. Specifically, some collectors provided multiple nails for small turtles which increased the successful extraction, some collectors cut nails deeper than others or the nails at certain sites were larger and provided more soft tissue. Overall, there are multiple tissue types that genotype successfully, and the selection of the tissue type for any future studies should consider these various factors. If a study desires high success rates in the lab and uses experienced collectors, then tail tips or blood would be preferred. However, if the study can tolerate some failed samples in the lab and/or uses inexperienced collectors then shell or toenail would be preferred. Sampling should be coordinated among collectors and the lab, and designed to best fit the questions and goals of the study.

Future Research

Connectivity and Movement – The extent and mechanisms related to connectivity among populations and associations with landscape (habitat) attributes needs more investigation to assist conservation planning. Demographic and movement studies should begin long-term efforts to identify individual, longer range movements. Additionally, the genetic data can be further investigated with a landscape genetics approach to examine correlations among the genetic data, landscape attributes and population demographics. This approach could explore possible habitat or population related correlates that may be associated with turtle movement among sites that could be important to identify areas where movement may be more critical to population dynamics. For example, the Potomac sites appear to support high movement among sites, whereas the northern Maine sites indicate no movement.

Evolution and Selection – Genomic studies to identify locations on the genome where selection or variation is occurring could inform conservation of the species as well as identify potential threats to Wood Turtles and other freshwater turtles (see Andrews et al. (2016) for conservation applications).

Genomic Sequencing – Single neucleotide polymorphism methods should be investigated for potential to identify finer-scale population structure. Panels of approximately 300 – 500 loci could be developed. Use

of these panels would increase the ability to differentiate population groups and would likely increase the success of genetic assignment.

Acknowledgments

Funding for this study was provided by a US Fish and Wildlife Service sponsored Competitive State Wildlife Grant to the Massachusetts Division of Fisheries and Wildlife and its partners, and by a Regional Conservation Needs Grant. M. Jones, L. Wiley, T. Akre, H. P. Roberts, provided direction and coordination for the study and comments on the draft report. Participants in the CSWG process and volunteers provided samples and associated data including: M. Jones, L. Wiley, D. Yorks, M. Marchand, J. Megyesy, T. Akre, E. Ruther, N. Gerard, N. Sherwood, H. P. Roberts, L. Johnson, E. Thompson, V. Young, B. Zarate, K. Gipe, and A. Rubillard. Collectors who contributed more than 30 samples (and not listed above) were: V. Young, C. Eichelberger, N. Girard, L. Johnson, and N. Sherwood. Numerous other collectors and volunteers contributed samples to this project. S. Painter and A. Lodmell provided assistance in the lab. Tissue samples from Iowa used as an out-group shown in Appendix F were collected by J. Tamplin.

All turtles were collected and handled according to designated scientific collection permits and followed ethical treatment and care under policies for each agency or institution.

Citations

Allendorf, F.W. and G. Luikart. 2007. Conservation and the Genetics of Populations. Blackwell Scientific Publishing; Malden, Massachusetts.

Allendorf, F.W., G. Luikart, and S. Aitken. 2013. Conservation and the Genetics of Populations. Blackwell Scientific Publishing, Malden, Massachusetts.

Amato, M.L., R.J. Brooks, and J. Fu. 2008. A phylogeographic analysis of populations of the Wood Turtle (*Glyptemys insculpta*) throughout its range. Molecular Ecology 17:570–581.

Andrews, K.R., J. M. Good, M. R. Miller, G. Luikart and P.A. Hohenlohe. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics 17: 81-92.

Blair, C., D. Weigel, M. Balazik, A. Keeley, F. Walker, E. Landguth, S. Cushman, M. Murphy, L. Waits and N. Balkenhol. 2012. A simulation-based evaluation of methods for inferring linear barriers to gene flow. Molecular Ecology Resources 12:822-833.

Castellano, C.M., J. L. Behler, and G. Amato. 2009. Genetic diversity and population genetic structure of the Wood Turtle (*Glyptemys insculpta*) at Delaware Water Gap National Recreation Area, USA. Conservation Genetics 10:1783–1788.

Earl, D. A. and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources vol. 4 (2) pp. 359–361.

Escalona, T., T.N. EnNJArom, O.E. Hernandez, B.C. Bock, R. C. Vogt and N. Valenzuela. 2009. Population genetics of the endangered South American freshwater turtle, *Podocnemis unifilis*, inferred from microsatellite DNA data. Conservation Genetics 10:1683–1696.

Faubet, P. and O. E. Gaggiotti. 2008. A new Bayesian method to identify the environmental factors that influence recent migration. Genetics 178:1491-1504.

Faubet, P., R. Waples and O. E. Gaggiotti. 2007. Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. Molecular Ecology 16:1149-116.

Fridgen, C., L. Finnegan, C. Reaume, J. Cebek, J. Trottier, and P.J. Wilson. 2013. Conservation genetics of Wood Turtle (*Glyptemys insculpta*) populations in Ontario, Canada. Herpetological Conservation and Biology 8(2):351–358.

Gaur, A., A. Reddy, S. Annapoorni, B. Satyarebala, and S. Shivaji. 2006. The origin of Indian Star tortoises (*Geochelone elegans*) based on nuclear and mitochondrial DNA analysis: A story of rescue and repatriation. Conservation Genetics 2006:231–240.

Goudet, J., M. Raymond, T deMeeus, and F. Rousset. 1996. Testing differentiation in diploid populations. Genetics 144(4):1933–1940.

Howeth, J.G., S.E. McGaugh and D. A. Hendrickson. 2008. Contrasting demographic and genetic estimates of dispersal in the endangered Coahuilan box turtle: a contemporary approach to conservation. Molecular Ecology 17:4209–4221.

Jensen, J.L., A. J. Bohonak, and S. T. Kelley. 2005. Isolation by distance, web service. BMC Genetics 6: 13. v.3.23 http://ibdws.sdsu.edu/

Jones, M.T. and L.L. Willey. 2015. Status and Conservation of the Wood Turtle in the Northeastern United States. Final report to the Northeast Association of Fish and Wildlife Agencies (NEAFWA) for a Regional Conservation Needs (RCN) award.

Jones, M. T. and P. Seivert. 2009. Effects of stochastic flood disturbance on adult Wood Turtles, Glyptemys insculpta, in Massachusetts. Canadian Field-Naturalist 123:313-322.

Kalinowski, S.T., A.P. Wagner, and M.L. Taper. 2006. ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes 6:576–579.

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

Kalinowski, S.T. 2011. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity 106(4):625–632.

Kimura, M. and G.H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561–576.

King, T.L. and S.E. Julian. 2004. Conservation of microsatellite DNA flanking sequence across 13 Emydid genera assayed with novel Bog Turtle (*Glyptemys muhlenbergii*) loci. Conservation Genetics 5:719–725.

Libants, S., A.M. Kamarainen, K.T. Scribner, and J.D. Congdon. 2004. Isolation and cross-species amplification of seven microsatellite loci from *Emydoidea blandingii*. Molecular Ecology Notes 4:300–302.

Lowe, W.H. and F.W. Allendorf. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19:3038–3051.

Lutterschmidt, W.I., J.C. Cureton, and A.L. Gaillard. 2010. "Quick" DNA extraction from claw clippings of fresh and formalin-fixed Box Turtle (*Terrapene ornata*) specimens. Herpetological Review 41(3):313–315.

Malenfant, R. M., D. W. Coltman and C. S. Davis. 2015. Design of a 9K illumine BeadChip for polar bears (*Ursus maritimus*) from RAD and transcriptome sequencing. Molecular Ecology Resources 15:587-600.

Manel, S., O.E. Gaggiotti, and R.S. Waples. 2005. Assignment methods: matching biological questions with appropriate techniques. Trends in Ecology and Evolution 20:136–142.

Meirmans, P.G. and P.H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms, Molecular Ecology Notes. 4:792–794.

Meirmans, P.G. 2014. Nonconvergence in Bayesian estimation of migration rates. Molecular Ecology Resources 14:726-733.

National Marine Fisheries Service (NMFS). 2018. http://www.westcoast.fisheries.noaa.gov/maps_data/Species_Maps_Data.html. Last accessed January 2018.

Osentoski, M.F., S. Mockford, J.M. Wright, M. Snyder, B. Herman, and C.R. Hughes. 2002. Isolation and characterization of microsatellite loci from the Blanding's turtle, *Emydoidea blandingii*. Molecular Ecology Notes 2:147–149.

Palsball, P.J., M. Berube and F.W. Allendorf. 2007. Identification of management units using population genetic data. Trends in Ecology and Evolution 22:11-16.

Pearse, D.E., F.J. Janzen and J.C. Avise. 2001. Genetic markers substantiate long-term storage and utilization of sperm by female painted turtles. Heredity 2001:378–384.

Paetkau, D., W. Calvert, I. Stirling and C. Strobeck. 1995. Microsatellite analysis of population structure in Canadian polar bears. Molecular Ecology 4:347–354.

Paetkau, D., R. Slade, M. Burden, and A. Estoup. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation based exploration of accuracy and power. Molecular Ecology 13:55-65.

Peakall, R. and P.E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics *2*8, 2537–2539.

Peakall, R. and P.E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes. 6, 288–295.

Piry, S., A. Alapetite, J.-M. Cornuet, D. Paetkau, L. Baudouin, A. Estoup. 2004. GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. Journal of Heredity 95:536-539.

Prichard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Puechmaille, S.J. 2016. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Molecular Ecology Resources 16:608–627.

R Corps Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raymond, M. and F. Rousset. 1995. Genepop Version 1.2: population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248–249.

Rice W.R. 1989. Analyzing tables of statistical test. Evolution, 43: 223–225.

Sethuraman, A., and 12 other authors. 2014. Population genetics of Blanding's turtle (Emys blandingii) in the Midwestern United States. Conservation Genetics 15:61–73.

Shoemaker, K.T. and J.P. Gibbs. 2013. Genetic connectivity among populations of the threatened Bog Turtle (*Glyptemys muhlenbergii*) and the need for a regional approach to turtle conservation. Copeia 2013:324–331.

Spradling, T.A., J.W. Tamplin, S.S. Dow, K.J. Meyer. 2010. Conservation genetics of a peripherally isolated population of the Wood Turtle (*Glyptemys insculpta*) in Iowa. Conservation Genetics 11:1667–1677.

Tessier, N., S.R. Paquette, and F.J. Lapointe. 2005. Conservation genetics of the Wood Turtle (*Glyptemys insculpta*) in Quebec, Canada. Canadian Journal of Zoology 83:765–772.

Wang, J. L. 2004. Sibship reconstruction from genetic data with typing errors. Genetics. 166:1963-1979.

Washington Department of Fish and Wildlife (WDFW). 2018. https://fortress.wa.gov/dfw/score/score/species/bulltrout.jsp?species=Bulltrout Last accessed January 2018.

Whiteley, A. R., J. A. Coombs, B. H. Letcher and K. H. Nislow. 2014. Simulation and empirical analysis of novel sibship-based genetic determination of fish passage. Canadian Journal of Fisheries and Aquatic Sciences 71:1667-1679.

Willoughby, J.R., M. Sudaram, T.L. Lewis, and B.J. Swanson. 2013. Population decline in a long-lived species: the Wood Turtle in Michigan. Herpetologica 69(2):186–198.

Appendices

Appendix A. Genetic Diversity Measures over all loci by site.

Site	n	AR	PA	Ho	He	Fis
CT Wh	9	5.13	0.03	0.60	0.60	-0.04
MAWor	23	4.34	0.03	0.55	0.56	0.00
MACas	5	3.94	0	0.66	0.64	-0.15
MACros	21	5.1	0.04	0.61	0.62	-0.04
MABum	19	4.99	0	0.56	0.62	0.08
MAChar	11	5.49	0.05	0.66	0.65	-0.07
MAWild	11	5.25	0.05	0.55	0.63	0.07
MALiBear	16	5.38	0.1	0.66	0.67	0.00
MDWolf	22	5.72	0	0.66	0.66	-0.02
MDMary	8	5.44	0.06	0.64	0.65	-0.08
MDPumF	18	5	0	0.55	0.58	0.04
MDMM	18	5.53	0	0.61	0.64	0.00
MDTom	23	5.57	0	0.62	0.66	0.06
MEArCoy	18	4.33	0.05	0.50	0.53	0.01
MEArF	25	4.85	0.11	0.55	0.60	0.12
MEArTL	41	4.78	0.05	0.56	0.59	0.02
MEArCol	24	4.77	0.01	0.60	0.60	0.05
MEArY	21	4.83	0.18	0.57	0.57	-0.03
MEBigCy	12	5.1	0.04	0.58	0.60	0.01
MEMon	12	3.97	0	0.56	0.57	-0.06
MEArTB	25	4.49	0.04	0.59	0.60	0.01
MECamH	20	5.1	0.01	0.63	0.62	-0.05
MERoarL	7	3.38	0.05	0.62	0.53	-0.21
MESm	23	4.97	0.09	0.59	0.61	0.01
METan	17	4.78	0	0.56	0.60	0.03
MEBax	21	4.51	0.03	0.56	0.59	0.06
NHTur	25	5.16	0.01	0.64	0.65	-0.01

NHCrow	7	4.44	0.13	0.57	0.63	0.02
NHBull	13	4.57	0.08	0.52	0.55	0.04
NHDLiz	8	5	0.24	0.64	0.67	0.01
NHArCu	12	4.92	0.01	0.57	0.58	-0.03
NHFor	27	4.82	0.03	0.62	0.62	0.00
NHArB	20	5.01	0.02	0.61	0.63	-0.02
NHYoo	10	5.19	0	0.62	0.63	-0.08
NHFlood	28	5.46	0.05	0.65	0.66	-0.02
NHSour	8	4.94	0.04	0.60	0.60	-0.03
NHCyc	17	5.47	0.02	0.63	0.66	-0.01
NHPick	7	5	0	0.62	0.62	-0.09
NHMills	8	4.63	0.04	0.54	0.62	0.10
NJPot	29	5.75	0.06	0.65	0.65	-0.05
NJBar	9	5.75	0.06	0.65	0.65	-0.08
NJSu	20	5.38	0.05	0.64	0.66	-0.02
NJJack	13	5.61	0.14	0.58	0.61	0.03
NJWill	5	3.63	0	0.60	0.59	-0.13
NJBull	14	5.61	0.07	0.58	0.62	0.06
NYBearV	11	5.61	0.04	0.57	0.64	0.04
NYBar	12	5.25	0.03	0.57	0.63	0.05
NYYan	10	4.91	0.07	0.63	0.65	-0.03
PASno	17	5.78	0.19	0.62	0.69	0.07
PAMy	27	5.75	0.06	0.55	0.65	0.14
PANan	12	5.97	0.06	0.60	0.70	0.10
PACor	23	6.2	0.1	0.55	0.67	0.13
RI	18	3.95	0	0.56	0.56	-0.04
VAStS	21	5.42	0.05	0.65	0.66	-0.01
VASil	21	5.53	0	0.65	0.66	0.00
VABox	23	5.6	0.05	0.70	0.67	-0.08

VAWatF	13	5.57	0.02	0.61	0.65	-0.01
VAAug	19	5.3	0.03	0.64	0.64	-0.04
VAChick	17	5.06	0.04	0.64	0.62	-0.06
VAHid	12	5.72	0.29	0.67	0.68	-0.02
VADiv	50	5.46	0.07	0.61	0.66	0.04
WV	25	5.45	0.08	0.66	0.66	-0.02
max	50.000	6.200	0.290	0.700	0.702	0.142
min	5.000	3.380	0.000	0.502	0.532	-0.213
avg	17.44	5.09	0.05	0.60	0.63	0.00

Appendix B. Individual assignments for unknown samples. Samples with an asterisk indicate no clear assignment.

Sample	rank 1	score %	rank 2	score %	rank 3	score %
/FWS1	MDMary	67.712	PANan	19.653	VAAug	5.992
/FWS2	WV	46.211	MAWor	16.316	PANan	15.657
/FWS3	METan	38.538	NHFor	8.782	NHYoo	8.306
/FWS4	MECamH	97.18	MEBigC	1.577	NYBar	0.393
/FWS5	PANan	70.292	NHCy	26.387	WV	1.225
/FWS6**	MEArF	28.242	PASno	25.696	NYBar	11.205
/FWS7	VAWat	47.691	WV	35.786	NJSu	10.448
/FWS8	VAWat	53.628	MDMM	40.937	WV	4.383
/FWS9	WV	46.123	MDPumF	32.424	PASno	12.815
/FWS10	WV	93.029	VAStS	1.567	MDMM	1.338
/FWS11	MDMary	34.66	NHFor	31.366	MEArF	18.451
/FWS12	WV	42.145	VASil	16.272	VAStS	12.17
/FWS13	WV	52.775	MDMM	27.008	PANan	8.081
/FWS14	NHCy	57.357	NJSu	13.09	NHYoo	10.396
/FWS15	MAWild	44.313	NHArB	20.654	MEArTL	10.156
/FWS16	NHCy	96.223	MDMary	1.085	MEArTL	0.403
/NJA1	NYBar	53.555	VAWatF	13.65	MDMM	10.526
/NJA2	NJSu	87.497	NHCy	4.889	MEBax	2.082
/NJA3	MAWild	64.087	CT	27.95	NHYoo	2.373
/NJA4	NHCy	73.922	NJSu	5.881	NJBull	5.002
/NJA5**	NHDLi	17.923	WV	15.72	NJSu	12.637
/NJA6	MDMary	93.87	WV	2.133	NJSu	1.606
/NJA7	MEArB	49.144	MEArTL	13.848	MECamH	9.764
/NJA8	MACros	93.674	NJSu	2.132	NJPot	1.518
/NJA9	VAHid	44.634	VAAug	27.45	MDMM	12.006
/NJA10	NHCy	71.454	PACor	10.536	NHYoo	6.578
/NJA11	MDMary	15.037	MDPumF	12.321	VAAug	12.236
/NJA12	MDMM	80.921	PACor	4.719	MACros	2.973
/NJA13**	MDMary	47.858	NHDLi	46.809	VAChick	2.843
/NJA14	PANan	42.609	METan	25.082	NHCy	18.843
/NJA15	WV	29.134	MDMary	21.61	VAHid	16.984
/NJA16	NHCy	55.893	NHFlood	16.666	MAWor	7.509
/NJA17**	MAChar	26.549	METan	22.292	PANan	15.762
/NJA18	NJSu	99.376	PACor	0.141	MEArY	0.136
/NJA19	MDMary	50.943	WV	21.284	VAHid	13.142
/NJA20	MACros	73.021	MABum	9.093	MALiBea	3.884
/NJA21	MACros	31.181	NYBar	15.012	MESm	11.159
/NJB1**	NJBull	38.763	WV	35.148	MEArF	9.835
/NJB2	NHCy	59.081	PANan	33.134	MECamH	3.125
/NJB3	MACros	37.54	PACor	37.246	VAHid	11.476

/NJB4	PANan	84.344	NJSu	7.297	NHCy	2.862
/NJB5	NHCy	55.983	NJBull	32.828	NHYoo	2.281
/NJB6	MDMary	63.196	MDMM	31.848	NJSu	2.366
/NJB7	NJBar	44.903	NJSu	25.406	MESm	15.803
/NJB8	MDTom	13.462	NJSu	11.718	PANan	9.737
/NJB9	MEArTL	85.284	MEArCol	10.162	NHCy	1.658
/NJB10	MDMary	84.402	WV	5.757	PANan	3.402
/NJB11	MDMM	92.552	VAAug	5.645	NJSu	1.366
/NJB12	MEArTL	84.925	MEBigCy	9.016	MEArCol	1.399
/NJB13	NHCy	81.658	MEBax	2.254	MEArCol	2.05
/NJB14	MAChar	20.674	MEArF	18.846	CT	12.406
/NJB15	MDMary	99.869	MDMM	0.077	NJSu	0.035
/NJB16	WV	32.695	MDMM	29.823	VAHid	16.162
/NJB17	WV	32.656	PANan	26.56	VAWat	24.746
/NJB18	MDMary	59.494	NJSu	13.922	NHCy	13.48
/MA690	MDMary	35.617	MDMM	34.649	VAWat	21.06

Appendix C. F_{ST} values for all sites included in the analyses. Note that sample sizes vary from n=5 to 50 (shown in Appendix A). Asterisks indicate pairwise allele frequency exact tests that are not significant.

pop	CT Wh	MA Wor	MA Cas	MA Cros	MA Bum	MA Char	MA Wild	MA Lbear	MD Wolf	MD Mary	MD Pum	MD MM	MD Tom	MEArCoy	MEArF	MEArTL	MEArCol	MEArY
MA Wor	0.096																	
MA Cas	0.087	0.082																
MA Cros	0.043	0.076	0.042															
MA Bum	0.059	0.062	0.013	0.031														
MA Char	0.048*	0.058	0.012	0.028	0.000*													
MA Wild	0.036	0.076	0.033	0.029	0.013*	0.015*												
MA Lbear	0.040	0.071	0.017	0.026	0.014*	0.000*	0.014*											
MD Wolf	0.065	0.117	0.091	0.069	0.076	0.061	0.033	0.062										
MD Mary	0.062	0.099	0.086	0.052	0.055	0.052	0.018*	0.047	0.008*									
MD Pum	0.088	0.134	0.148	0.097	0.111	0.107	0.058	0.103	0.027	0.038								
MD MM	0.080	0.120	0.106	0.078	0.087	0.066	0.030	0.067	0.001*	0.010*	0.022							
MD Tom	0.080	0.137	0.116	0.082	0.099	0.081	0.058	0.082	0.005*	0.023*	0.041	0.007*						
MEArCoy	0.080	0.094	0.109	0.063	0.077	0.082	0.074	0.072	0.118	0.120	0.131	0.118	0.129					
MEArF	0.083	0.086	0.073	0.073	0.053	0.049	0.056	0.061	0.097	0.077	0.115	0.095	0.108	0.064				
MEArTL	0.091	0.097	0.080	0.062	0.051	0.048	0.049	0.047	0.096	0.087	0.124	0.096	0.114	0.057	0.039			
MEArCol	0.068	0.080	0.054	0.045	0.031	0.028	0.050	0.034	0.093	0.085	0.121	0.092	0.108	0.036	0.031	0.045		
MEArY	0.065	0.087	0.116	0.059	0.074	0.072	0.045	0.059	0.094	0.080	0.113	0.090	0.104	0.041	0.056	0.039	0.049	
MEBCy	0.111	0.092	0.088	0.095	0.075	0.083	0.072	0.073	0.122	0.108	0.144	0.120	0.133	0.085	0.063	0.066	0.069	0.064
MEMon	0.145	0.125	0.095	0.104	0.050	0.087	0.098	0.089	0.144	0.131	0.178	0.152	0.154	0.106	0.104	0.094	0.077	0.118
MEArTB	0.086	0.072	0.071	0.055	0.054	0.043	0.058	0.057	0.095	0.078	0.118	0.101	0.107	0.057	0.046	0.045	0.045	0.056
MECamH	0.097	0.064	0.076	0.064	0.045	0.055	0.046	0.056	0.099	0.065	0.115	0.097	0.106	0.058	0.050	0.051	0.050	0.049
MERLion	0.132	0.142	0.118	0.093	0.081	0.113	0.096	0.093	0.152	0.158	0.180	0.169	0.163	0.091	0.118	0.114	0.087	0.097
MESm	0.084	0.064	0.061	0.053	0.038	0.047	0.037	0.050	0.087	0.064	0.105	0.086	0.094	0.066	0.049	0.059	0.049	0.055
METan	0.095	0.080	0.058	0.055	0.035	0.024	0.050	0.042	0.087	0.085	0.130	0.090	0.107	0.057	0.040	0.036	0.032	0.059
MEBax	0.130	0.116	0.104	0.110	0.074	0.080	0.098	0.082	0.144	0.137	0.168	0.141	0.144	0.091	0.081	0.093	0.075	0.104
NH Tur	0.087	0.048	0.046	0.057	0.030	0.038	0.038	0.034	0.096	0.074	0.129	0.095	0.112	0.081	0.058	0.065	0.057	0.064
NH Crow	0.066	0.079	0.049	0.038	0.022*	0.044*	0.017*	0.036	0.068	0.050	0.079	0.071	0.076	0.050	0.057	0.062	0.043	0.037
NH Bull	0.076	0.091	0.073	0.046	0.042	0.059	0.050	0.062	0.107	0.100	0.128	0.117	0.125	0.082	0.098	0.097	0.071	0.087
NH Dliz	0.086	0.107	0.059	0.062	0.033	0.033*	0.052	0.039	0.095	0.080	0.139	0.101	0.102	0.094	0.084	0.071	0.049	0.084
NHArCu	0.091	0.093	0.082	0.046	0.042	0.045	0.052	0.048	0.102	0.093	0.139	0.103	0.115	0.052	0.048	0.024	0.034	0.044
NHFor	0.064	0.063	0.056	0.059	0.025	0.028	0.034	0.037	0.077	0.061	0.099	0.076	0.090	0.056	0.047	0.050	0.042	0.053
NHArB	0.062	0.082	0.034	0.052	0.027	0.035	0.034	0.032	0.082	0.061	0.107	0.084	0.096	0.069	0.069	0.068	0.050	0.071
NHYoo	0.093	0.083	0.063	0.073	0.051	0.062	0.057	0.053	0.109	0.083	0.131	0.108	0.114	0.069	0.053	0.063	0.051	0.057

pop	CT Wh	MA Wor	MA Cas	MA Cros	MA Bum	MA Char	MA Wild	MA Lbear	MD Wolf	MD Mary	MD Pum	MD MM	MD Tom	MEArCoy	MEArF	MEArTL	MEArCol	MEArY
NHFlood	0.055	0.067	0.031	0.051	0.030	0.030	0.024*	0.027	0.075	0.059	0.099	0.074	0.087	0.058	0.042	0.055	0.041	0.049
NHSour	0.081	0.072	0.043	0.062	0.024*	0.030*	0.017*	0.033*	0.088	0.062	0.110	0.072	0.102	0.089	0.077	0.082	0.060	0.076
NHCy	0.060	0.050	0.029	0.052	0.024	0.036	0.047	0.036	0.094	0.078	0.123	0.105	0.108	0.070	0.057	0.066	0.039	0.068
NH Pic	0.076*	0.072	0.048	0.059	0.03*	0.033*	0.021*	0.030*	0.085	0.064*	0.120	0.078	0.093	0.067	0.053	0.049	0.045	0.051
NHMill	0.070*	0.064	0.022	0.054	0.009*	0.022*	0.043*	0.028	0.094	0.080	0.134	0.107	0.109	0.099	0.069	0.081	0.048	0.090
NJPot	0.077	0.130	0.087	0.061	0.078	0.076	0.059	0.062	0.061	0.046	0.092	0.076	0.077	0.120	0.099	0.099	0.091	0.100
NJBar	0.054	0.115	0.081	0.051	0.048	0.035*	0.018	0.033	0.028	0.014*	0.064	0.040	0.056	0.107	0.079	0.068	0.075	0.077
NJSu	0.082	0.113	0.085	0.065	0.060	0.065	0.061	0.048	0.058	0.040	0.090	0.071	0.075	0.099	0.075	0.067	0.066	0.083
NJJack	0.097	0.147	0.112	0.077	0.080	0.083	0.036	0.072	0.052	0.044*	0.079	0.062	0.067	0.128	0.115	0.115	0.109	0.104
NJWill	0.112	0.133	0.145	0.096	0.098	0.090	0.079	0.095	0.071	0.081	0.096	0.084	0.084	0.170	0.117	0.136	0.120	0.147
NJBull	0.056	0.121	0.081	0.066	0.077	0.059	0.038	0.053	0.055	0.039	0.076	0.052	0.065	0.097	0.091	0.088	0.086	0.085
NYBearV	0.044	0.095	0.063	0.039	0.042	0.036	0.021*	0.041	0.050	0.043	0.079	0.056	0.069	0.099	0.083	0.069	0.069	0.071
NYBarr	0.063	0.117	0.062	0.058	0.052	0.057	0.049	0.046	0.070	0.049	0.083	0.066	0.070	0.096	0.060	0.065	0.063	0.074
NYYan	0.053	0.092	0.056	0.049	0.044	0.036	0.029	0.035	0.060	0.047	0.081	0.064	0.067	0.100	0.078	0.082	0.067	0.077
PASno	0.059	0.099	0.072	0.064	0.078	0.062	0.040	0.057	0.005	0.017*	0.041	0.021	0.020	0.105	0.090	0.094	0.088	0.088
PAMy	0.059	0.108	0.056	0.054	0.050	0.034	0.026	0.041	0.042	0.023	0.081	0.038	0.050	0.105	0.066	0.069	0.066	0.085
PANan	0.049	0.091	0.060	0.053	0.052	0.040	0.020	0.028	0.025	0.002*	0.054	0.027	0.034	0.091	0.070	0.077	0.070	0.062
PACor	0.047	0.086	0.029	0.038	0.037	0.024	0.019	0.033	0.034	0.017*	0.074	0.037	0.047	0.091	0.066	0.073	0.060	0.074
RI	0.111	0.012*	0.077	0.079	0.067	0.059	0.088	0.081	0.125	0.106	0.145	0.129	0.142	0.131	0.116	0.126	0.092	0.124
VAStS	0.084	0.129	0.106	0.075	0.079	0.068	0.045	0.071	0.013*	0.015*	0.043	0.012*	0.004*	0.127	0.107	0.112	0.099	0.110
VASil	0.075	0.118	0.095	0.070	0.079	0.065	0.037	0.070	0.005*	0.014*	0.042	0.013*	0.015*	0.113	0.100	0.100	0.094	0.099
VABox	0.080	0.122	0.103	0.083	0.089	0.077	0.046	0.076	0.005*	0.023*	0.039	0.019*	0.006*	0.129	0.112	0.114	0.106	0.108
VAWat	0.069	0.134	0.102	0.080	0.078	0.066	0.037	0.064	0.008*	0.001*	0.023*	0.001*	0.010*	0.126	0.103	0.104	0.098	0.103
VAAug	0.120	0.132	0.108	0.088	0.095	0.073	0.064	0.079	0.031	0.039	0.084	0.024*	0.031	0.142	0.110	0.115	0.108	0.120
VAChick	0.134	0.167	0.142	0.123	0.126	0.104	0.096	0.105	0.049	0.055	0.092	0.045	0.046	0.179	0.141	0.147	0.132	0.150
VAHid	0.088	0.127	0.093	0.083	0.078	0.062	0.039	0.059	0.013*	0.029*	0.053	0.019*	0.021*	0.133	0.108	0.103	0.104	0.115
VADiv	0.081	0.116	0.102	0.078	0.080	0.071	0.043	0.073	0.010	0.014*	0.038	0.019	0.018	0.123	0.099	0.101	0.101	0.104
WVLT	0.087	0.123	0.107	0.084	0.083	0.076	0.046	0.079	0.015	0.022*	0.046	0.021	0.017	0.127	0.100	0.111	0.104	0.109

pop	CT Wh	MA Wor	MA Cas	MA Cros	MA Bum	MA Char	MA Wild	MA Lbear	MD Wolf	MD Mary	MD Pum	MD MM	MD Tom	MEArCoy	MEArF	MEArTL	MEArCol	MEArY
NHFlood	0.055	0.067	0.031	0.051	0.030	0.030	0.024*	0.027	0.075	0.059	0.099	0.074	0.087	0.058	0.042	0.055	0.041	0.049
NHSour	0.081	0.072	0.043	0.062	0.024*	0.030*	0.017*	0.033*	0.088	0.062	0.110	0.072	0.102	0.089	0.077	0.082	0.060	0.076
NHCy	0.060	0.050	0.029	0.052	0.024	0.036	0.047	0.036	0.094	0.078	0.123	0.105	0.108	0.070	0.057	0.066	0.039	0.068
NH Pic	0.076*	0.072	0.048	0.059	0.03*	0.033*	0.021*	0.030*	0.085	0.064*	0.120	0.078	0.093	0.067	0.053	0.049	0.045	0.051
NHMill	0.070*	0.064	0.022	0.054	0.009*	0.022*	0.043*	0.028	0.094	0.080	0.134	0.107	0.109	0.099	0.069	0.081	0.048	0.090
NJPot	0.077	0.130	0.087	0.061	0.078	0.076	0.059	0.062	0.061	0.046	0.092	0.076	0.077	0.120	0.099	0.099	0.091	0.100
NJBar	0.054	0.115	0.081	0.051	0.048	0.035*	0.018	0.033	0.028	0.014*	0.064	0.040	0.056	0.107	0.079	0.068	0.075	0.077
NJSu	0.082	0.113	0.085	0.065	0.060	0.065	0.061	0.048	0.058	0.040	0.090	0.071	0.075	0.099	0.075	0.067	0.066	0.083
NJJack	0.097	0.147	0.112	0.077	0.080	0.083	0.036	0.072	0.052	0.044*	0.079	0.062	0.067	0.128	0.115	0.115	0.109	0.104
NJWill	0.112	0.133	0.145	0.096	0.098	0.090	0.079	0.095	0.071	0.081	0.096	0.084	0.084	0.170	0.117	0.136	0.120	0.147
NJBull	0.056	0.121	0.081	0.066	0.077	0.059	0.038	0.053	0.055	0.039	0.076	0.052	0.065	0.097	0.091	0.088	0.086	0.085
NYBearV	0.044	0.095	0.063	0.039	0.042	0.036	0.021*	0.041	0.050	0.043	0.079	0.056	0.069	0.099	0.083	0.069	0.069	0.071
NYBarr	0.063	0.117	0.062	0.058	0.052	0.057	0.049	0.046	0.070	0.049	0.083	0.066	0.070	0.096	0.060	0.065	0.063	0.074
NYYan	0.053	0.092	0.056	0.049	0.044	0.036	0.029	0.035	0.060	0.047	0.081	0.064	0.067	0.100	0.078	0.082	0.067	0.077
PASno	0.059	0.099	0.072	0.064	0.078	0.062	0.040	0.057	0.005	0.017*	0.041	0.021	0.020	0.105	0.090	0.094	0.088	0.088
PAMy	0.059	0.108	0.056	0.054	0.050	0.034	0.026	0.041	0.042	0.023	0.081	0.038	0.050	0.105	0.066	0.069	0.066	0.085
PANan	0.049	0.091	0.060	0.053	0.052	0.040	0.020	0.028	0.025	0.002*	0.054	0.027	0.034	0.091	0.070	0.077	0.070	0.062
PACor	0.047	0.086	0.029	0.038	0.037	0.024	0.019	0.033	0.034	0.017*	0.074	0.037	0.047	0.091	0.066	0.073	0.060	0.074
RI	0.111	0.012*	0.077	0.079	0.067	0.059	0.088	0.081	0.125	0.106	0.145	0.129	0.142	0.131	0.116	0.126	0.092	0.124
VAStS	0.084	0.129	0.106	0.075	0.079	0.068	0.045	0.071	0.013*	0.015*	0.043	0.012*	0.004*	0.127	0.107	0.112	0.099	0.110
VASil	0.075	0.118	0.095	0.070	0.079	0.065	0.037	0.070	0.005*	0.014*	0.042	0.013*	0.015*	0.113	0.100	0.100	0.094	0.099
VABox	0.080	0.122	0.103	0.083	0.089	0.077	0.046	0.076	0.005*	0.023*	0.039	0.019*	0.006*	0.129	0.112	0.114	0.106	0.108
VAWat	0.069	0.134	0.102	0.080	0.078	0.066	0.037	0.064	0.008*	0.001*	0.023*	0.001*	0.010*	0.126	0.103	0.104	0.098	0.103
VAAug	0.120	0.132	0.108	0.088	0.095	0.073	0.064	0.079	0.031	0.039	0.084	0.024*	0.031	0.142	0.110	0.115	0.108	0.120
VAChick	0.134	0.167	0.142	0.123	0.126	0.104	0.096	0.105	0.049	0.055	0.092	0.045	0.046	0.179	0.141	0.147	0.132	0.150
VAHid	0.088	0.127	0.093	0.083	0.078	0.062	0.039	0.059	0.013*	0.029*	0.053	0.019*	0.021*	0.133	0.108	0.103	0.104	0.115
VADiv	0.081	0.116	0.102	0.078	0.080	0.071	0.043	0.073	0.010	0.014*	0.038	0.019	0.018	0.123	0.099	0.101	0.101	0.104
WVLT	0.087	0.123	0.107	0.084	0.083	0.076	0.046	0.079	0.015	0.022*	0.046	0.021	0.017	0.127	0.100	0.111	0.104	0.109

pop	MEBCy	MEMon	MEArTB	MECamH	MERLion	MESm	METan	MEBax	NH Tur	NH Crow	NH Bull	NH Dliz	NHArCu	NHFor	NHArB	NHYoo	NHFlood	NHSour
MA Wor																		
MA Cas																		
MA Cros																		
MA Bum																		
MA Char																		
MA Wild																		
MA Lbear																		
MD Wolf																		
MD Mary																		
MD Pum																		
MD MM																		
MD Tom																		
MEArCoy																		
MEArF																		
MEArTL																		
MEArCol																		
MEArY																		
MEBCy																		
MEMon	0.088																	
MEArTB	0.057	0.091																
MECamH	0.030*	0.073	0.038															
MERLion	0.109	0.130	0.110	0.097														
MESm	0.053	0.069	0.040	0.026	0.093													
METan	0.090	0.106	0.051	0.059	0.111	0.065												
MEBax	0.037	0.076	0.072	0.042	0.109	0.060	0.096											
NH Tur	0.044	0.069	0.056	0.030	0.089	0.035	0.063	0.065										
NH Crow	0.045*	0.067	0.037	0.017*	0.067	0.019*	0.055	0.054	0.035									
NH Bull	0.124	0.115	0.070	0.089	0.094	0.070	0.081	0.133	0.060	0.030*								
NH Dliz	0.069	0.065	0.047	0.060	0.094	0.045	0.060	0.064	0.060	0.017*	0.068							
NHArCu	0.057	0.079	0.039	0.051	0.112	0.058	0.026	0.085	0.057	0.037*	0.069	0.024*						
NHFor	0.058	0.086	0.039	0.038	0.094	0.033	0.053	0.071	0.036	0.024	0.052	0.048	0.048					
NHArB	0.072	0.071	0.049	0.038	0.079	0.033	0.070	0.065	0.033	0.006*	0.050	0.029	0.060	0.030				
NHYoo	0.012*	0.065	0.039	0.032	0.072	0.042	0.078	0.030	0.026*	0.001*	0.078	0.037*	0.042*	0.035	0.023*			

pop	MEBCy	MEMon	MEArTB	MECamH	MERLion	MESm	METan	MEBax	NH Tur	NH Crow	NH Bull	NH Dliz	NHArCu	NHFor	NHArB	NHYoo	NHFlood	NHSour
NHFlood	0.031	0.060	0.038	0.027	0.074	0.030	0.058	0.048	0.014*	0.014*	0.058	0.048	0.043	0.025	0.015	0.010*		
NHSour	0.058*	0.078	0.070	0.040*	0.088	0.032*	0.075	0.053	0.028*	0.008*	0.054*	0.049*	0.063	0.026	0.016*	0.018*	0.006*	
NHCy	0.038	0.056	0.046	0.040	0.074	0.042	0.061	0.057	0.015	0.028*	0.053	0.042	0.052	0.032	0.030	0.016*	0.012*	0.035*
NH Pic	0.011*	0.060	0.042	0.027*	0.084	0.028	0.065	0.030	0.011*	0.016	0.082	0.041*	0.046*	0.022	0.033*	0.002*	0.006*	0.013*
NHMill	0.061	0.096	0.063	0.045	0.084	0.040	0.080	0.069	0.020*	0.022*	0.061	0.050	0.083	0.025	0.017*	0.029*	0.020*	0.017*
NJPot	0.127	0.148	0.103	0.094	0.146	0.080	0.097	0.134	0.103	0.078	0.107	0.102	0.107	0.080	0.080	0.105	0.078	0.096
NJBar	0.114	0.144	0.084	0.083	0.140	0.074	0.062	0.138	0.082	0.062	0.091	0.084	0.083	0.048	0.071	0.096	0.063	0.065
NJSu	0.096	0.114	0.081	0.071	0.124	0.071	0.078	0.116	0.080	0.056	0.098	0.072	0.085	0.059	0.075	0.081	0.072	0.093
NJJack	0.131	0.153	0.105	0.096	0.157	0.086	0.104	0.127	0.114	0.068*	0.120	0.100	0.120	0.088	0.090	0.110	0.082	0.084
NJWill	0.175	0.191	0.132	0.142	0.230	0.123	0.125	0.198	0.118	0.107	0.126	0.139	0.159	0.099	0.120	0.152	0.111	0.142
NJBull	0.113	0.153	0.087	0.088	0.141	0.077	0.086	0.122	0.094	0.071	0.104	0.094	0.093	0.066	0.076	0.093	0.063	0.074
NYBearV	0.115	0.126	0.073	0.081	0.115	0.070	0.054	0.128	0.062	0.065	0.062	0.078	0.065	0.064	0.058	0.085	0.050	0.062
NYBarr	0.097	0.098	0.069	0.070	0.128	0.042	0.074	0.109	0.070	0.026*	0.075	0.051	0.068	0.051	0.040	0.061	0.046	0.070
NYYan	0.123	0.108	0.078	0.074	0.110	0.056	0.067	0.104	0.072	0.036	0.066	0.059	0.084	0.059	0.048	0.076	0.052	0.054
PASno	0.093	0.130	0.081	0.078	0.125	0.072	0.086	0.114	0.080	0.048	0.102	0.072	0.092	0.067	0.058	0.076	0.056	0.067
PAMy	0.087	0.117	0.078	0.074	0.130	0.066	0.069	0.107	0.072	0.064	0.093	0.064	0.067	0.059	0.067	0.069	0.052	0.048
PANan	0.076	0.103	0.066	0.057	0.111	0.058	0.063	0.089	0.053	0.031*	0.086	0.051	0.070	0.052	0.044	0.049	0.039	0.036*
PACor	0.086	0.102	0.060	0.062	0.106	0.051	0.059	0.094	0.057	0.041	0.072	0.054	0.070	0.051	0.043	0.065	0.038	0.032*
RI	0.138	0.142	0.096	0.098	0.171	0.095	0.109	0.149	0.067	0.094	0.100	0.104	0.119	0.087	0.083	0.106	0.086	0.080
VAStS	0.132	0.131	0.103	0.097	0.154	0.090	0.097	0.132	0.101	0.063	0.107	0.090	0.106	0.078	0.080	0.103	0.075	0.073
VASil	0.122	0.129	0.090	0.089	0.139	0.082	0.091	0.130	0.093	0.064	0.099	0.083	0.096	0.074	0.068	0.092	0.067	0.074
VABox	0.127	0.142	0.106	0.102	0.155	0.090	0.107	0.144	0.104	0.070	0.118	0.099	0.115	0.083	0.082	0.111	0.077	0.090
VAWat	0.128	0.148	0.091	0.094	0.165	0.079	0.109	0.134	0.104	0.050	0.119	0.090	0.114	0.072	0.063	0.100	0.072	0.080
VAAug	0.139	0.152	0.112	0.113	0.180	0.106	0.090	0.146	0.109	0.093	0.133	0.091	0.105	0.092	0.101	0.119	0.094	0.100
VAChick	0.170	0.196	0.139	0.143	0.209	0.135	0.134	0.176	0.151	0.134	0.175	0.146	0.157	0.119	0.132	0.155	0.121	0.141
VAHid	0.133	0.149	0.101	0.103	0.152	0.103	0.095	0.140	0.102	0.081	0.130	0.103	0.113	0.083	0.088	0.108	0.077	0.091
VADiv	0.123	0.133	0.095	0.089	0.150	0.080	0.103	0.129	0.099	0.071	0.113	0.098	0.110	0.078	0.075	0.104	0.073	0.081
WVLT	0.125	0.135	0.094	0.089	0.147	0.084	0.108	0.128	0.100	0.075	0.115	0.097	0.112	0.076	0.077	0.100	0.072	0.083

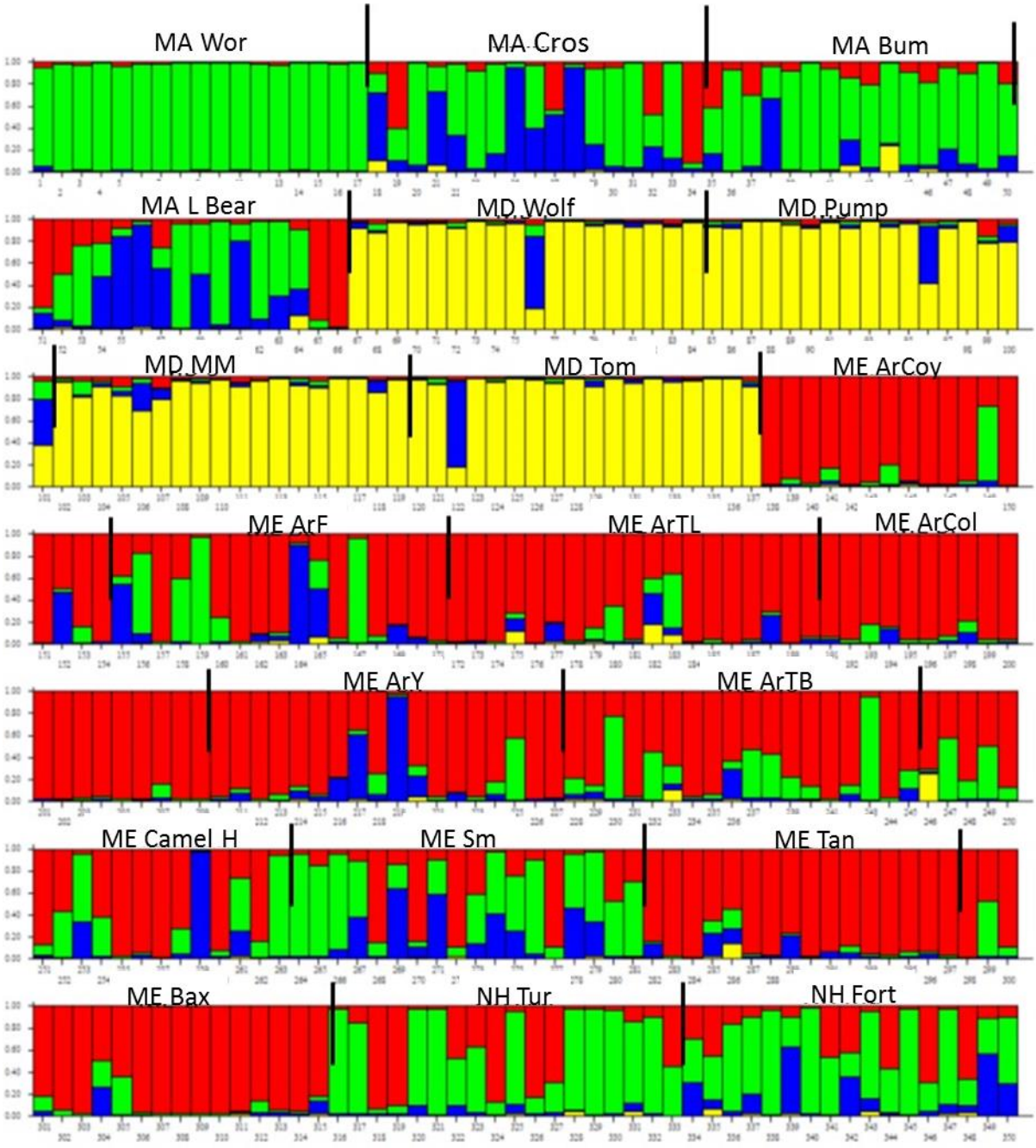
pop	NHCy	NH Pic	NHMill	NJPot	NJBar	NJSu	NJJack	NJWill	NJBull	NYBearV	NYBarr	NYYan	PASno	PAMy	PANan	PACor	RI	VAStS
MA Wor																		
MA Cas																		
MA Cros																		
MA Bum																		
MA Char																		
MA Wild																		
MA Lbear																		
MD Wolf																		
MD Mary																		
MD Pum																		
MD MM																		
MD Tom																		
MEArCoy																		
MEArF																		
MEArTL																		
MEArCol																		
MEArY																		
MEBCy																		
MEMon																		
MEArTB																		
MECamH																		
MERLion																		
MESm																		
METan																		
MEBax																		
NH Tur																		
NH Crow																		
NH Bull																		
NH Dliz																		
NHArCu																		
NHFor																		
NHArB																		
NHYoo																		

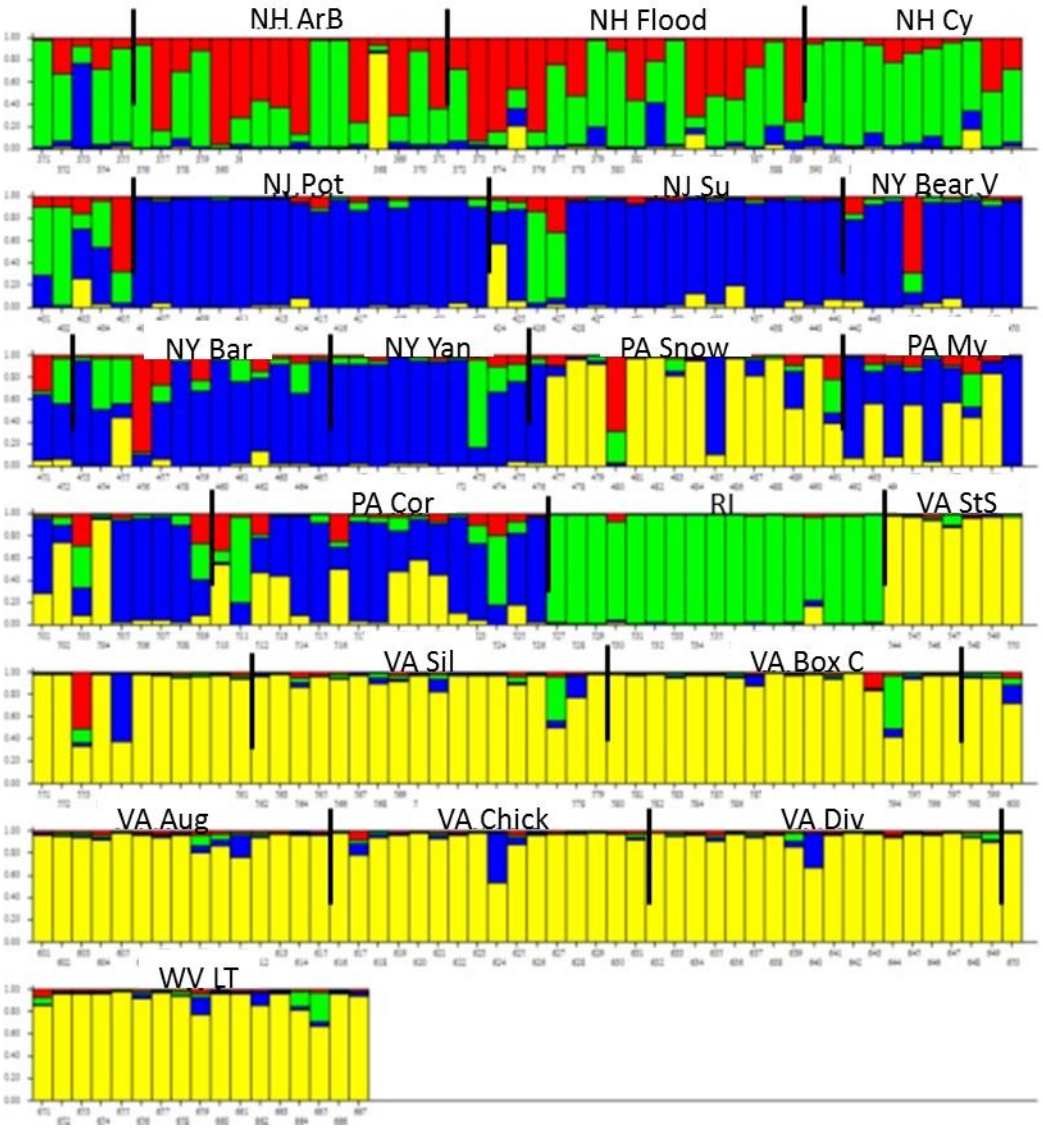
pop	NHCy	NH Pic	NHMill	NJPot	NJBar	NJSu	NJJack	NJWill	NJBull	NYBearV	NYBarr	NYYan	PASno	PAMy	PANan	PACor	RI	VAStS
NHFlood																		
NHSour																		
NHCy																		
NH Pic	0.016*																	
NHMill	0.010*	0.023*																
NJPot	0.089	0.092	0.086															
NJBar	0.087	0.074*	0.082	0.028*														
NJSu	0.070	0.076	0.071	0.033	0.020*													
NJJack	0.108	0.075	0.107	0.026*	0.032*	0.058												
NJWill	0.114	0.135	0.100	0.087	0.061	0.075	0.083											
NJBull	0.087	0.068	0.090	0.024	0.017*	0.052	0.034	0.109										
NYBearV	0.065	0.071	0.070	0.043	0.017*	0.060	0.049*	0.075	0.041									
NYBarr	0.060	0.064	0.067	0.056	0.051	0.041	0.078	0.090	0.056	0.035								
NYYan	0.066	0.060*	0.065	0.049	0.041	0.054	0.061	0.088	0.047	0.018	0.008*							
PASno	0.067	0.066	0.075	0.056	0.030	0.056	0.057	0.085	0.046	0.051	0.058	0.049						
PAMy	0.068	0.048	0.061	0.051	0.023	0.044	0.044	0.097	0.043	0.036	0.053	0.060	0.041					
PANan	0.052	0.038*	0.063	0.042	0.014*	0.037	0.038*	0.095	0.027	0.032	0.043	0.038	0.010*	0.027				
PACor	0.056	0.049	0.045	0.051	0.026	0.054	0.050	0.087	0.037	0.020	0.044	0.030	0.023	0.018	0.017*			
RI	0.069	0.100	0.067	0.138	0.127	0.123	0.163	0.128	0.133	0.109	0.122	0.095	0.111	0.111	0.107	0.089		
VAStS	0.099	0.083	0.103	0.073	0.049	0.078	0.059	0.093	0.066	0.065	0.069	0.056	0.025	0.048	0.027	0.043	0.131	
VASil	0.091	0.081	0.096	0.068	0.039	0.070	0.054	0.094	0.063	0.049	0.066	0.054	0.011	0.046	0.027	0.032	0.125	0.002*
VABox	0.094	0.093	0.106	0.075	0.051	0.079	0.072	0.090	0.071	0.071	0.075	0.061	0.012*	0.062	0.036	0.044	0.127	0.004*
VAWat	0.098	0.082	0.099	0.066	0.044	0.066	0.056	0.099	0.055	0.062	0.051	0.048	0.016	0.050	0.026*	0.035	0.138	0.000*
VAAug	0.118	0.100	0.118	0.086	0.052	0.076	0.078	0.108	0.069	0.085	0.091	0.077	0.031	0.054	0.038	0.044	0.139	0.026
VAChick	0.137	0.139	0.147	0.091	0.075	0.096	0.091	0.135	0.087	0.101	0.115	0.105	0.053	0.079	0.073	0.070	0.172	0.043
VAHid	0.101	0.091	0.100	0.074	0.036	0.073	0.053	0.096	0.061	0.054	0.078	0.057	0.013	0.041	0.035	0.032	0.131	0.010*
VADiv	0.097	0.089	0.101	0.066	0.041	0.072	0.055	0.086	0.065	0.057	0.065	0.056	0.015	0.055	0.038	0.038	0.124	0.008*
WVLT	0.096	0.093	0.102	0.069	0.043	0.073	0.052	0.097	0.060	0.057	0.071	0.060	0.016	0.049	0.033	0.036	0.132	0.008*

pop	VASil	VABox	VAWat	VAAug	VAChick	VAHid	VADiv	WVLT
MA Wor								
MA Cas								
MA Cros								
MA Bum								
MA Char								
MA Wild								
MA Lbear								
MD Wolf								
MD Mary								
MD Pum								
MD MM								
MD Tom								
MEArCoy								
MEArF								
MEArTL								
MEArCol								
MEArY								
MEBCy								
MEMon								
MEArTB								
MECamH								
MERLion								
MESm								
METan								
MEBax								
NH Tur								
NH Crow								
NH Bull								
NH Dliz								
NHArCu								
NHFor								
NHArB								
NHYoo								

pop	VASil	VABox	VAWat	VAAug	VAChick	VAHid	VADiv	WVLT
NHFlood								
NHSour								
NHCy								
NH Pic								
NHMill								
NJPot								
NJBar								
NJSu								
NJJack								
NJWill								
NJBull								
NYBearV								
NYBarr								
NYYan								
PASno								
PAMy								
PANan								
PACor								
RI								
VAStS								
VASil								
VABox	0.008*							
VAWat	0.006*	0.006*						
VAAug	0.032	0.039	0.041					
VAChick	0.046	0.045	0.029*	0.058				
VAHid	0.003*	0.016*	0.006*	0.032	0.033			
VADiv	0.004*	0.006*	0.000*	0.040	0.034	0.006*		
WVLT	0.000*	0.011*	0.004*	0.038	0.032	0.005*	0.000*	

Appendix D. Expanded STRUCTURE plot for K=4 (run 17) shown in Figure 5.




NH ArB
NH Flood
NH Cy
NJ Pot
NJ Su
NY Bear V
NY Bar
NY Yan
PA Snow
PA My
PA Cor
RI
VA StS
VA Sil
VA Box C
VA Aug
VA Chick
VA Div
WV LT

Appendix E. Sites tested for full sibling families and the full sibling families identified within sites.

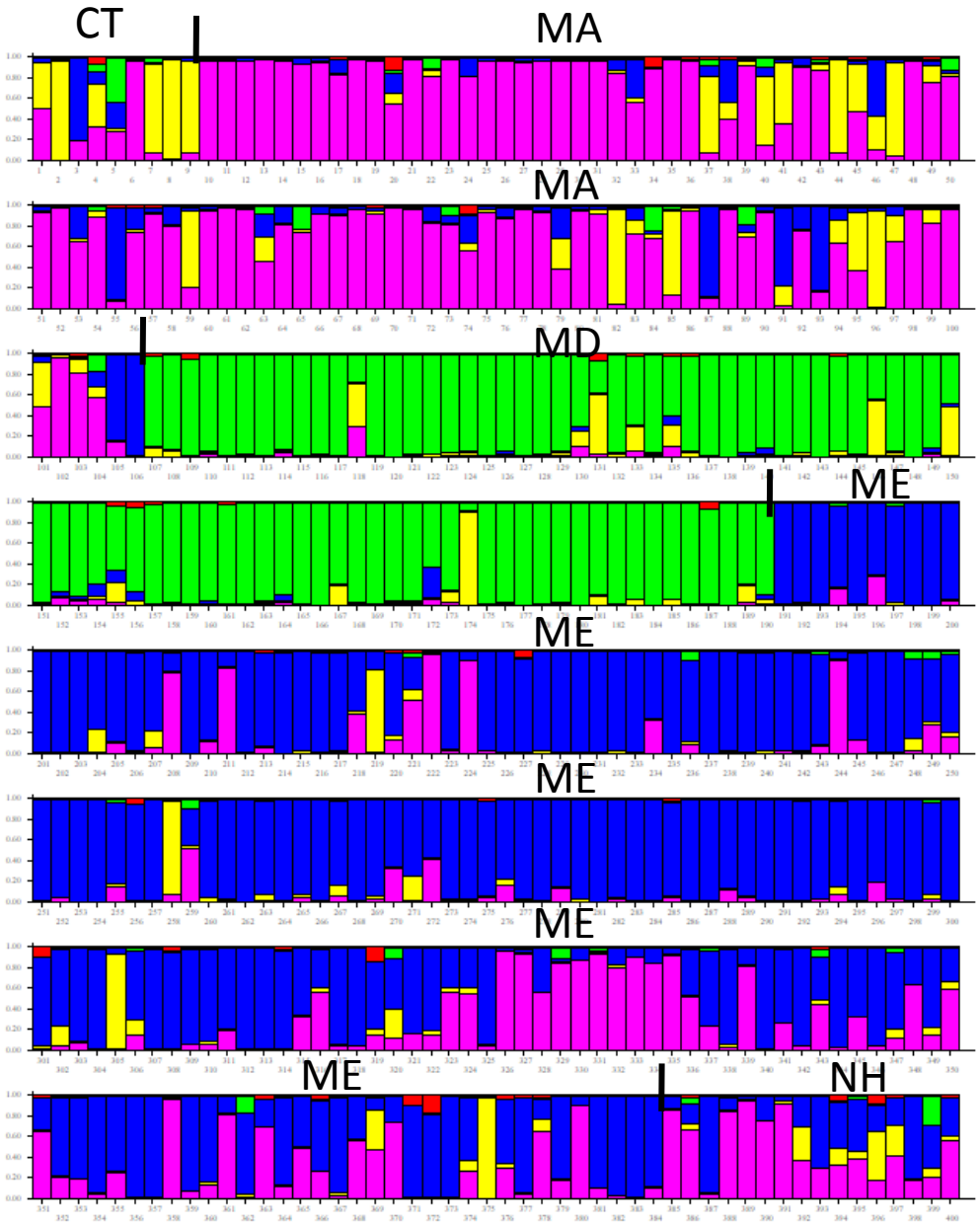
Table E.1. Sites used in the full sibling family tests. Sites included in each group are listed down the column.

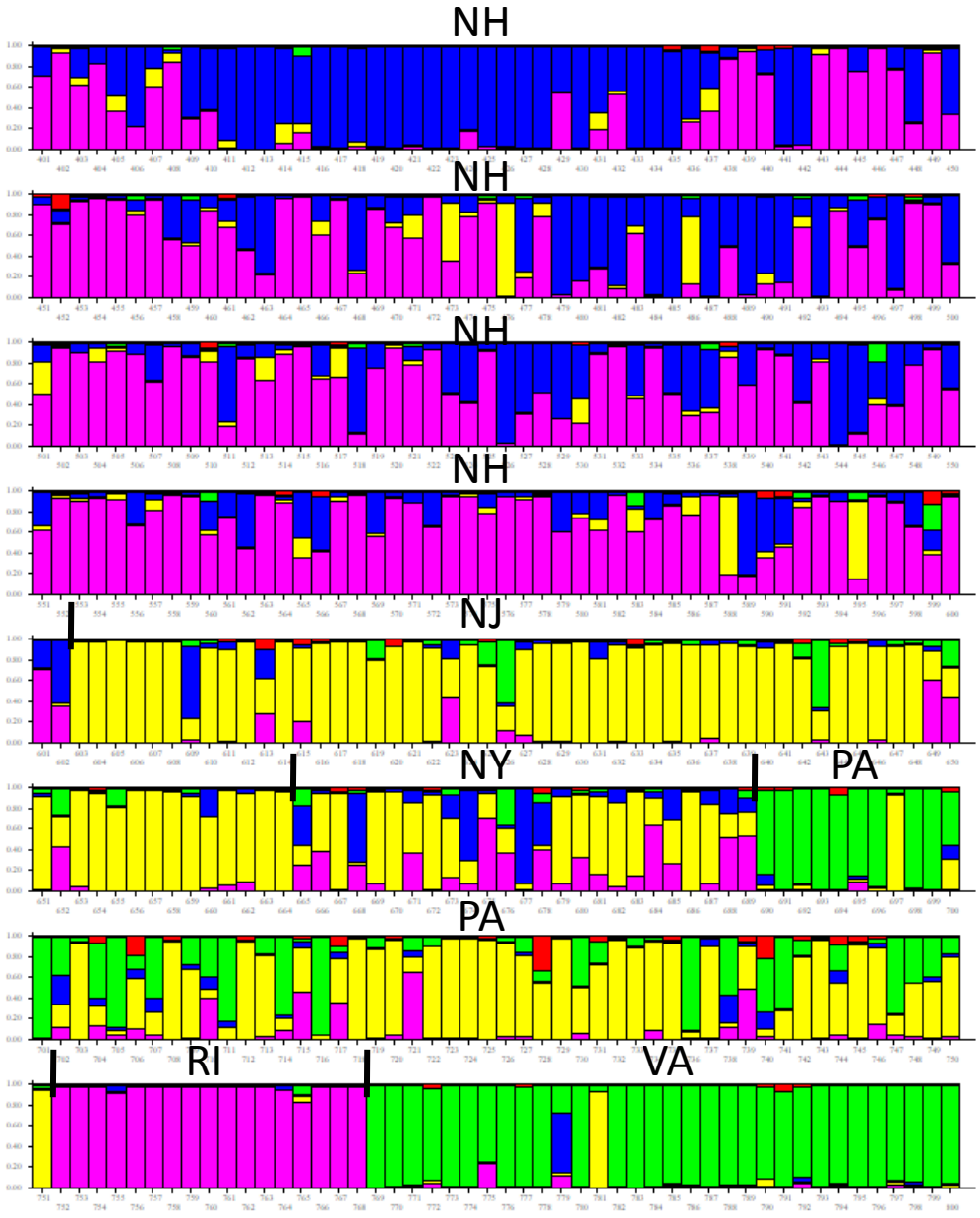
Northern ME	ME/NH	West MA	Potomac
Ar Coyote	Big Cypress	Crosby	St Sebastian
Ar Frijoles	Camel Hut	Bumblebee	Lone Tule
Ar Tio Lino	Smiley	Charcoal House	Hidden
Ar Colorado	Turpentine	Wildcat	Silvertip
Ar Yupa	Bullhead	L Bearskin	Box Canyon
Big Cypress	Ar del Cuervo		Waterfall
Monroe	Fortification		August
Ar Tierra Blanca	Ar los Barrancos		Chicken
Camel Hut	Yoosa		Diversion
Tanager	Flood		Tomahawk
Baxter	Cyclone		Pumpkin

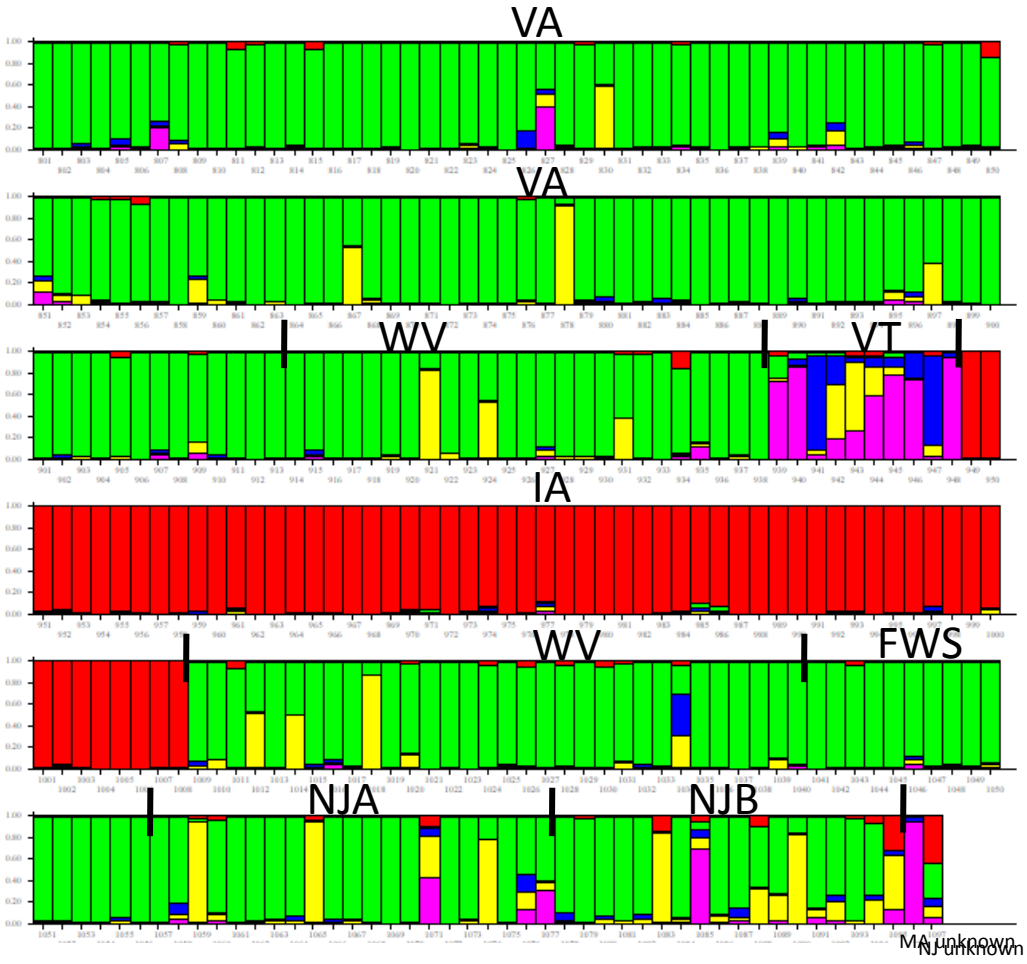
Table E.2. Full sibling families within sites tested for the Potomac, northern Maine and New Hampshire/Maine groups.

Site	Indiv. No.					
ME Smiley	455	457	474			
ME Ar. Tio Lino	280	295	307			
ME Ar. Frijoles	257	271	274			
ME. Ar. Tierro Blanca	411	412	416			
	415	420	421			
ME Ar. Yupa	344	360	365			
ME Monroe	388	390	399			
MA Crosby	48	64	66			
VA Chicken	1093	1095	1100	1106	1107	1108

Appendix F. Combined STRUCTURE output from NE and Midwest samples and confiscations (FWS) and captive populations (NJ) and unknown individuals (MA, NJ). The STRUCTURE output is shown as K=5 to be comparable to the NE groups presented above; however another K value may be more optimal for these data. These data are shown as preliminary data and are used to include as wide a data set as possible to classify the unknown samples (captives, confiscations and unknowns). The major population groups are shown as: eastern MA/RI=pink; Potomac (VA, WV, MD)=green; north ME=blue; NJ/NY=yellow; and Midwest=red. A higher K-value allows for finer scale population differentiation, but also may over-cluster the data. Therefore, individuals that are not clearly assigning to a major population cluster should be interpreted as 'unassigned' and not necessarily due to complex admixture.




NH
NH
NH
NH
NJ
NY
PA
PA
RI
VA


VA
VA
WV
VT
IA
WV
FWS
NJA
NJB
MA unknown
NJ unknown