Received: 9 November 2019 | Revised: 11 June 2020 | Accepted: 15 June 2020

DOI: 10.1111/mec.15524

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Population genetic structure and demographic history of the lone star tick, *Amblyomma americanum* (Ixodida: Ixodidae): New evidence supporting old records

Paula Lado | Megan L. Smith | Bryan C. Carstens | Hans Klompen

Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, OH, USA

Correspondence
Paula Lado, Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, OH, USA.
Emails: ladohenaise.1@osu.edu; pau.parasito@gmail.com

Funding information
American Phylosophical Society; National Science Foundation, Grant/Award Number: DEB-1457519

Abstract

Range expansions are a potential outcome of changes in habitat suitability, which commonly result as a consequence of climate change. Hypotheses on such changes in the geographic distribution of a certain species can be evaluated using population genetic structure and demography. In this study we explore the population genetic structure, genetic variability, demographic history of, and habitat suitability for *Amblyomma americanum*, a North American tick species that is a known vector of several pathogenic microorganisms. We used a double digestion restriction site-associated DNA sequencing technique (dd-RAD seq) and discovered 8,181 independent single nucleotide polymorphisms (SNPs) in 189 ticks from across the geographic range of the species. Genetic diversity was low, particularly when considering the broad geographic range of this species. The edge populations were less diverse than populations belonging to the historic range, possibly indicative of a range expansion, but this hypothesis was not statistically supported by a test based on genetic data. Nonetheless, moderate levels of population structure and substructure were detected between geographic regions. For New England, demographic and species distribution models support a scenario where *A. americanum* was present in more northern locations in the past, underwent a bottleneck, and subsequently recovered. These results are consistent with a hypothesis that this species is re-establishing in this area, rather than one focused on range expansion from the south. This hypothesis is consistent with old records describing the presence of *A. americanum* in the northeastern US in the early colonial period.

KEYWORDS
genetic diversity, population genetics, range expansion, species distribution, ticks

1 | INTRODUCTION

An understanding of the natural history, ecology, demography, and population genetics of disease vectors is a prerequisite to understanding the dynamics of vector-borne diseases (VBDs) because these factors may influence the transmission of the pathogen. Despite the public and veterinary health significance of VBDs, knowledge about the genetic structure and demographic history of vectors is lacking for many species, particularly ticks (Khatchikian et al., 2015). Disease outbreak and prevalence are dependent on genetic structure and demographic processes in the vector and knowledge of these factors may help identify regions that are newly at risk for disease outbreaks. Here, we assess the population genetic structure, genetic variability, demographic history of, and habitat

suitability for *Amblyomma americanum* (Linnaeus, 1758), commonly known as the lone star tick. This tick is the main vector of tick-borne diseases (TBDs) such as ehrlichiosis, tularemia, and STARI (southern tick-associated rash illness; Childs & Paddock, 2003; Goddard & Varela-Stokes, 2009), as well as the likely cause of Alpha-gal syndrome ("red meat allergy") in humans (Jackson, 2018).

Amblyomma americanum feeds on medium to large mammals, including humans, in all instars. It is an aggressive biter (Childs & Paddock, 2003) accounting for most reported human bites in the southeastern and Atlantic portions of North America (Merten & Durden, 2000). In 1986 its geographic range was described as ranging from west-central Texas (TX), north to Missouri (MO) and east to the Atlantic Ocean, with a northernmost coastal population in New Jersey (NJ) (Hair & Bowman, 1986). Specimens from more northern localities were occasionally reported, but it is unclear whether these ticks were established at those sites. More recent reports suggest that this species has been expanding its range westwards and northwards during the last few decades (Barrett et al., 2015; Cortinas & Spomer, 2013; Monzón, Atkinson, Henn, & Benach, 2016), resulting in a historic/expanded distribution like the one presented in Monzón et al. (2016) and illustrated in Figure 1. Furthermore, it has been reported that population sizes are increasing in other locations such as Long Island (New York, NY) and Maryland (MD) (Carroll, 2011; Ginsberg et al., 1991). Taken together, these sources document a dynamic recent history of *A. americanum*, and highlight that the western and northern boundaries of this species geographic distribution are not well established.

New England can serve to illustrate the uncertainties observed for the range of *A. americanum*. The Pleistocene glaciations clearly affected species distribution and demography in North America (Hewitt, 2000; Prentice, Bartlein, & Webb, 1991; Sakamoto, Goddard, & Rasgon, 2014), and many species are assumed to have expanded their range northwards following the retreat of the most recent glaciers (~10,000 years ago) (Cox, Stringer, Moseley, Chippindale, & Streicher, 2018; Pielou, 1991). *Amblyomma americanum* could be one of those species. Under this scenario, *A. americanum* has been absent from the north for a very long time (or never occurred there). Recent records from more northern sites are therefore interpreted as consistent with a hypothesis of a northward range extension.

However, some historical observations do not fit well with this hypothesis. In 1912, a map of the "probable" distribution of *A. americanum* included Michigan (MI), NY, and the New England states (Hooker, Bishopp, Hunter, & Wood, 1912). A few decades later, the distribution map was revised and all northern locations were removed (Bishopp & Trembley, 1945). Interestingly, Bishop and

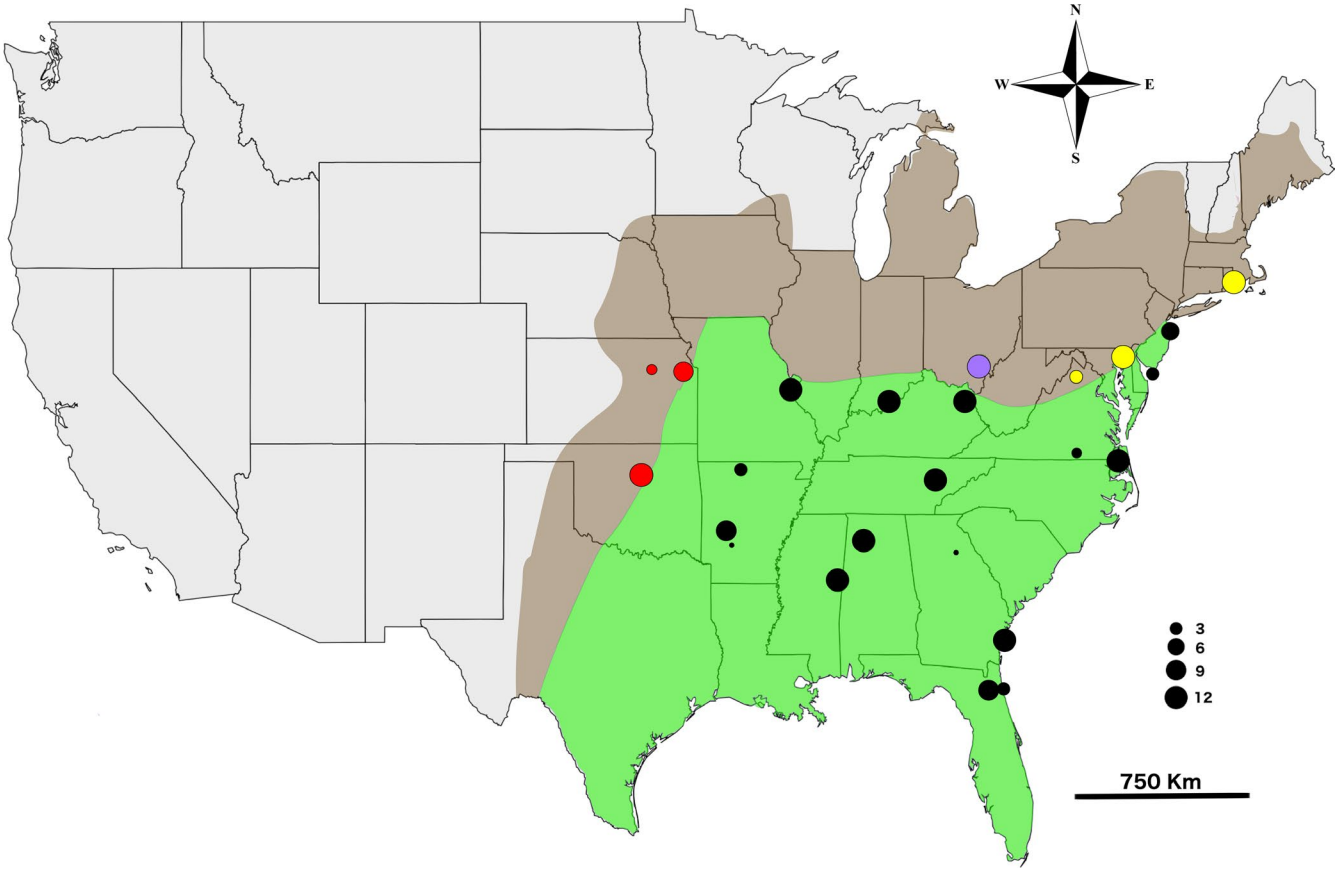

FIG. 1 Map showing the geographic origin of the sampled ticks. The expanded range of *A. americanum*, following Monzón et al. (2016) is shown in dark brown, and the historic range in light green. The circles indicate sampling locations, and the size of the circle reflects the number of specimens analysed for each location. The colour of the circles indicates different regions: black, historic range; yellow, northeast; purple, midwest; red, west

Trembley (1945) noted that lone star ticks might have been more numerous in the northern states in the recent past. In this context, it is worth noting a report by Fitch (1872) of a specimen from NY collected in 1,830. The potential presence of this tick in New England states during early colonial times was also discussed in Paddock and Yabsley (2007). These reports suggest a hypothesis of re-establishment in northern areas, rather than a northward expansion. On the other hand, several efforts on Long Island during the 1940s and 1950s failed to collect any *A. americanum*, so it still remains unclear if and when *A. americanum* was established in that area, or in any northern inland locations, in the past.

Investigations of the population genetic structure of lone star ticks may help resolve these types of questions, and to date, several such studies have been conducted (Hilburn & Sattler, 1986; Mixson, Lydy, Dasch, & Real, 2006; Reichard et al., 2005; Trout, Steelman, & Szalanski, 2010). However, most studies have important limitations, such as: sample size, the use of a single gene, and the use of ticks raised in laboratory conditions. Hilburn and Sattler

TABLE 1 Summary of the sampling locations included in this study. The Region column corresponds to each of the four regions mentioned throughout the manuscript: Midwest (MW), northeast (NE), historic (HC), and western (W). A complete table with sampling information and voucher information for the individual ticks analysed is available in File S1

Raw data ID	Locality	State	Region	Coordinates
115028	Prudence Island	RI	NE	41.6033, −71.3175
115029	Prudence Island	RI	NE	41.6033, −71.3175
115030	Prudence Island	RI	NE	41.6033, −71.3175
N8441B	Smithsonian conservation biology institute	VA	NE	38.8876, −78.1651
N8442A	Smithsonian conservation biology institute	VA	NE	38.8876, −78.1651
N8448A	Smithsonian conservation biology institute	VA	NE	38.8876, −78.1651
119275	Aberdeen Proving ground	MD	NE	39.4669, −76.1307
119239	Vinton-Furnace Exp forest	OH	MW	39.2067, −82.3961
119240	Vinton-Furnace Exp forest	OH	MW	39.2067, −82.3961
N8186A	Konza prairie biological station, Riley Co	KS	W	39.1068, −96.6087
N8204A	Konza prairie biological station, Riley Co	KS	W	39.1068, −96.6087
N8215A	The University of Kansas Biological station	KS	W	39.0389, −95.2063
N8216H	The University of Kansas Biological station	KS	W	39.0389, −95.2063
128238	Near stillwater	OK	W	36.1106, −97.0430
115002	O'bannon, Fox Hollow	IN	HC	38.1917, −86.2856
115075	Noxubee Co	MS	HC	33.1121, −88.5364
115098	Camp Blanding training site, St Agustine	FL	HC	29.9803, −81.9826
115122	Guana River, St John's Co	FL	HC	30.0233, −81.3274
119568	Bethel Valley Rd, Oak Ridge	TN	HC	35.9529, −84.2738
119582	Grayson lake state Park	KY	HC	38.2041, −83.0184
119936	Tyson Research center	MO	HC	38.51, −90.57
119957	Sapelo Island	GA	HC	31.3999, −81.2862
128211	ASU Biological Station, Marion Co	AR	HC	36.0979, −92.5518
128212	ASU Biological Station, Marion Co	AR	HC	36.0979, −92.5518
128231	Camp clearfork	AR	HC	34.508, −93.391
128232	Camp clearfork	AR	HC	34.508, −93.391
128234	Arkadelphia	AR	HC	34.1134, −93.1269
128251	South hill	VA	HC	36.7258, −78.1503
129127	William B Bankhead NF	AL	HC	34.2338, −87.3936
129129	Martha Washington trail	NC	HC	36.5057, −76.3551
129737	Athens, botanical garden	GA	HC	33.8997, −83.3827
129759	Mount Laurel	NJ	HC	39.9868, −74.8559
99030	Shark river park	NJ	HC	40.2007, −74.0945
99031	Shark river park	NJ	HC	40.2007, −74.0945
99036	Shark river park	NJ	HC	40.2007, −74.0945

(1986) employed variations in 21 enzymes to assess nine populations from TX, Oklahoma (OK), Kentucky (KY), Florida (FL), South Carolina (SC), Virginia (VA) and NJ. The authors concluded that *A. americanum* is genetically homogeneous throughout its geographic distribution. Some years after, two studies used the mitochondrial gene marker 16SrDNA to analyse populations from Georgia (GA) (Mixson et al., 2006) and Arkansas (AR) (Trout et al., 2010), both reaching the same conclusion as Hilburn and Sattler. On the other hand, Reichard et al. (2005) detected genetic divergence between two OK populations through the analysis of the nuclear Inter transcribed spacer 2 (ITS-2). However, one of the populations included colony ticks (ticks reared in the lab). Colony ticks can be genetically distinct (Monzón et al., 2016; Reichard et al., 2005), and thus the differences observed could reflect wild versus laboratory ticks rather than population structuring among OK populations.

In order to overcome most of the limitations of the above-mentioned studies we follow Monzón et al. (2016) and incorporate next generation sequencing (NGS) technologies. At present, these technologies are available and relatively cost-effective, albeit novel in tick research. Monzón et al. (2016) investigated the population genetic structure of *A. americanum* using GBS (Genotyping by sequencing) and detected population genetic structure between geographic regions (historic: the Carolinas versus. expanded: NY and OK), a result that differs from the conclusions reached by an earlier investigation based on enzymes. However, this study did not aim at exploring genetic variability across the entire range of *A. americanum*. In order to accurately assess the genetic structure of an entire species, comprehensive sampling of natural *A. americanum* populations across its current geographic range is required. Monzón et al. (2016) used specimens from only four localities. Second, that study used colony ticks for the OK samples, which, as noted above, can be problematic. Ideally, only ticks sampled from natural populations should be included in population genetics analyses.

The primary goal of this study is to examine population genetic structure and genetic variability across the range of *A. americanum*, employing genomic data from nearly 200 adult ticks. The assembled data allow us to test the following hypotheses: (i) *A. americanum* populations from the expanded range are genetically distinct from populations of the historic range; (ii) the genetic variability is lower in expanded areas than in historic areas of the range; (iii) *A. americanum* is expanding its geographical range north- and westwards; and (iv) the demographic history of, and habitat suitability for *A. americanum* in the Northeastern U.S. is consistent with re-establishment in this area, rather than range expansion from the south.

2 | MATERIALS AND METHODS

2.1 | Sampling and barcoding of tick specimens

A total of 189 adult *A. americanum* specimens were used for this study. Ticks were collected at 24 localities in 17 states (number of ticks in parenthesis): Alabama (AL, 12), Arkansas (AR, 12), Florida (FL, 12), Georgia (GA, 12), Indiana (IN, 11), Kansas (KS, 9), Kentucky (KY, 12), Maryland (MD, 12), Mississippi (MS, 12), Missouri (MO, 11), New Jersey (NJ, 10), North Carolina (NC, 12), Oklahoma (OK, 12), Ohio (OH, 12), Rhode Island (RI, 12), Tennessee (TN, 11), and Virginia (VA, 5) (Figure 1). All ticks were collected from vegetation and preserved in 95% ethanol. Detailed information about each individual tick is provided in Table 1. All ticks were accessioned in the Ohio State Acarology Collection (OSAL). The collection data are accessible online through the OSAL database (https://acarology.osu.edu/database), and OSAL numbers are presented in Table 1 and File S1.

2.2 | DNA extraction, library preparation, and SNP calling

Genomic DNA from ticks was extracted using QIAGEN Blood & Tissue kit. Tick exoskeletons were preserved following previously published protocols (Beati & Keirans, 2001; Beati et al., 2012). A small portion of the posterolateral idiosoma of each tick was cut-off using a disposable scalpel to facilitate the penetration of the buffer into the tick's tissues, and both fragments (the rest of the tick and the cut fragment of the idiosoma) were incubated overnight in 180 µl Qiagen ATL lysis buffer (Qiagen) and 20 µl of a 14.3 mg/ml solution of proteinase K (Roche Applied Sciences). After complete lysis of the tick tissues and repeated vortexing, the exoskeleton was stored in 70% ethanol and kept as a voucher specimen at the OSAL.

For the discovery of SNPs (single nucleotide polymorphisms), we used double digest restriction site associated DNA sequencing (dd-RADseq) following a modified protocol based on Peterson, Weber, Kay, Fisher, and Hoekstra (2012). Extracted genomic DNA samples were quantified using a Qubit 2.0 fluorometer and normalized to 10 ng/µl. A total of 20 µl of each sample were used for the dd-RAD library preparation. The two restriction enzymes used for DNA digestion were SbfI and MspI. Once the DNA samples were digested and the barcodes ligated, the libraries were amplified using conventional PCR to increase the number of fragments. After fragment amplification, a size selection step was performed employing a Blue Pippin (Sage Sciences) to select fragments between 300 and 600 bp. Samples were quantified using a bioanalyser and sequenced using an Illumina HiSeq with single end 150-base pair reads. Raw sequence reads were demultiplexed and processed using the ipyrad pipeline (Eaton & Overcast, 2016). This pipeline inputs raw sequence reads and outputs loci, alleles, and SNPs. Base calls with Phred score below 33 were replaced by *N*s, and a clustering threshold of 90% was employed to assemble reads into loci. Both the minimum depth at which statistical base calls are made during consensus base calling and the minimum depth at which majority rule base calls are made was set to six, such that only statistical base calls are made. The maximum depth was set to 10,000 to account for PCR duplicates, plastid markers, and potential paralogues. We also used a strict filter for adapters (filter_adapters 2 in ipyrad), reads were trimmed to 145 bp (first five bases of the reads were removed), and reads shorter than

50 bp were eliminated. Furthermore, we required that a locus was present in at least 40 tick individuals for it to be retained for downstream analyses.

2.3 | Data sets

For this study, we followed the division of historic and expanded range of the geographic distribution of *A. americanum* (see Table 1, Figure 1 and File S1) of Monzón et al. (2016). For most of the analyses we then divided the tick samples into four regions. Throughout the manuscript these regions are arbitrarily designated as: (i) historic, corresponding to the historic range (abbreviated as HC); (ii) midwest, belonging to the expanded range and including all samples from OH (abbreviated as MW); (iii) west, belonging to the expanded range and including all samples from KS and OK (abbreviated as W); and (iv) northeast, also belonging to the expanded range and including samples from RI, MD, and VA (abbreviated as NE). Thus, the historic group represents populations that are believed to have been established in those areas for a fairly long time (historic range); whereas the MW, W, and NE regions represent 3 edge zones along what is considered to be the expanded range (three areas of the border of the distribution range). This data set is referred to as "the four regions data set" throughout the manuscript. For a few analyses (e.g., DAPC), however, we divided the samples per state, because this grouping made it easier to identify and show specific patterns. We refer to that data set as "the states data set" throughout the manuscript, to distinguish it from the four regions data set.

2.4 | Genetic structure

To infer genetic population structure within *A. americanum* we used STRUCTURE v2.3.4 software, a model based approach developed by Pritchard, Stephens, and Donnelly (2000) that finds genetically differentiated populations. STRUCTURE is a Bayesian clustering method that assigns individuals in the sample to populations, or to two or more populations if their genotypes indicate that they are admixed. The program assigns individuals to clusters (denoted by *k*) assuming linkage equilibrium within populations. For this analysis we used the ".ustr" output file from ipyrad, a file that includes unlinked markers. We assessed 10 *k* values, from $k = 1$ to $k = 10$, and ran five independent runs for each *k* value. The number of MCMC was set to half a million, and burnin to 75,000. Additionally. Additionally, STRUCTURE was run without any priors about population information; thus, the individual samples were not preassigned to specific clusters.

Independent STRUCTURE runs were combined using CLUMPP (Jakobsson & Rosenberg, 2007), and the optimal number of clusters chosen using the Evanno method (Evanno, Regnaut, & Goudet, 2005) in STRUCTURE Harvester (Earl & vonHoldt, 2012). Structure plots for the chosen value of *k* were built in R using the barplot function (R Development Core Team, 2011).

Discriminant analysis of principal components (DAPC) is a model free approach developed by Jombart, Devillard, and Balloux (2010). DAPC is a multivariate statistical method that does not make any assumptions about Hardy–Weinberg or linkage equilibrium. As for the STRUCTURE analysis, we used the ".ustr" output file from ipyrad. We employed this approach to further explore the relationships among samples for the states data set (the individual samples were assigned to their respective US state). The scatterplot function was used to visualize those relationships and to plot the densities of individuals on the first discriminant function.

Given the genetic distinctiveness of RI samples (see Results section), both STRUCTURE and DAPC analyses were performed including and excluding RI samples.

2.5 | Genetic diversity and genetic differentiation between regions

Basic diversity statistics by geographic region (HC, MW, W, NE), including the number of alleles ($N_{alleles}$), allelic richness (A_r), observed heterozygosity (H_{OBS}), expected heterozygosity/gene diversity (H_{EXP}), and inbreeding coefficient (F_{IS}) were calculated in R using the Adegenet (Jombart, 2008) and Hierfstat (Goudet, 2005) packages. To test if observed heterozygosity was equal to expected heterozygosity we used the Bartlett test included in the Adegenet package. Departures from Hardy-Weinberg (HW) equilibrium per locus were determined using the hw.test function in Fisher's exact test for HW equilibrium per geographic region (null hypothesis: "random union of gametes") was performed in R, using the Genepop package (Rousset, 2008). The Genepop package was also used to test for heterozygous deficit, to calculate genic differentiation between regions, and to calculate inbreeding coefficients for each geographic region.

Fixation index (F_{ST}) values were used to quantify genetic differentiation. Global F_{ST} (function *basic.stats*), as well as pairwise F_{ST} (function pairwise.neifst) values between regions were calculated using the Hierfstat package (Goudet, 2005). Genetic differentiation between regions was assessed by analysis of molecular variance (AMOVA) in GenAlEx (Peakall & Smouse, 2006, 2012).

2.6 | Isolation by distance and range expansion

We tested for isolation by distance (IBD) (Wright, 1943) using Mantel tests. These analyses (as well as associated plots) were performed in R, using the Adegenet package (Jombart, 2008). From the different genetic distances available in Adegenet, the absolute Provesti's genetic distance was chosen since it is not Euclidean and can handle missing data. Geographic distances were calculated from coordinates following IBD analysis directions in the Adegenet tutorial. The coordinates used for this analysis are presented in Table 1.

In addition to the IBD test, we ran a specific analysis to test for a signal of range expansion using genetic data (Peter & Slatkin, 2013, 2015). The R package called RangeExpansion

developed by Peter and Slatkin (2013) calculates the directionality index, *psi*, and estimates the most likely origin of expansion. The parameter psi measures asymmetries in derived allele frequencies between populations to evaluate the direction of the expansion event, if any. The input for the analysis includes an snp file and a coordinates file. For this analysis we did not group the locations a priori, we provided the coordinates for each of the collection locations.

2.7 | Demographic history and habitat suitability

To test whether populations in the Northeastern US colonized the region recently or have persisted and underwent a bottleneck, we used a model-based approach in delimitR (Smith & Carstens, 2020). We constructed a folded multidimensional Site Frequency Spectrum (mSFS) following the approach described in Satler and Carstens (2016) using scripts available on github (https://github.com/meganlsmith/). Briefly, we defined two populations: the Northeastern US (NE), and all other sampled sites (Other). Due to the genetic distinctiveness of the Rhode Island samples (see Results section), two different tests were run. In the first one we considered the regions following Monzón et al. (2016), and thus, the NE included samples from RI, MD, and VA (as described in the data sets section). For the second test, however, we kept only the RI samples as part of the NE, with all other samples in the Other grouping. The objective of this design was to determine whether NE populations behave similarly, or if RI population has evolved differently, and should be treated as a separate unit. For both analyses we required that SNPs were sequenced in at least 50% of individuals from each population to be included in the mSFS. For SNPs sampled in more than 50% of individuals, we randomly downsampled. We used only unlinked SNPs and considered only biallelic SNPs.

delimitR allows users to define custom model sets in fastsimcoal (Excoffier, Dupanloup, Huerta-Sánchez, Sousa, & Foll, 2013). We defined a model set based on two hypotheses: (i) The NE population persisted in the region for more than 2,000 years, undergoing a bottleneck in recent years (Figure 2, Models 1 and 2); and (ii) The NE was recently colonized from Other population (i.e., the NE population and the Other population share a recent common ancestor) (Figure 2, Models 3 and 4). We also included the possibility of population expansion in the Other population (Figure 2, Models 2 and 4), corresponding to population expansion since the Last Glacial Maximum.

Effective population sizes were drawn from a uniform (200,000, 4,000,000 haploid individuals) prior for the NE population and a uniform (1,000,000, 10,000,000 haploid individuals) prior for the Other population. For models including population expansion in the Other population, at time zero, an exponential population growth rate was drawn from a uniform (−0.001, −0.00035) prior, and exponential population growth stopped at time drawn from a uniform (100, 10,000 generation) prior. This tick species has on average one generation per year, and the priors here are presented as number of generations (which is equivalent to the number of years). In models where the NE population persisted and underwent a bottleneck, at some time drawn from a uniform (0, 200 generations) prior, the population was resized. The proportion of the population that persisted was drawn from a uniform (0.001, 0.1) prior. After 100 generations, the population returned to the original size. When the NE population persisted, divergence times between NE and Other populations were drawn from a uniform (1,000, 1,000,000 generations) prior. In models

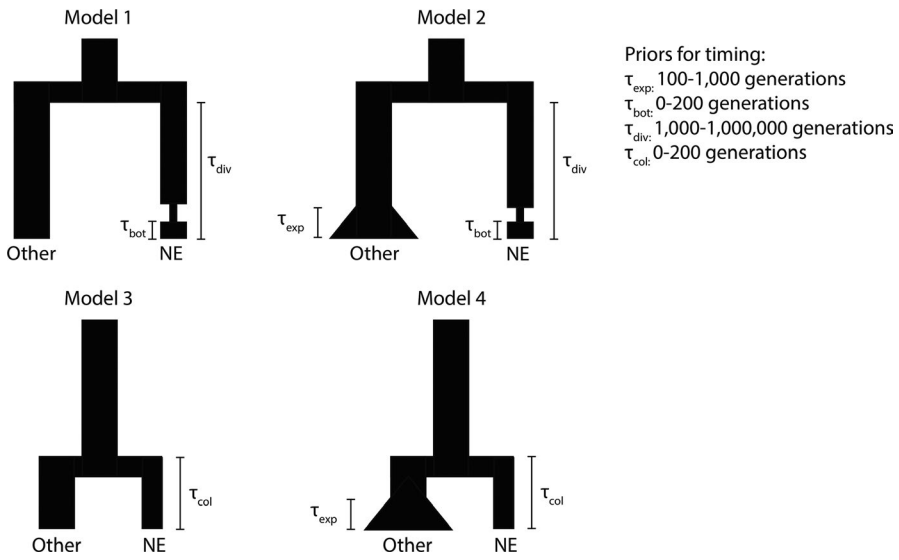

FIG. 2 Demographic models tested. Model 1: northeastern (NE) population has persisted in the region, undergoing a bottleneck in recent years; Model 2: NE population has persisted in the region, undergoing a bottleneck in recent years, and the "Other" populations has undergone population expansion; Model 3: NE population originated recently by colonization from "Other"; Model 4: NE population originated recently by colonization from "Other", and "Other" experienced population expansion. τ_{bot}, time to the bottleneck; τ_{div}, time to the divergence of both NE and "Other" population; τ_{exp}, duration of the population expansion; τ_{col}, time to the colonization event

that did not include persistence of the NE population, divergence times between NE and Other populations were drawn from a uniform (0, 200 generations) prior.

We used the fastsimcoalsims() function in the R package delimitR, which simulates mSFS in fsc26 (Excoffier et al., 2013) to simulate 10,000 data sets under each model. To build the prior in delimitR, we used five classes to summarize the SFS. Then, we constructed a Random Forest (RF) classifier using 500 decision trees. We calculated error rates, selected the best model, and approximated the posterior probability of the best model using functions in the delimitR package.

We built species distribution models (SDMs) to explore the habitat suitability of *A. americanum* in North America under past (after the retreat of the last ice sheet) and future climatic conditions (year 2070). The occurrence data to build the models was retrieved from the OSAL database, the USNTC (United States National Tick Collection), and File S1 in Raghavan, Townsend Peterson, Cobos, Ganta, and Foley (2019) (File S2). All occurrences without location information, those including only the country as location, and all duplicated occurrences were eliminated from the data set for downstream analysis. The USNTC records were first georeferenced using geo-locate (http://www.geo-locate.org), and then added to the OSAL and Raghavan et al. (2019) records, for a total of 1,392 unique occurrence data points (File S2).

We used the 19 bioclimatic variables available in the WorldClim v1.4 climate data archive (www.worldclim.org) (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) to summarize climatic conditions, and we employed climatic data at 2.5 arc-min spatial resolution (4.5 Kms). Only uncorrelated ($r < 0.7$) climatic variables were used to build the models. To assess past and future habitat suitability, we downloaded the Mid-Holocene data available in WorldClim (~6,000 yrs ago), and the predictions for 2070. For past conditions, we used the Mid-Holocene because it is the only time period for which climatic data are available between present and the last glacial maximum.

We built the SDMs using the ensemble method, from the "biomod2" R package (Thuiller, Georges, & Engler, 2013), and four different models: MaxEnt, Generalized linear model (GLM), Generalized boosting model (GBM), and Random Forest (RF). Because our data set consists of only presence data, 10,000 pseudoabsences were randomly sampled from the background area. From our occurrences data, 80% was used to train the models, and 20% to test the models, with five replicates to assess models' performance. The models were evaluated using a ROC (receiver operating characteristics) curve, and models with ROC < 0.85 were ignored when building the ensemble model. Models passing the filters were rescaled so that we could combine them. Once the models were combined, we calculated the average of all runs to plot the habitat suitability in a single plot for each time window evaluated (Mid-Holocene, present, and future).

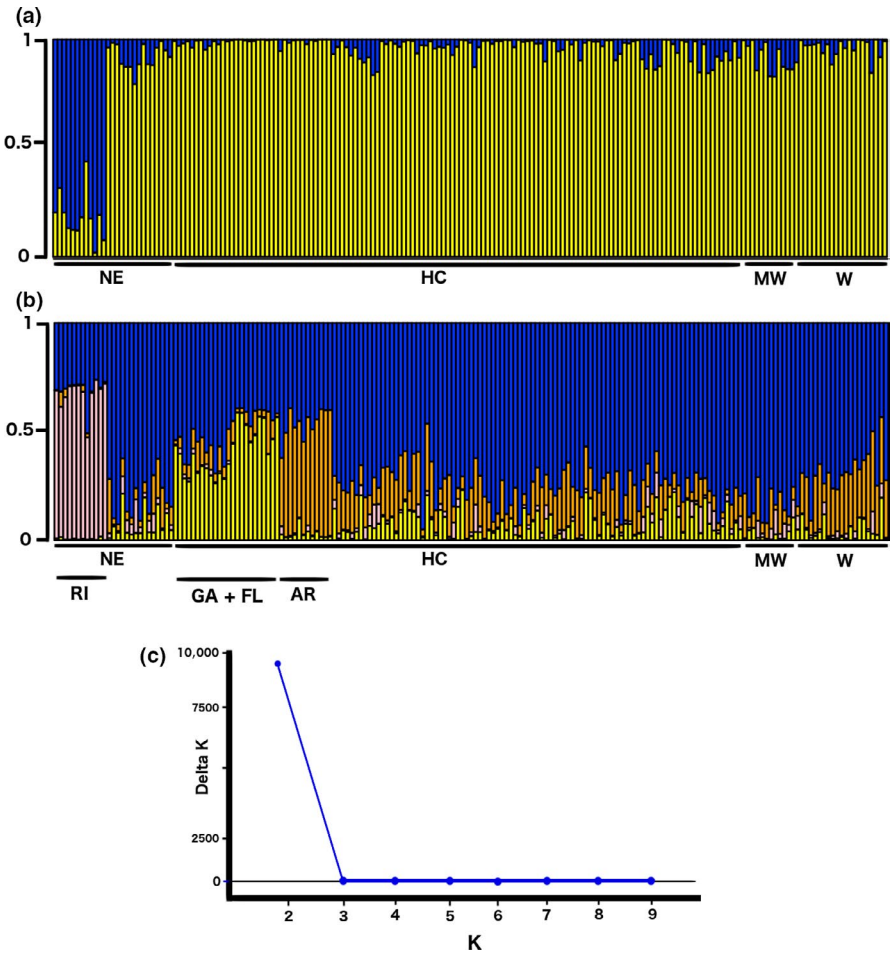

FIG. 3 Plots showing the population genetic structure of the tick populations and the optimal number of clusters. Structure plots for (a) $k = 2$; (b) $k = 4$; and (c) Evanno plot showing that $k = 2$ is the best k (Delta K value versus. k). In (a) and (b) barplots, each bar corresponds to a tick specimen, and the vertical axis represents the membership probability of an individual to each cluster. NE, Northeast; HC, historic; MW, Midwest; W, west, RI, Rhode Island; GA, Georgia; FL, Florida

3 | RESULTS

3.1 | Sequencing and raw data processing

Sequencing yield was 619.5 million reads for an average of 3.3 million reads per sample. From those raw reads 608 million passed the quality filters (average of 3.2 million per sample). The total number of prefiltered loci was 147,853. After all the filters, we retained 8,641 loci. All downstream analyses were performed using unlinked markers, and the total number of SNPs used was 8,181. Additional details about filtering steps and coverage are provided in File S3 (Table 1).

3.2 | Genetic structure

Overall, moderate levels of population structure were detected among individuals in the Structure results. The ideal number of clusters determined by the Evanno method was $k = 2$ (Figure 3). For $k = 2$, the samples from RI comprise one group, and all other samples cluster together. The same is observed at greater k values, for example at $k = 4$ (Figure 3), the clear split between RI and everything else remains. Ticks from RI appeared very similar to one another, and distinctive from all other samples, including other samples belonging to the NE region (Figure 3). Moderate levels of substructure can be observed within the heterogeneous "everything else" cluster: samples from AR are distinct, and the same for FL and GA samples. There were moderate levels of admixture between different locations within this group, with very little differentiation among geographic areas. To determine whether additional substructure was present within the "everything else" cluster, the structure analysis was repeated excluding RI samples. The best k was $k = 2$, followed by $k = 3$. The plot for $k = 2$ showed no genetic structure, and the plot for $k = 3$ showed a substructure consistent with that observed when the RI samples were included in the analysis (File S3). Thus, the exclusion of RI samples did not reveal additional substructure among the rest of the tick samples.

For the DAPC analysis, the first 50 principal components and three discriminant functions were retained, explaining 40.8% of the variance. The scatterplot showed the RI samples separated from the rest, as one distant and distinctive unit. All the other samples were closer to one another than to those from RI. Individuals from GA clustered together and were fairly separated from the remaining samples. FL specimens were the closest to the GA cluster (Figure 4). The DAPC analysis was repeated excluding RI samples to determine if additional groupings were present within non-RI samples. The scatterplot showed GA and AR as distinct groups, and FL was again the closest to GA (Figure 4).

3.3 | Genetic diversity and genetic differentiation between regions

Basic diversity statistics indicate moderate levels of genetic diversity (Table 2). The lowest values of allelic richness (A_r) and H_{EXP} correspond to the Midwest region. The per locus HW test performed revealed that 4,183 of the 8,181 loci were not in HW equilibrium. Fisher's exact tests were highly significant ($p < .05$) for the four regions, rejecting the hypothesis of random union of gametes (File S3). Overall; the observed heterozygosity was lower than the expected heterozygosity (Figure S1 in File S3, Table 2) suggesting a deficit of heterozygotes as compared to a population in HW equilibrium (Bartlett test $p < 2.2e\text{-}16$, File S3). Heterozygous deficit was confirmed by a global HW statistical test performed in Genepop ($p = 0$) (File S3).

Genetic differentiation test showed a highly significant differentiation between the NE and MW regions ($p = .0004$), and between NE and W regions ($p < .05$) (File S3). Positive and significant inbreeding coefficients (F_{IS}) were found for all four populations, ranging from 0.40 to 0.47, with the minimum (0.40) corresponding to the historic (HC) region (Table 2).

AMOVA results (File S3) also showed that the majority of the variance, 95.4%, corresponds to "within samples" variation, followed by "between samples within population" (3.4%) variation. Between populations variation accounted only for a 1.2% of the total variance.

The overall F_{ST} value for all loci and all populations was 0.05. The paired F_{ST} values for the four regions data set ranged between 0.011 and 0.087 and showed that all combinations including the NE group were considerably higher than the rest (Table 3). Thus, the NE group

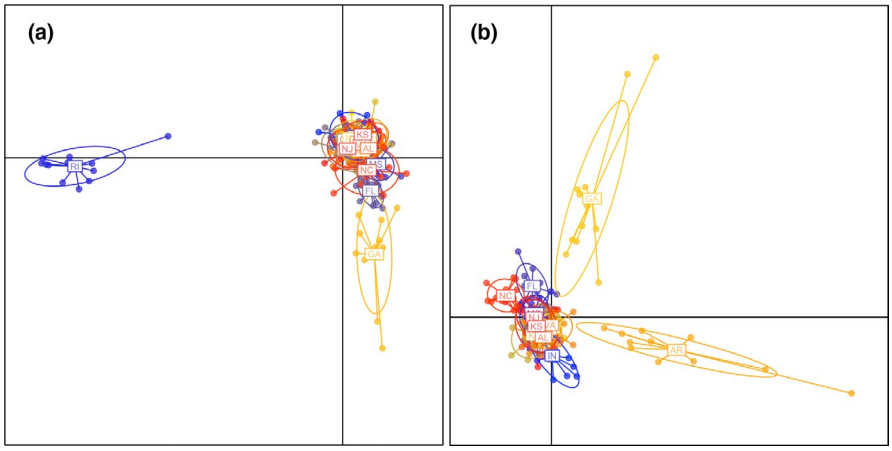

FIG. 4 DAPC scatterplots showing the relationships between ticks from different locations. The labels in the boxes correspond to states, and the ellipses represent the variation within each group. (a) Scatterplot for all 189 samples included in this study. The percentage of cumulative variance explained is 40.8%; and (b) scatterplot excluding RI samples from the analysis. The percentage of cumulative variance explained is 47.2%

TABLE 2 Basic genetic diversity statistics

Region	N	N_{ALL}	A_r	H_O	H_{EXP}	F_{IS}
W	21	10,653	1.078	0.029	0.085	0.45
MW	12	7,367	1.063	0.026	0.078	0.47
NE	27	10,726	1.075	0.029	0.079	0.46
HC	129	15,452	1.086	0.034	0.087	0.41

Note: N, number of individuals in each region (W, western; MW, Midwest; NE, northeast; HC, historic regions). N_{ALL}, number of alleles; Ar, allelic richness; H_O, observed heterozygosity; H_{EXP}, expected heterozygosity, F_{IS}, inbreeding coefficient.

is well differentiated from the other three, whereas the combinations between the HC, W, and MW groups revealed low, although statistically significant, genetic differentiation (Table 3).

3.4 | Isolation by distance and range expansion

Results from the IBD analysis suggested the genetic and geographic distance were not correlated. The individual points were dispersed and there was no support for a relationship between geographic and genetic distance (Figure S2 in File S3). Additionally, in the histogram of simulated values the correlation between the genetic and geographic matrices lay within the distribution expected by chance (no IBD) (Figure S3 in File S3).

The analysis used to test range expansion from the molecular data failed to detect signals of range expansion for *A. americanum* ($p = .097$). As described in the Methods section, for this analysis the coordinates of all sampling locations were used; the ticks were not grouped a priori.

3.5 | Demographic history and habitat suitability

For the demographic analyses two separate analysis were run. In the first one, the NE region included samples from RI, VA, and MD; whereas in the second analysis the NE included samples from RI only.

For the first analysis, after downsampling, we constructed the multispecies site frequency spectrum (mSFS) using 162 alleles from the Other population and 27 alleles from the NE population. We used 2,770 unlinked SNPs to build the mSFS. Error rates in *delimitR* indicate a high power (error rates < 1%; File S3) to distinguish among models with persistence and bottlenecks in NE populations (Figure 2; Models 1–2) and models in which NE populations were recently colonized (Figure 2; Models 3–4). However, it was more difficult to distinguish between models where the Other populations did or did not experience population expansion, particularly when colonization of the NE populations occurred recently (error rates ~ 50 percent between these two models; Table S2 in File S3).

When the RF classifier was applied to the observed data, we selected a model in which NE populations persisted and underwent a population bottleneck (posterior probability = 0.557; Table 4). The model receiving the second highest number of votes differed only in whether or not the Other populations experienced population expansion.

For the second analysis (including only RI), the mSFS was built using 177 alleles from the Other population, and 12 alleles from the NE (RI) population. 2,759 unlinked SNPs were employed to construct the mSFS. Error rates (Table S3 in File S3) indicate a high power to distinguish among models with persistence and bottlenecks in NE populations (Figure 2; Models 1–2) and models in which NE populations were recently colonized (Figure 2; Models 3–4). The most supported model was that in which NE populations persisted and underwent a population bottleneck (posterior probability = 0.774, Table 4).

Species distribution modeling was employed to explore the past, current, and present habitat suitability for *A. americanum* across the United States. The mean ROC score for the models was 0.96. After the elimination of all correlated variables, the bioclimatic variables employed for building the models were: mean diurnal range (BIO2), maximum temperature of the warmest month (BIO5), temperature annual range (BIO7), mean temperature wettest quarter (BIO8), annual precipitation (BIO12), and precipitation seasonality (BIO15). Current distribution models show high habitat suitability for *A. americanum* across its currently known distribution range, and moderate suitability in peripheric areas (northern and western). These models also indicate that areas where the habitat suitability is high have changed since the Mid-Holocene, and they will continue to change. There is a general shift of the highly suitable areas; losing territory in western inland locations and gaining territory towards northern and northeastern locations (Figure 5).

TABLE 3 Pairwise F_{ST} and p-values for the four regions: HC, historic; MW, Midwest; W, western; NE, northeast. F_{ST} values are in bold, below the diagonal; and p-values are above the diagonal

	HC	MW	W	NE
HC	–	0.003	0.001	0.006
MW	**0.035**	–	0.001	0.002
W	**0.011**	**0.024**	–	0.001
NE	**0.069**	**0.087**	**0.08**	–

4 | DISCUSSION

4.1 | Population genetic structure and genetic diversity

Moderate levels of population genetic structure were detected, with congruent clustering between a model based (Structure) and a non-model based approach (DAPC).

Overall, two groups were identified: (i) all samples from RI; and (ii) all remaining samples clustering together in an admixed and heterogeneous group. The distinctiveness of the RI samples could be due to the demographic history of New England populations (see below),

TABLE 4 Results of model selection in delimitR for both data sets: NE population comprising RI, VA, and MD samples, and NE comprising RI samples only. We report the number of decision trees that voted for each model when the classifier was applied to the observed data, and the approximated posterior probability of the best model (Model 1 for both data sets)

NE population	Model 1	Model 2	Model 3	Model 4	Post. probability
RI, VA, MD	**193**	154	107	46	0.556
RI	**345**	140	11	4	0.774

Best model, model with the majority of the votes.

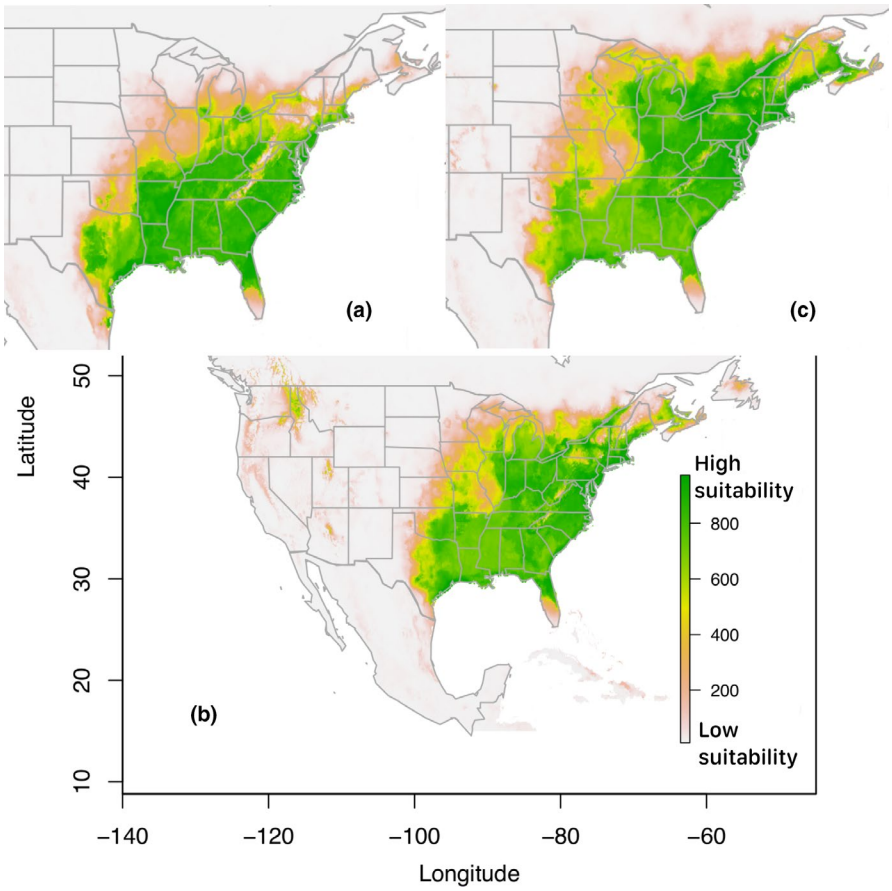

FIG. 5 Species distribution models built using the ensemble method, from the "biomod2" R package. The images show the habitat suitability for *A. americanum* during (a) the Mid-Holocene (~6,000 yrs ago); (b) present conditions; and (c) forecasted for the future, year 2070. Areas in green indicate high suitability, and grey and brown areas represent less suitable habitats for this species. Bioclimatic variables included in the models: mean diurnal range (BIO2), maximum temperature of the warmest month (BIO5), temperature annual range (BIO7), mean temperature wettest quarter (BIO8), annual precipitation (BIO12), and precipitation seasonality (BIO15)

or to the fact that Prudence island (where these samples were collected) is an isolated island population, possibly shaped strongly by evolutionary forces such as genetic drift. Within the heterogeneous "all remaining samples" cluster, two groupings could be observed: a GA-FL group which can be explained by the historical stability of southeastern populations (older large populations with little population size fluctuations); and an AR group, suggesting that ticks from that region may have gone through different evolutionary processes. It is possible that AR colonization was different from that of many other inland locations (e.g., OK, KS, MO), for example a single event instead a constant movement of ticks from nearby locations. It is also possible that the distinctiveness of these ticks is due to isolation by environment. A further exploration of this local population should be performed to evaluate the reasons for the distinctiveness of these samples.

Genetic diversity is one of the main components of biodiversity, and a critical factor in predicting colonization success during range expansion (Chapman, Nakagawa, Coltman, Slate, & Sheldon, 2009; Crawford & Whitney, 2010; Frankham, 2005; Willi, Van Buskirk, & Hoffmann, 2006). Because of the large geographic range of *A. americanum*, we were surprised at the relatively low genetic diversity, although this finding is consistent with the low levels of genetic structure detected. The observed heterozygosity was lower than expected in all four regions, consistent with a deficit of heterozygous individuals (File S3). The inbreeding coefficient was fairly high (average of 0.45) for the different regions, and the lowest F_{IS} corresponded to the historic region, congruent with admixture between subpopulations and relatively higher genetic diversity.

Thus, populations from the expanded range do not appear genetically distinct from populations belonging to the historic range. However, RI populations, which are included within the expanded range, are genetically distinct from all other populations sampled regardless of their geographic origin.

4.2 | Isolation by distance and range expansion

The Mantel tests (considering each coordinate point as one location) did not detect signals of isolation by distance among tick samples, suggesting that geographic distance alone does not explain the differences in genetic diversity. Moreover, the test for range expansion did not support a range expansion for the ticks tested, which belonged to 24 unique collection locations. It is possible that expansion events are too recent to be detected by genetic data alone, or that the signals are not strong enough to be detected by the range expansion test.

Even though *A. americanum* seems to be less variable than expected, our results show edge populations to be less diverse than historic populations (Table 2). The decreased diversity in edge populations aligns well with population genetic theory expectations for the front of the distribution during range expansions, mainly due to founder events (Excoffier, Foll, & Petit, 2009; Nei, Maruyama, & Chakraborty, 1975). This expectation is empirically supported by previous reports in other organisms, such as plants (*Petunia*), lice, and voles (Demastes, Hafner, Hafner, Light & Spradling; Silva-Arias et al., 2017; White, Perkins, Heckel, & Searle, 2013). These results contrast, to some extent, with those of Monzón et al. (2016), which indicated that lone star edge populations are as diverse as historic populations. The partial discrepancies between Monzón et al. (2016) and this study may be due to the methodologies employed, or to the use of wild populations only (in Monzón et al., (2016) OK ticks, one of the two populations from the expanded range, were from a laboratory colony).

All in all, the analyses provide evidence consistent with range expansion (i.e., decreased variation in edge populations), even though the range expansion test failed to statistically support that hypothesis. Our results reject the hypothesis of IBD, suggesting that geographic distance alone does not explain the observed variation. There are several potential explanations for these findings: (i) migration and colonization events are too recent to be detected by the range expansion analysis used; (ii) genetic diversity is maintained through gene flow and therefore the test (based on founder effects) fails to support a "real" range expansion; (iii) some of the populations were present in the past, went through a bottleneck (see demography results below) and thus, the process is a re-establishment, instead of a range expansion; (iv) it is possible that some of the populations belonging to the expanded area already had very small established populations, and they are only noticed during recent years due to a significant increase in local population sizes or surveillance efforts; and (v) a combination of the aforementioned reasons.

4.3 | Demographic history and habitat suitability

The demographic modeling results presented here support old records stating that *A. americanum* was present in northern (specifically northeastern) locations. Our results indicate that after being present in the northeast for a long time, lone star ticks went through a bottleneck, nearly wiping them out in this region, and are currently in a recovery phase. Even though both demographic modeling tests suggested this, the results appear to be driven by the RI population (for example, the signal of bottleneck and recovery is stronger when the NE is represented by RI samples only). This could be indicating that the history of New England as a whole is different from that of the RI population included in this study (which is an island population). It is possible that New England as a whole went through a bottleneck and is in recovery phase, with the alternative that the RI island population is the sole remnant of older populations present in the coastal area. These hypotheses are worth exploring further but require the inclusion of more sampling locations.

Even though the best supported demographic models did not include population demographic expansion of the "Other" group (historic + western + Midwestern populations), the second-best model did, and thus, the possibility of demographic expansion cannot be excluded. It is also worth reiterating that the demographic approaches employed are usually used to detect older demographic events (thousands or millions of years), and thus, it is possible that these events are too recent (few centuries) to be detected. To further test this idea, alternative approaches should be assessed, as well as sampling efforts redirected to test demographic expansion in specific areas, rather than focus on a larger scale. For example, signatures of such a local population expansion have been reported for some locations within the *A. americanum* distribution (Carroll, 2011; Ginsberg et al., 1991; Trout et al., 2010).

Our species distribution models show that the suitable areas for this species have shifted since the Mid-Holocene and will continue to change according to the projections for 2070 (Figure 5). Western inland localities are more suitable in the present than 6,000 years ago, although they may become less suitable in the future (Figure 5). In contrast, some northern locations, including parts of New England, were moderately suitable for *A. americanum* in the past, and have become somewhat more suitable since. A future range expansion and establishment in more northern and northeastern locations (including southeastern Canada) seems likely, as habitat suitability will continue to improve (Figure 5) probably as a result of climate change. Notably, the tick has already been detected in southeastern Canada (Nelder et al., 2019; Raghavan et al., 2019).

Precipitation and the maximum temperature of the warmest month consistently appear as limiting factors for this tick distribution (Raghavan et al., 2019; Sagurova et al., 2019; Springer, Jarnevich, Barnett, Monaghan, & Eisen, 2015). The majority of the bioclimatic variables retained in the analysis (BIO2, mean diurnal range; BIO5, max temperature of the warmest month; and BIO15, precipitation seasonality) were also highlighted as relevant by Raghavan et al. (2019) and Springer et al. (2015). Continuous passive and active surveillance efforts will help track changes in species distribution and establishment, allowing better preparation if the species and associated pathogens reach and establish in new locations.

The hypothesis that New England, or some northeastern (including Rhode Island) *A. americanum* populations underwent a

bottleneck with recent recovery is consistent with the history of this region, and habitat preference of this species (wooded areas). New England was a predominantly forested region before the European settlement that has undergone about 300 years of intense logging, deforestation for agriculture, and, recently, natural reforestation (Barton, White, & Cogbill, 2013; Thompson, Carpenter, Cogbill, & Foster, 2013). The area underwent approximately 100 years (between 1,760 and 1,860) of intense land clearance for agriculture (O'Keefe & Foster, 1998), which destroyed most of the tick habitat and refugia, along with strong population decreases of potential hosts (especially white-tailed deer). This would explain the collapse of *A. americanum* populations. Beginning in the second half of the 19th century, a decrease in agriculture was followed by a regrowth of the forest (although current forest is very different from that presettlement). Along with reforestation of the region, populations of some animals, like deer, exploded (O'Keefe & Foster, 1998; Thompson et al., 2013), with the possible consequence of dispersing ticks from other areas.

There are good records of this species for some geographic regions in the literature, for example for Long Island, New York. Surveys in this location during the decade of 1940 failed to find lone star ticks, although ticks of other species were numerous (Anastos, 1947; Collins, Nardy, & Glasgow, 1949). In the 1970s *A. americanum* was detected in Long Island (Good, 1972), and subsequent increases in population sizes have been reported for this location (Ginsberg et al., 1991; Means & White, 1997). This can be explained by a range expansion from more southern populations with successful establishment across the area. But it could also be explained by a recolonization of the area after older populations went through a bottleneck. Even though a recent range expansion and colonization may seem the most likely explanation, it is worth considering the possibility that lone stars may have been in northeastern locations in the past (18th and 19th centuries), and then went through a bottleneck, with subsequent recovery. This is still consistent with the absence of the species in the early nineties, and later presence and increased densities. It is also consistent with the record of *A. americanum* presence in New York in 1,830 (Fitch, 1872). Thus, records since the late eighties may represent either a range expansion (Monzón et al., 2016), recolonization of the area (Bishopp & Trembley, 1945; Monzón et al., 2016), or perhaps a combination of the two.

Amblyomma americanum demography appears to be more complex than previously thought. Given the considerations mentioned above, it is likely that demographic histories differ between regions, and therefore, breaking it down by region may be necessary depending on the question asked. Such studies need to explicitly consider population genetic structure in their design, including localized sampling efforts. This study can be used as a starting point for testing these follow-up hypotheses. Unfortunately, our sampling of northeastern locations was very limited. A sampling focused on northeastern and New England populations should help either reject or support the hypothesis of recolonization (bottleneck and subsequent recovery). Overall, this species evolutionary history and demography merits further exploration, as it can substantially change the way we think about this species (its success in the past, remnant populations), and ultimately about its management and the control of the diseases it transmits.

In conclusion, lone star tick populations showed modest population genetic structure across their geographic range and are generally not very diverse at the genetic level. The RI island population (collected on Prudence island) is isolated from continental populations and it is genetically distinct from all other populations. Demographic modeling results suggest that *A. americanum* was present in northeastern locations in the past (including RI), underwent a bottleneck event, and then recovered. These results are supported by species distribution models, which show that many areas in the northeast were moderately suitable for this species in the past. Despite the limitations of this study, such as limited sampling of the northeastern region, these findings are worth exploring further with a comprehensive sampling of that area. The species distribution modeling results suggest that the northeast will be even more suitable for lone stars in 2070 than it is today. That, together with the likely range expansion of this species, emphasize the importance of continuing ongoing surveillance efforts. This work highlights the importance of study design, geographic scale, and approaches used to study wide-spread species and emphasizes the need for phylogeographic investigations into disease vectors.

ACKNOWLEDGEMENTS

We are grateful to Jerome Goddard, Howard Ginsberg, Andrew Hoffman, Benedict Pagac, Graham Hickling, Solny Adalsteinsson, Bruce Noden, Brent Newman, Jim Occi, Shelby Ford, Renn Tumlison, and Lance Durden who provided some of the tick specimens used in this study. We also thank The National Ecological Observatory Network, a programme sponsored by the National Science Foundation and operated under cooperative agreement by Battelle Memorial Institute, for providing some specimens used in this work. In addition, we thank Lorenza Beati, curator of the USNTC for the tick records, and Karine Posbic Leydet for her help with the range expansion analysis. The authors would also like to thank the reviewers whose suggestions and comments helped improving previous versions of this manuscript. This work was partially funded by NSF grant DEB-1457519 to Bryan Carstens, and the American Philosophical Society Lewis and Clark fund for exploration and field research to Paula Lado.

AUTHOR CONTRIBUTIONS

P.L., H.K., and B.C.C designed and planned the study. P.L. carried out the fieldwork in several locations in the United States and performed the laboratory work. P.L., and M.L.S., analysed the data. All authors contributed to writing the manuscript.

DATA AVAILABILITY STATEMENT

All genetic data generated in this study is available in the Dryad data base, https://doi.org/10.5061/dryad.pk0p2ngjq. Collection data will

be accessible through the Ohio State Acarology collection (OSAL) online database.

ORCID

Paula Lado https://orcid.org/0000-0003-4852-3495
Bryan C. Carstens https://orcid.org/0000-0002-1552-227X

REFERENCES

Anastos, G. (1947). Hosts of Certain New York Ticks. *Psyche (New York)*, *54*(3), 178–180. https://doi.org/10.1155/1947/81979

Barrett, A. W., Noden, B. H., Gruntmeir, J. M., Holland, T., Mitcham, J. R., Martin, J. E., … Little, S. E. (2015). County scale distribution of *Amblyomma americanum* (Ixodida: Ixodidae) in Oklahoma: Addressing local deficits in tick maps based on passive reporting. *Journal of Medical Entomology*, *52*(2), 269–273. https://doi.org/10.1093/jme/tju026

Barton, A., White, A., & Cogbill, C. (2013). *Reconstructing the Past: Maine Forests Then and Now*, Lyme, NH: The Northern Woodlands.

Beati, L., & Keirans, J. E. (2001). Analysis of the systematic relationships among ticks of the genera *Rhipicephalus* and *Boophilus* (Acari: Ixodidae) based on mitochondrial 12S ribosomal DNA gene sequences and morphological characteres. *Journal of Parasitology*, *87*(1), 32–48.

Beati, L., Patel, J., Lucas-Williams, H., Adakal, H., Kanduma, E. G., Tembo-Mwase, E., … Kelly, P. (2012). Phylogeography and Demographic History of *Amblyomma variegatum* (Fabricius) (Acari: Ixodidae), the Tropical Bont Tick. *Vector-Borne and Zoonotic Diseases*, *12*(6), 514–525. https://doi.org/10.1089/vbz.2011.0859

Bishopp, A. F. C., & Trembley, H. L. (1945). Distribution and Hosts of Certain North American Ticks. *The Journal of Parasitology*, *31*(1), 1–54. https://doi.org/10.2307/3273061

Carroll, J. F. (2011). An Increasing Presence: The Lone Star Tick, *Amblyomma americanum* (Linnaeus) (Acari: Ixodidae) in Maryland. *Maryland Entomologist*, *5*(3), 66–76.

Chapman, J. R., Nakagawa, S., Coltman, D. W., Slate, J., & Sheldon, B. C. (2009). A quantitative review of heterozygosity-fitness correlations in animal populations. *Molecular Ecology*, *18*(13), 2746–2765. https://doi.org/10.1111/j.1365-294X.2009.04247.x

Childs, J. E., & Paddock, C. D. (2003). The Ascendancy of *Amblyomma americanum* as a vector of pathogens affecting humans in the United States. *Annual Review of Entomology*, *48*(1), 307–337. https://doi.org/10.1146/annurev.ento.48.091801.112728

Collins, D. L., Nardy, R. V., & Glasgow, R. D. (1949). Some host relationships of Long Island ticks. *Journal of Economic Entomology*, *42*(1), 110–112. https://doi.org/10.1093/jee/42.1.110

Cortinas, R., & Spomer, S. (2013). Lone Star Tick (Acari: Ixodidae) occurrence in Nebraska: Historical and current perspectives. *Journal of Medical Entomology*, *50*(2), 244–251. https://doi.org/10.1603/ME12207

Cox, C. L., Stringer, J. F., Moseley, M. A., Chippindale, P. T., & Streicher, J. W. (2018). Testing the geographical dimensions of genetic diversity following range expansion in a North American snake. *Biological Journal of the Linnean Society*, *125*(3), 586–599. https://doi.org/10.1093/biolinnean/bly137

Crawford, K. M., & Whitney, K. D. (2010). Population genetic diversity influences colonization success. *Molecular Ecology*, *19*(6), 1253–1263. https://doi.org/10.1111/j.1365-294X.2010.04550.x

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eaton, D. A., & Overcast, I. (2016). *ipyRAD: An interactive assembly and analysis of RADseq data sets*.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLOS Genetics*, *9*(10), e1003905. https://doi.org/10.1371/journal.pgen.1003905

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414

Fitch, A. (1872). Fourteenth Report on the noxious, beneficial, and other insects of the State of New York. *Transactions of the New York State Agricultural Society*, *30*, 355–381.

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Ginsberg, H. S., Ewing, C. P., O'Connell, A. F., Bosler, E. M., Daley, J. G., & Sayre, M. W. (1991). Increased Population Densities of *Amblyomma americanum* (Acari: Ixodidae) on Long Island. *New York. the Journal of Parasitology*, *77*(3), 493–495. https://doi.org/10.2307/3283144

Goddard, J., & Varela-Stokes, A. S. (2009). Role of the lone star tick, *Amblyomma americanum* (L.), in human and animal diseases. *Veterinary Parasitology*, *160*(1–2), 1–12. https://doi.org/10.1016/j.vetpar.2008.10.089

Good, N. (1972). Tick locality and host records from Long Island and southeastern New York State. *Entomological News*, *83*, 165–168.

Hair, J., & Bowman, J. (1986). Behavioral ecology of *Amblyomma americanum*. In J. R. Sauer, & J. H. Sauer (Eds.), *Morphology, Physiology, and Behavioral Biology of Ticks* (pp. 406–427). New York, NY: John Wiley & Sons Inc.

Hewitt, G. (2000). The genetic legacy of the quaternary ice ages. *Nature*, *405*, 907–913. https://doi.org/10.1038/35016000

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). WORLDCLIM - a set of global climate layers (climate grids). *International Journal of Climatology*, *25*, 1965–1978.

Hilburn, L. R., & Sattler, P. W. (1986). Electrophoretically detectable protein variation in natural populations of the lone star tick, *Amblyomma americanum* (Acari: Ixodidae). *Heredity*, *57*, 67–74. https://doi.org/10.1038/hdy.1986.88

Hooker, W. A., Bishopp, F. C., Hunter, W. D., & Wood, H. P. (1912). *The life history and bionomics of some North American ticks*. Washington, D.C.: US Department of Agriculture, Bureau of Entomology.

Jackson, W. L. (2018). Mammalian meat allergy following a tick bite: A case report. *Oxford Medical Case Reports*, *2018*(2), omx098 https://doi.org/10.1093/omcr/omx098

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94.

Khatchikian, C. E., Prusinski, M. A., Stone, M., Backenson, P. B., Wang, I., Foley, E., … Brisson, D. (2015). Recent and rapid population growth and range expansion of the Lyme disease tick vector, *Ixodes scapularis*. *North America. Evolution*, *69*(7), 1678–1689. https://doi.org/10.1111/evo.12690

Linnaeus, C. (1758). Acarus. In *Systema Naturae* (10th edn., pp. 615–618). Stockholm: Salvius.

Means, R. G., & White, D. J. (1997). New distribution records of *Amblyomma americanum* (L.) (Acari: Ixodidae) in New York State. *Journal of Vector Ecology: Journal of the Society for Vector Ecology*, *22*(2), 133–145.

Merten, H., & Durden, L. (2000). A state-by-state survey of ticks recorded from humans in the United States. *Journal of Vector Ecology*, *25*, 102–113.

Mixson, T. R., Lydy, S. L., Dasch, G. A., & Real, L. A. (2006). Inferring the population structure and demographic history of the tick, *Amblyomma americanum* Linnaeus. *Journal of Vector Ecology*, *31*(1), 181–192. https://doi.org/10.3376/1081-1710(2006)31[181:ITPSAD]2.0.CO;2

Monzón, J. D., Atkinson, E. G., Henn, B. M., & Benach, J. L. (2016). Population and evolutionary genomics of *Amblyomma americanum*, an expanding arthropod disease vector. *Genome Biology and Evolution*, *8*(5), 1351–1360. https://doi.org/10.1093/gbe/evw080

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, *29*(1), 1–10. https://doi.org/10.2307/2407137

Nelder, M. P., Russell, C. B., Clow, K. M., Johnson, S., Weese, J. S., Cronin, K., … Patel, S. N. (2019). Occurrence and distribution of *Ambylomma americanum* as determined by passive surveillance in Ontario, Canada (1999–2016). *Ticks and Tick-Borne Diseases*, *10*(1), 146–155. https://doi.org/10.1016/j.ttbdis.2018.10.001

O'Keefe, J., & Foster, D. (1998). III. Changes After European Settlement. *Arnoldia*, *58*(2), 11–25.

Peakall, R., & Smouse, P. E. (2006). GENALEX *6*: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*(1), 288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Peakall, R., & Smouse, P. (2012). Quick Start GenAlEx 6.5. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Peter, B. M., & Slatkin, M. (2013). Detecting range expansions from genetic data. *Evolution*, *67*(11), 3274–3289. https://doi.org/10.1111/evo.12202

Peter, B. M., & Slatkin, M. (2015). The effective founder effect in a spatially expanding population. *Evolution*, *69*(3), 721–734. https://doi.org/10.1111/evo.12609

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*(5), e37135 https://doi.org/10.1371/journal.pone.0037135

Pielou, E. C. (1991). *After the ice age: The return of life to glaciated North America*, Chicago, IL: University of Chigago Press.

Prentice, I. C., Bartlein, P. J., & Webb, T. (1991). Vegetation and climate change in eastern North America since the last glacial maximum. *Ecology*, *72*(6), 2038–2056. https://doi.org/10.2307/1941558

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing (Vol. 1). doi: https://doi.org/10.1007/978-3-540-74686-7

Raghavan, R. K., Townsend Peterson, A., Cobos, M. E., Ganta, R., & Foley, D. (2019). Current and Future Distribution of the Lone Star Tick, *Amblyomma americanum* (L.) (Acari: Ixodidae) in North America. *PLoS One*, *14*(1), 1–13. https://doi.org/10.1371/journal.pone.0209082

Reichard, M. V., Kocan, A. A., Van Den Bussche, R. A., Barker, R. W., Wyckoff, J. H., & Ewing, S. A. (2005). Sequence variation of the ribosomal DNA second internal transcribed spacer region in two spatially-distinct populations of *Amblyomma americanum* (L.) (Acari: Ixodidae). *The Journal of Parasitology*, *91*(2), 260–263. https://doi.org/10.1645/GE-292R1

Rousset, F. (2008). genepop'007: A complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Sagurova, I., Ludwig, A., Ogden, N. H., Pelcat, Y., Dueymes, G., & Gachon, P. (2019). Predicted northward expansion of the geographic range of the tick vector *Amblyomma americanum* in North America under future climate conditions. *Environmental Health Perspectives*, *127*(10), 1–14. https://doi.org/10.1289/EHP5668

Sakamoto, J. M., Goddard, J., & Rasgon, J. L. (2014). Population and demographic structure of *Ixodes scapularis* Say in the eastern United States. *PLoS One*, *9*(7), e101389 https://doi.org/10.1371/journal.pone.0101389

Satler, J. D., & Carstens, B. C. (2016). Phylogeographic concordance factors quantify phylogeographic congruence among co-distributed species in the *Sarracenia alata* pitcher plant system. *Evolution*, *70*(5), 1105–1119. https://doi.org/10.1111/evo.12924

Silva-Arias, G. A., Reck-Kortmann, M., Carstens, B. C., Hasenack, H., Bonatto, S. L., & Freitas, L. B. (2017). From inland to the coast: Spatial and environmental signatures on the genetic diversity in the colonization of the South Atlantic Coastal Plain. *Perspectives in Plant Ecology, Evolution and Systematics*, *28*, 47–57. https://doi.org/10.1016/j.ppees.2017.06.006

Smith, M. L., & Carstens, B. C. (2020). Process-based species delimitation leads to identification of more biologically relevant species*. *Evolution*, *74*(2), 216–229. https://doi.org/10.1111/evo.13878

Springer, Y. P., Jarnevich, C. S., Barnett, D. T., Monaghan, A. J., & Eisen, R. J. (2015). Modeling the present and future geographic distribution of the lone star tick, *Amblyomma americanum* (ixodida: Ixodidae), in the continental United States. *American Journal of Tropical Medicine and Hygiene*, *93*(4), 875–890. https://doi.org/10.4269/ajtmh.15-0330

Thompson, J. R., Carpenter, D. N., Cogbill, C. V., & Foster, D. R. (2013). Four centuries of change in Northeastern United States forests. *PLoS One*, *8*(9), e72540 https://doi.org/10.1371/journal.pone.0072540

Thuiller, W., Georges, D., & Engler, R. (2013). biomod2: Ensemble platform for species distribution modeling. *R Package Version*.

Trout, R. T., Steelman, C. D., & Szalanski, A. L. (2010). Population genetics of *Amblyomma americanum* (Acari: Ixodidae) collected from Arkansas. *Journal of Medical Entomology*, *47*(2), 152–161. https://doi.org/10.1603/ME09106

White, T. A., Perkins, S. E., Heckel, G., & Searle, J. B. (2013). Adaptive evolution during an ongoing range expansion: The invasive bank vole (*Myodes glareolus*) in Ireland. *Molecular Ecology*, *22*(11), 2971–2985. https://doi.org/10.1111/mec.12343

Willi, Y., Van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 433–458. https://doi.org/10.1146/annurev.ecolsys.37.091305.110145

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114–138.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Lado P, Smith ML, Carstens BC, Klompen H. Population genetic structure and demographic history of the lone star tick, *Amblyomma americanum* (Ixodida: Ixodidae): New evidence supporting old records. *Mol Ecol*. 2020;00:1–14. https://doi.org/10.1111/mec.15524