



OPEN

Northern and southern blacklegged (deer) ticks are genetically distinct with different histories and Lyme spirochete infection rates

Guang Xu¹✉, Ben Wielstra^{2,3} & Stephen M. Rich¹

Lyme borreliosis (LB) is the archetypal emerging zoonosis and is dependent on transmission by ticks in the genus *Ixodes*. Understanding the origin, maintenance, and spread of these ticks contributes much to our understanding of the spread of LB and other disease agents borne by these ticks. We collected 1232 *Ixodes scapularis* ticks from 17 east coast sites ranging from New Hampshire to Florida and used mtDNA, three nuclear genetic loci, and incorporated Bayesian analyses to resolve geographically distinct tick populations and compare their demographic histories. A sparse, stable, and genetically diverse population of ticks in the Southeastern US, that is rarely infected with the agent of LB is genetically distinct from an abundant, expanding, and comparatively uniform population in the Northeast, where epidemic LB now constitutes the most important vector borne disease in the United States. The contrasting geography and demography of tick populations, interpreted in the context of the geological history of the region, suggests that during the last glacial period such ticks occupied distinct refugia, with only the northern-most site of refuge giving rise to those ticks and pathogens now fueling the epidemic.

Lyme disease is the most commonly reported vector borne disease in North America though its risk varies widely by geographic region¹. *Ixodes scapularis* is the main vector tick of *Borrelia burgdorferi*, the bacterium that causes Lyme disease. In the United States, most human Lyme disease cases are reported in the Northeast and upper Midwest, while cases in western and southern regions of the country are low². This contrast is also evident in the entomological record where the frequency of *B. burgdorferi* is 2–8% in the main vector *I. scapularis* adult ticks in the South^{3,4}, compared to infection rates averaging 50% in New England⁵. Environmental factors may affect survival rates of immature *I. scapularis* and the geographical distribution of Lyme disease⁶. In addition to the persistent difference in Lyme infection rates, there are major differences of questing behavior, human attacking rate, host preferences, and survival rate between northern and southern populations of *I. scapularis* ticks^{7–11}. There are also vast differences in abundance, with *I. scapularis* tick densities in the South being much lower than in the North^{4,12}.

Due to these phenotypic variations, it is not surprising that large genetic differences occur between north and south *I. scapularis* populations. In 1985, the original description of the north population as a new species *I. dammini* was driven in part by the report of an abundant *I. ricinus*-like tick with northern distribution several hundred miles disjunct from that of the southern *I. scapularis*¹³. However, the morphological differences between northern and southern populations were cryptic, tending to manifest only during particular life stages. In 1993, Oliver and colleagues suggested that *I. dammini* should be reduced to a junior synonym of *I. scapularis*¹⁴. Population genetic studies showed no differences in chromosome morphology or allozyme frequencies among populations^{14,15}. The nuclear ribosomal DNA sequences of the ITS-1 and ITS-2 regions were also continuously distributed among populations^{16,17}. However, the distribution of mitochondrial haplotypes is distinct and forms two clades: (1) the American clade, found in the northeastern US down to the Carolinas, and (2) the Southern clade, which overlaps in the Carolinas and extends into the southeastern states. More genetic variation and diversity have been observed in ticks collected from southern states^{18–21} and extremely abundant single nucleotide

¹Department of Microbiology, University of Massachusetts, Amherst, United States of America. ²Institute of Biology Leiden, Leiden University, Leiden, The Netherlands. ³Naturalis Biodiversity Center, Leiden, The Netherlands. ✉e-mail: gxu@umass.edu

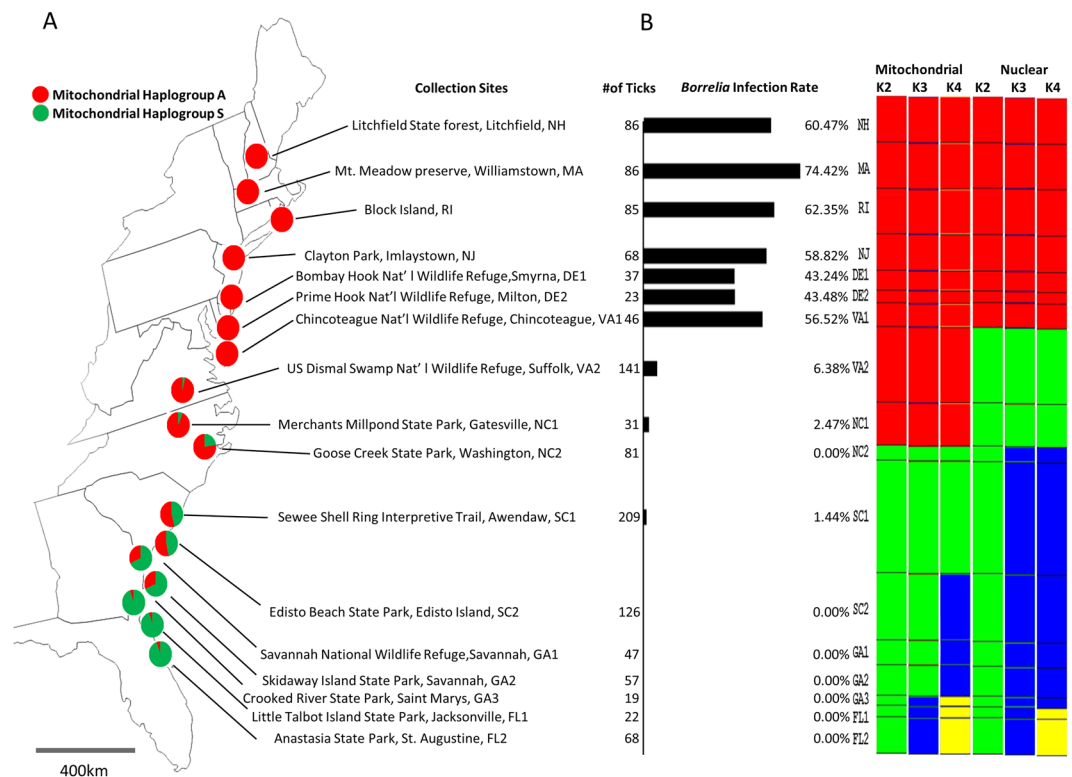


Figure 1. Collection sites, percentage of mitochondrial Haplogroups, *Borrelia* infection rates and BAPS population structures. **(A)** Colored circles are 17 collection sites of 1232 ticks from eastern US. Red and green colors show the percentage of mitochondrial Haplogroup A and S. **(B)** Total tick numbers in each collection site with *Borrelia* infection rate. Population structure analysis is based on mitochondrial and 3 nuclear genes using BAPS. Each color represents one population.

polymorphisms (SNPs) have been found in *I. scapularis*²². SNPs analysis for ten ticks showed that samples from New Jersey and Virginia formed a homogeneous group with low genetic diversity, whereas southern ticks from Georgia and Mississippi consisted of two separate groups, each with high genetic diversity. SNPs analysis also revealed a predominantly North to South gene flow²³.

While the population genetic structure of blacklegged ticks is a fundamental basis for understanding and predicting Lyme disease epidemiology, relevant studies have mainly utilized phylogenetic methods based on mitochondrial markers from a small number of individuals at limited collection sites. As a result, we still do not know the border between the northern and southern populations. Studies of nuclear, segregating genetic markers that could elucidate the evidence for genetic exchange across the broad geographic range of the east coast of the United States are limited²⁴. Although ticks from northern and southern populations appear to be equally competent *B. burgdorferi* vectors in laboratory conditions^{25,26}, data on the prevalence of the Lyme spirochete has not been linked to mitochondrial Southern clade ticks. The present study is based on mtDNA and three nuclear genetic loci and incorporated Bayesian analyses to detect evidence of population structure and geographically distinct tick populations. We have also used ecological niche modeling to determine the historical distribution and spread of these ticks. We found that northeastern (northern) and southeastern (southern) *I. scapularis* tick populations are genetically distinct with different histories and Lyme spirochete infection rates.

Results

Species and PCR amplification. A total of 1236 adult ticks were flagged from vegetation from 17 east coast sites in November to December 2007 (Fig. 1). Four ticks collected at Awendaw, South Carolina were not identified as the species of interest (*I. affinis* N = 3; and *I. brunneus* N = 1) based on the mitochondrial 16S sequences. Fragments of mitochondrial 16S ribosomal RNA (mt DNA 16S), nuclear (nuc) Calreticulin (CRT), Glycerol-3-phosphate dehydrogenase (G3PD) and Isocitrate dehydrogenase (IDH) genes were amplified for 1232 adult *I. scapularis* ticks.

Phylogeny and distribution of haplotypes. In agreement with the findings of previous studies^{18,19}, the phylogenetic analysis of 16S gene revealed two mitochondrial haplogroups (Fig. 2). The proportion and geographic distribution of these two haplogroups clearly shows that one is distributed mainly in the North, but extending to the South, while the second is restricted only to the Southeast locations. From Suffolk, Virginia to Florida the haplotype ratios appear as a latitudinal cline, with the frequency of the southern haplogroup starting

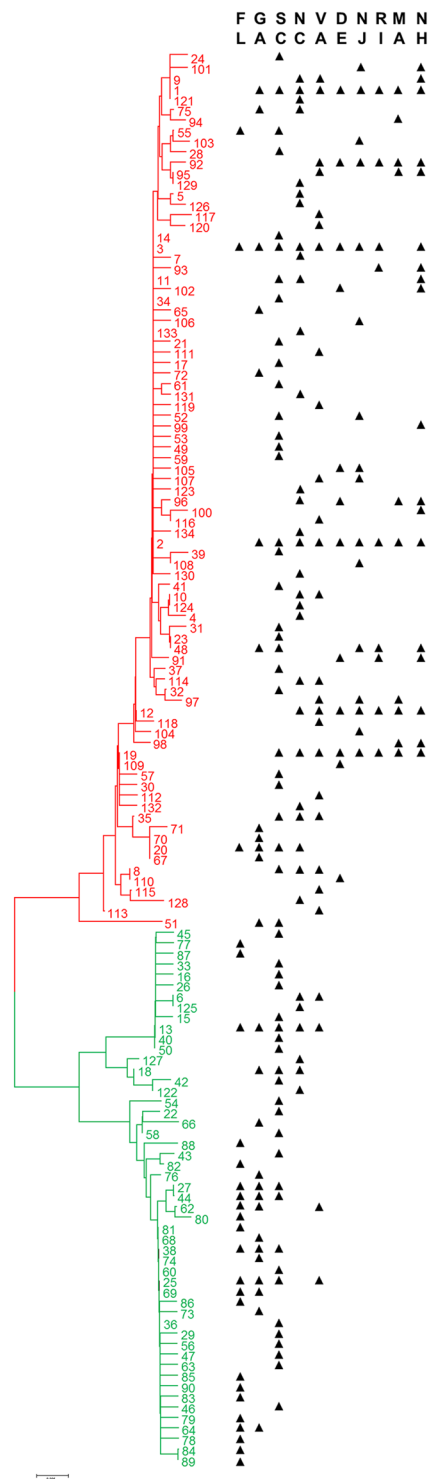


Figure 2. Phylogenetic tree of mtDNA haplotypes. The red color represents Haplogroup A and the green color Haplogroup S. Triangles show geographic distribution of each haplotype (Software used to generate the maximum likelihood (ML) tree was PAUP Version 4.0b10, <https://paup.phylosolutions.com>).

at 94.1% in St. Augustine, Florida and declining to 3.5% at Suffolk Virginia, and no individuals within that haplogroup occurring further north (Fig. 1).

Analysis of mitochondrial 16S sequences yielded 134 unique haplotypes. The greater number of haplotypes was detected in the southern locations (Fig. 2) (Supplementary S2 Table 1). The cluster analysis of mitochondrial haplotypes showed two groups. The samples from Florida, Georgia and South Carolina formed one group, and the samples from North Carolina, Virginia, Delaware, New Jersey, Rhode Island, Massachusetts and New

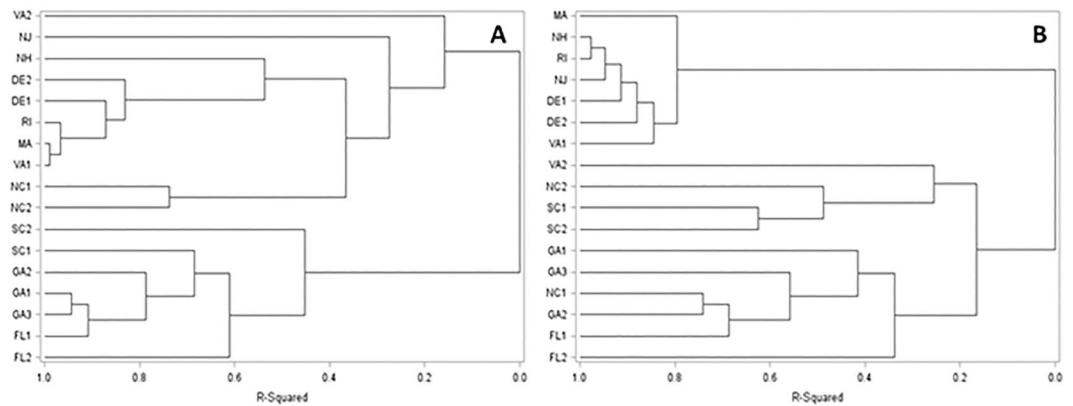


Figure 3. Cluster Analysis of mtDNA (A) and nucDNA (B) haplotypes distribution among 17 east coast collection sites (New Hampshire: NH; Massachusetts: MA; Rhode Island: RI; New Jersey: NJ; Delaware: DE1 and DE2; Virginia: VA1 and VA2; North Carolina: NC1 and NC2; South Carolina: SC1 and SC2; Georgia: GA1, GA2 and GA3; Florida: FL2 and FL2), (cluster analysis was performed by SAS software Version 9.0, <https://www.sas.com>).

Hampshire formed another group. Clearly, the 16S haplotype distributions are different between the South and the North (Fig. 2).

DNA sequences of 3 nuclear genes yielded 342, 201 and 124 unique haplotypes for CRT, IDH and G3PD, respectively. G3PD, CRT and IDH haplotype distributions also showed significant differences between the South and the North (Fig. 3). However, compared to the mitochondrial gene haplotype distribution, the cluster analysis of nuclear gene haplotype distribution suggested the boundary between southern and northern populations to be positioned more to the north, allocating the two Virginia populations to the southern population.

Population comparison. The mitochondrial 16S gene, and nuclear genes CRT and IDH supported a separation between northern and southern populations when pairwise differences were used to analyze the population relationships (See Supplementary S1 Fig. A to D). The evaluated F_{st} values indicated limited gene flow between the southern and the northern populations. However, pairwise comparisons of F_{st} and Nei's distance of G3PD gene showed no significant correlations between geographic locations and pairwise differentiation. Strong population structure was detected using AMOVA among populations, both within and between the North and South (Supplementary S2 Table 2).

Demographic history. According to the mismatch distribution analysis of 4 individual genes, the spatial expansion model could not be rejected for all 17 populations. However, the demographic model showed different results (Supplementary S2 Table 3). A mismatch distribution is the frequency distribution of the observed number of differences between pairs of haplotypes. The mitochondrial 16S gene suggested a sudden expansion for northern populations and stable tick populations in Florida, Georgia and South Carolina and North Carolina. The mismatch distribution of CRT and IDH sequences also indicated a different demographic history between southern and northern populations.

Bayesian analysis of population structure. The seventeen populations were clustered using Structure and BAPS software population mixture analysis. Both Structure and BAPS software suggest $K = 4$ for nuclear genes and $K = 2$ for mtDNA gene. The mitochondrial and combined nuclear data all suggested two major black-legged tick population clusters along the eastern U.S. coast, with a single population containing NH, MA, RI, NJ, DE1, DE2 and VA1 and 2–3 possible subpopulations in the Southern area (Fig. 1). The mitochondrial and nuclear genes identified different boundaries of these two populations. Ticks (VA2) from Suffolk, Virginia belonged to the northern population based on mtDNA but to the southern group based on nuclear DNA.

Borrelia infection rate. The *Borrelia* infection rate ranged from 43.24 to 74.42% in NH, MA, RI, NJ, DE1, DE2 and VA1 tick populations (Fig. 1). The highest infection rate was 74.42% in Massachusetts. The infection rates of the two Virginia populations that belong to the southern population based on nuclear DNA but to the northern population based on mtDNA were very different, with an infection rate of 56.52% in Chincoteague (VA1) population and 6.38% in the Suffolk (VA2) population. In populations allocated to the southern group based on both mtDNA and nuclear DNA the infection rate was very low, ranging from 0 to 2.47%. The *Borrelia* infection was strictly limited to individuals with mitochondrial Haplogroup A; no Haplogroup S individuals were found positive for *Borrelia*.

Ecological niche model. Projection of the ecological niche models (AUC values for the northern and southern population were 0.94 and 0.96) on the two Last Glacial Maximum climate reconstructions (CCSM and MIROC) suggested that the northern tick population was derived from a small refuge located just south of the Pleistocene ice sheet on the Atlantic continental shelf off the current Delmarva Peninsula. The southern *I.*

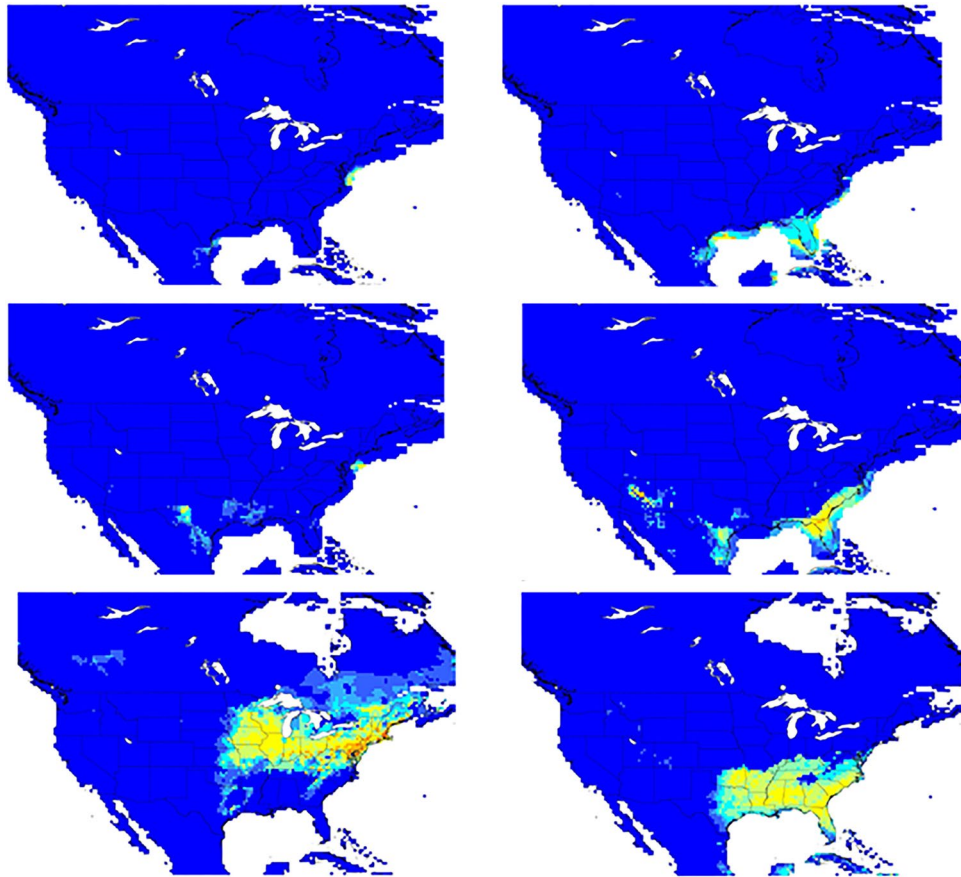


Figure 4. The predicted distribution for the northern (left) and southern population (right) of *Ixodes scapularis* at the Last Glacial Maximum (top), the mid-Holocene (middle) and the present day (bottom), based on ecological niche modeling. The warmer a grid cell's color, the higher its predicted suitability (Software used to draw the map was Maxent Version 3.3.3k, https://biodiversityinformatics.amnh.org/open_source/maxent/).

scapularis population probably originated from a suitable refuge located in the southern coastal plain of today's Florida, Georgia and South Carolina (Fig. 4).

Discussion

Based on haplotype distribution, population comparison, phylogenetic analysis, demographic history and Bayesian analysis of population structure, we conclude that there are two distinct blacklegged tick populations (northern and southern) along the US eastern coast, with distinct *Borrelia* infection rates and post-glacial histories.

In agreement with previous mtDNA studies^{18–20}, this study not only suggested that blacklegged ticks are geographically structured along the east US coast, but also suggested that two distinct clades meet at North Carolina and Virginia. The southern *I. scapularis* population is characterized by greater heterogeneity and absence of widely distributed haplotypes. In contrast to the results from the ITS-2 gene¹⁷, analysis of three nuclear genes confirms the existence of two major structured populations of *I. scapularis* ticks. Both mitochondrial and nuclear genes are comparatively homogeneous in northern *I. scapularis* populations, but more diverse in southern *I. scapularis* populations. We found nuclear genes suggested a boundary between the northern and southern population in Virginia, whereas mtDNA suggested this boundary to be positioned in North Carolina.

The demographic model strongly supports a recent and rapid increase in population size for all Northeastern sites. By contrast, only several southern sites (Florida, Georgia and South Carolina) showed no evidence of changes in historical population size. The ecological niche model suggests that there were two possible distinct refugia in eastern North America during the Last Glacial Maximum. The northern *I. scapularis* populations are suggested to be derived from a small refugia located to the south of the Pleistocene ice sheet and on the Atlantic continental shelf off the current Delmarva Peninsula. Although this region has been submerged as sea level raised from melted glacial ice, it was a broad coastal plain and well covered with land vegetation during the Last Glacial Maximum²⁷. The southern *I. scapularis* population may have originated from a suitable refuge that is more to the South, located in the Southern coastal plain. The later expansion of the southern population seemed to occur through much of the Southeastern quadrant of the United States. However, since *I. scapularis* is expanding its range and has recently shown tolerance for colder conditions in northern US and southern Canada, the current distribution of *I. scapularis* may not represent the full spectrum of its habitable environment, and the inferred

historical range might be underestimated. The New York–Connecticut region²⁰ and the areas further north of Delmarva have been suggested as a likely location of refugia populations of *I. scapularis*²⁸.

We detected a homogenous northern population and a heterogeneous southern population. The heterogeneity of the southern population is due to introgression from the northern population. In agreement with previous studies^{18,19}, we found that the two major mitochondrial haplogroups co-occur throughout the south. All northern ticks carry only Haplogroup A, while Haplogroups A and S occur mixed throughout the South. This has previously been explained as evidence of continuous gene flow among populations and a northward movement of the southern population³. We, however, found that the relative frequency of A and S haplogroups is a latitudinal cline with S increasing towards the South. The introgression of the northern population mitochondrial DNA might suggest that gene flow between two tick populations has migrated southwards. Positive selection or migrating birds may play an important role for the directional gene flow because they would carry ticks southward more effectively than northward²⁸. However, the expansion of *I. scapularis* includes southward as well as northward expansion. For example, the recent northward invasion of *I. scapularis* and *B. burgdorferi* in Canada has occurred via the introduction by migratory birds²⁹. The environmental conditions and selective pressure from warmer temperatures may change the host-seeking behavior of southern tick population⁶. Alternatively, the northern mtDNA in the Southeastern population may represent a genetic footprint left after the southern population displaced an expanding northern population^{30,31}. Bayesian Analysis of Population Structure revealed two populations with a possible hybridization zone located in Virginia.

We found that the southern *I. scapularis* populations appear to be at a demographic equilibrium, while northern populations have gone through a recent demographic expansion, especially after glacial retreat. However, human activities have accelerated the expanding of *I. scapularis* ticks in the eastern United States and southern Canada during the 20th century³². *I. scapularis* was first collected on Naushon Island, Massachusetts in 1920. With the return of white-tailed deer to other areas of the Northeast and Midwest, the range of *I. scapularis* began expanding²⁸. During the 1960s, focal *I. scapularis* infestations were recognized on Long Island and Nantucket Island in the Northeast and in northwestern Wisconsin. By the mid-1970s, the known range of *I. scapularis* had expanded, particularly extending in a southerly direction along the middle-Atlantic coast²⁸. Within 100 years, this tick has largely spread throughout most of the eastern United States and into southern Canada³². The expansion of *I. scapularis* and *B. burgdorferi* is determined by distinct environmental factors: climate, habitat and host range and abundance^{29,33,34}. In this study, only climatic factors were included in the ecological niche models.

The most interesting finding is that a high infection rate of *Borrelia* (43–74%) is mostly linked to the northern *I. scapularis* population and where *Borrelia* does occur in the southern *I. scapularis* population (as defined by nuclear DNA), it is always found linked to the northern mtDNA type. The infection rate drops substantially to 0–6% with a change of the nuclear DNA defined population boundary in Virginia. Whether the differences in the infection rate is environmentally determined, is a heritable trait for two populations or is a combination of genetic traits and hosts needs further study. For example, warmer temperatures and a lower density of host population may change the host-seeking behavior of southern tick population and eventually lower the incidence of Lyme disease in areas at low altitudes⁶. However, environmental factors alone cannot explain the virtual absence of infection rate in *I. scapularis*. Although different hosts were believed to be the reason for low infection rates in the South, a comparative study between two sites in eastern Virginia³⁵ showed the infection of white-footed mice is 66.5% in the coastal site and 15.5% in the inland sites. *Borrelia* infection is 35.7% and 0% in the nymphal ticks removed from these mice in the coastal site and in the inland sites, respectively. Recently, the absence of detectable *B. burgdorferi* infection was found among blacklegged ticks in North Carolina³⁶ and Tennessee³⁷. These may suggest a genetic basis of tick behaviors that influence the epidemiology of vector-borne pathogens. It is very likely that the Northeastern tick population is the principal tick population responsible for transmitting Lyme disease. This same population is solely carrying two zoonotic agents in the Eastern US: *Anaplasma phagocytophilum* and *Babesia microti*¹³.

In this paper, we only sampled adult ticks from eastern coastal sites. Tick-borne disease in the Midwest of the US and Canada is increasing as *I. scapularis* populations continue to spread and grow^{29,38}. It is worth noting that tick population ecology and spirochete transmission patterns are different between the northeast and northern Midwest. The two foci of *B. burgdorferi* in the northeastern and midwestern have a shared past and once belonged to an admixed population but are now isolated^{11,39,40}. Transport of *B. burgdorferi* between northeastern and midwestern *I. scapularis* populations appears to be relatively infrequent.

The suggestion of a genetic basis to the transmission of *Borrelia* by Northeastern populations of *I. scapularis* is of significance for public health. We believe that, currently, the risk of acquiring Lyme disease through *I. scapularis* population is low in the South. First because they are low in density and they rarely bite humans^{7–9}; more importantly, this population is not likely to be carrying the Lyme spirochete. However, this does not mean that there are no enzootic cycles of spirochetes in the South. Enzootic transmission of Lyme disease spirochetes among rodents and ticks has been documented in Southern and South-central states. Although they rarely bite humans, the sympatric tick species *I. affinis*, *I. minor* and *I. dentatus* may maintain a different enzootic cycle in the South⁴¹. In this study, two of three collected *I. affinis* ticks are *Borrelia* positive. In North Carolina, *Borrelia* DNA was detected in 63.2% of 155 *I. affinis*, but in 0% of 298 *I. scapularis*³⁶; so-called “*Ixodes scapularis*” with a 35% *Borrelia* positive rate was confirmed as a non-accurate identification of *I. affinis*⁴². The current and potentially enhanced future role of the southern blacklegged tick population *I. scapularis* in Lyme epidemiology requires additional studies. Monitoring dynamics of the two populations will help us to predict future changes of tick and tick-borne pathogens.

Materials and Methods

Sampling natural tick populations. Adult ticks were collected from 17 sites in 10 states (New Hampshire: NH; Massachusetts: MA; Rhode Island: RI; New Jersey: NJ; Delaware: DE1 and DE2; Virginia: VA1 and VA2; North Carolina: NC1 and NC2; South Carolina: SC1 and SC2; Georgia: GA1, GA2 and GA3; Florida: FL2 and FL2) throughout the eastern United States by dragging a 1-m² white flannel flag through vegetation in November - December 2007 (Fig. 1). Ticks were kept alive until we returned to the laboratory, then frozen individually at -80 °C. Total DNA was extracted from individual ticks using the Epicenter Master Complete DNA & RNA Purification Kits (Epicenter Technologies, Madison, WI), following the manufacturer protocols.

Gene markers. We examined multiple genetic loci, including partial sequences of one mitochondrial gene and three nuclear genes. Both mitochondrial and nuclear gene primers were developed from the available sequences in GenBank. Primers TK16S_105F (5'-CGGTCTGAACTCAGATCATGTA-3') and TK16S_546RC (5'-AATTGCTGTGGTATTTGACTATAC-3') were used to amplify a mitochondrial 16S (441 bp) region. PCR was performed at 95 °C 15 sec, 45 °C 15 sec and 72 °C 60 sec for 40 cycles. Primers CRT_849F (5'-TGGACGAGCCGATGGGTA-3') and CRT_1657RC (5'-GGGCTGTACTCTGGGTATCGA-3'), G3PD_3574F (5'-GCGGTGGCACCTTGGTAG-3') and G3PD_4307RC (5'-CCGTCCTTGTCCCGTTTCT-3') and IDH_707F (5'-GGTACTTCGTGACCCTCTGC-3') and IDH_1436RC (5'-CGACCATAGTTGCGACCAGAC-3') were designed for genes encoding Calreticulin (CRT, 808 bp), Glycerol-3-phosphate dehydrogenase (G3PD, 733 bp) and Isocitrate dehydrogenase (IDH, 729 bp), respectively. These three nuclear genes were amplified at 95 °C 15 sec, 57 °C 15 sec and 72 °C 60 sec for 40 cycles. The PCR products were treated using Exosap-IT then sequenced with an ABI 3130 sequencer following the manufacturer protocols.

Borrelia burgdorferi real-time PCR Detection. Taqman real-time PCR assays were performed in a duplex format (to detect DNA of tick and *B. burgdorferi*) with a reaction volume of 20 µl, by using the Brilliant II QPCR Master Mix in a Stratagene MX3000P QPCR System (Agilent, La Jolla, Calif.). The detection of tick DNA served as an internal control with 300 nM of primers Tick16S_F (5'-AATACTCTAGGGATAACAGCGTAATAATTTT-3') and Tick16S_R (5'-CGGTCTGAACTCAGATCAAGTAGGA-3'), and 125 nM of Tick16S_Probe (FAM-5'-AAATAGTTTGGCAGCTCGATGTTGGATTAGGAT 3'-BHQ1). The detection of *B. burgdorferi* was modified from (Courtney *et al.* 2004) with 700 nM primers Borrelia23S_F (5'-CGAGTCTTAAAAGGGCGATTTAGT-3') and Borrelia23S_R (5'-GCTTCAGCCTGGCCATAAATAG-3'), and 300 nM of Borrelia_Probe (HEX-5'-AGATGTGGTAGACCCGAAGCCGAGTG-3'-BHQ1). Cycling conditions included an initial activation of the Taq DNA polymerase at 95 °C for 10 min, followed by 40 cycles of a 15 s denaturation at 95 °C and a 1 min annealing-extension step at 60 °C.

Phylogeny and population structure. DNA sequences were aligned using the ClustalW software. Individual mitochondrial haplotypes were identified and converted to Arlequin format using FaBox⁴³. Phylogenetic analyses were conducted using maximum likelihood (ML) method in PAUP 4.0b10⁴⁴.

Nuclear haplotypes for heterozygotes were reconstructed using PHASE 2.1⁴⁵ with default parameters. Genetic structure was examined by an analysis of molecular variance (AMOVA), implemented in Arlequin version 3.1⁴⁶. Structure analysis was performed using Structure⁴⁷ and BAPS version 5.3⁴⁸. Structure analysis was carried out with 10 independent runs per each K value (K = 1-17), with a 100,000 burn-in period and 900,000 Markov chain Monte Carlo iterations. These analyses were performed using a model with admixture, correlated allele frequencies and with no a priori information on the sample location of individuals. The ΔK approach was used to estimate the most likely K value. For BAPS, a spatial clustering algorithm and a mixture analysis of individuals without geographic information was chosen, with 10 replicates from K = 2 to K = 17 ran. When we run clustering of 17 sample populations, K = 4 was calculated based on nuclear genes and K = 2 was justified based on mtDNA. We, therefore, set the maximum number of clusters to 2, 3 and 4 for BAPS.

Locality and climate data used for ecological niche modeling. We composed a data set of 1423 county-level reported occurrences of blacklegged ticks (See Supplementary S3)^{49,50}. These occurrences were separated into two partitions based on genetic results: northern *I. scapularis* population (n = 825) and southern *I. scapularis* population (n = 598). For climate layers we used bioclimatic variables for the present, the Mid-Holocene (~6Ka) and the Last Glacial Maximum (~21Ka) available from the WorldClim database 1.4 (<http://www.worldclim.org>) and resampled these to a 25 arcminute resolution (c. 50 × 50 km). We selected a subset of bioclimatic variables that show a Pearson's correlation of $r < 0.7$, and are deemed biologically meaningful based on life history knowledge of the model system, in this case, seasonality: bio2 = mean diurnal temperature range (mean of the monthly difference between the minimum and maximum temperature), bio8 = mean temperature of the wettest quarter, bio9 = mean temperature of the driest quarter, bio16 = precipitation of the wettest quarter, and bio17 = precipitation of the driest quarter^{51,52} (Supplementary S2 Table 4).

Ecological niche modeling. We created ecological niche models for northern and southern *I. scapularis* populations with Maxent 3.3.3k⁵³. We restricted the feature type to hinge features as this facilitates extrapolation⁵⁴. The area between 20 and 55 degrees latitude and -60 and -130 degrees longitude was chosen as environmental background. Ecological niche models were projected on the current and past climate layers. Statistical significance of ecological niche models was confirmed⁵⁵ by testing their AUC values against a null model based on random localities created with ENMTools 1.3⁵⁶.

Received: 27 February 2020; Accepted: 4 June 2020;
Published online: 24 June 2020

References

- Bacon, R. M., Kugeler, K. J., Griffith, K. S. & Mead, P. S. Lyme disease - United States, 2003-2005 (Reprinted from MMWR, vol 56, pg 573-576, 2007). *Jama-Journal of the American Medical Association* **298**, 278–279 (2007).
- Diuk-Wasser, M. A. *et al.* Human Risk of Infection with *Borrelia burgdorferi*, the Lyme Disease Agent, in Eastern United States. *Am. J. Trop. Med. Hyg.* **86**, 320–327, <https://doi.org/10.4269/ajtmh.2012.11-0395> (2012).
- Oliver, J. H. Lyme borreliosis in the southern United States: A review. *J. Parasitol.* **82**, 926–935, <https://doi.org/10.2307/3284201> (1996).
- Diuk-Wasser, M. A. *et al.* Field and climate-based model for predicting the density of host-seeking nymphal *Ixodes scapularis*, an important vector of tick-borne disease agents in the eastern United States. *Global Ecol. Biogeogr.* **19**, 504–514 (2010).
- Barbour, A. G. & Fish, D. The Biological and Social Phenomenon of Lyme Disease. *Science* **260**, 1610–1616, <https://doi.org/10.1126/science.8503006> (1993).
- Ginsberg, H. S. *et al.* Environmental Factors Affecting Survival of Immature *Ixodes scapularis* and Implications for Geographical Distribution of Lyme Disease: The Climate/Behavior Hypothesis. *PLoS One* **12**, e0168723, <https://doi.org/10.1371/journal.pone.0168723> (2017).
- Falco, R. C. & Fish, D. Ticks Parasitizing Humans in a Lyme-Disease Endemic Area of Southern New-York State. *Am. J. Epidemiol.* **128**, 1146–1152 (1988).
- Felz, M. W., Durden, L. A. & Oliver, J. H. Ticks parasitizing humans in Georgia and South Carolina. *J. Parasitol.* **82**, 505–508 (1996).
- Merten, H. A. & Durden, L. A. A state-by-state survey of ticks recorded from humans in the United States. *J. Vector Ecol.* **25**, 102–113 (2000).
- Arsnoe, I., Tsao, J. I. & Hickling, G. J. Nymphal *Ixodes scapularis* questing behavior explains geographic variation in Lyme borreliosis risk in the eastern United States. *Ticks Tick Borne Dis.* **10**, 553–563 (2019).
- Gatewood, A. G. *et al.* Climate and Tick Seasonality Are Predictors of *Borrelia burgdorferi* Genotype Distribution. *Appl. Environ. Microbiol.* **75**, 2476–2483 (2009).
- Stromdahl, E. Y. & Hickling, G. J. Beyond Lyme: Aetiology of Tick-borne Human Diseases with Emphasis on the South-Eastern United States. *Zoonoses Public Health.* **59**, 48–64 (2012).
- Spielman, A., Clifford, C. M., Piesman, J. & Corwin, M. D. Human Babesiosis on Nantucket-Island, USA - Description of the Vector, *Ixodes (Ixodes) dammini*, n sp-Sp (Acarina, Ixodidae). *J. Med. Entomol.* **15**, 218–234 (1979).
- Oliver, J. H. *et al.* Conspecificity of the Ticks *Ixodes scapularis* and *Ixodes dammini* (Acari, Ixodidae). *J. Med. Entomol.* **30**, 54–63 (1993).
- Chen, C. S., Munderloh, U. G. & Kurtti, T. J. Cytogenetic Characteristics of Cell-Lines from *Ixodes scapularis* (Acari, Ixodidae). *J. Med. Entomol.* **31**, 425–434 (1994).
- Mclain, D. K., Wesson, D. M., Oliver, J. H. & Collins, F. H. Variation in Ribosomal DNA Internal Transcribed Spacers-1 among Eastern Populations of *Ixodes scapularis* (Acari, Ixodidae). *J. Med. Entomol.* **32**, 353–360 (1995).
- Wesson, D. M., Mclain, D. K., Oliver, J. H., Piesman, J. & Collins, F. H. Investigation of the Validity of Species Status of *Ixodes dammini* (Acari, Ixodidae) Using rDNA. *Proc. Natl. Acad. Sci. USA* **90**, 10221–10225 (1993).
- Norris, D. E., Klompen, J. S. H., Keirans, J. E. & Black, W. C. Population genetics of *Ixodes scapularis* (Acari: Ixodidae) based on mitochondrial 16S and 12S genes. *J. Med. Entomol.* **33**, 78–89 (1996).
- Rich, S. M. *et al.* Distribution of the *Ixodes ricinus*-Like Ticks of Eastern North-America. *Proc. Natl. Acad. Sci. USA* **92**, 6284–6288, <https://doi.org/10.1073/pnas.92.14.6284> (1995).
- Qiu, W. G., Dykhuizen, D. E., Acosta, M. S. & Luft, B. J. Geographic uniformity of the Lyme disease spirochete (*Borrelia burgdorferi*) and its shared history with tick vector (*Ixodes scapularis*) in the northeastern United States. *Genetics* **160**, 833–849 (2002).
- Humphrey, P. T., Caporale, D. A. & Brisson, D. Uncoordinated Phylogeography of *Borrelia burgdorferi* and Its Tick Vector, *Ixodes scapularis*. *Evolution* **64**, 2653–2663, <https://doi.org/10.1111/j.1558-5646.2010.01001.x> (2010).
- Van Zee, J. *et al.* High SNP density in the blacklegged tick, *Ixodes scapularis*, the principal vector of Lyme disease spirochetes. *Ticks Tick Borne Dis.* **4**, 63–71, <https://doi.org/10.1016/j.ttbdis.2012.07.005> (2013).
- Van Zee, J., Piesman, J. E., Hojgaard, A. & Black, W. C. Nuclear Markers Reveal Predominantly North to South Gene Flow in *Ixodes scapularis*, the Tick Vector of the Lyme Disease Spirochete. *Plos One* **10** (2015).
- Araya-Anchetta, A., Busch, J. D., Scoles, G. A. & Wagner, D. M. Thirty years of tick population genetics: A comprehensive review. *Infect. Genet. Evol.* **29**, 164–179 (2015).
- Dolan, M. C., Maupin, G. O., Panella, N. A., Golde, W. T. & Piesman, J. Vector competence of *Ixodes scapularis*, *I. spinipalpis*, and *Dermacentor andersoni* (Acari: Ixodidae) in transmitting *Borrelia burgdorferi*, the etiologic agent of Lyme disease. *J. Med. Entomol.* **34**, 128–135 (1997).
- Sanders, F. H. & Oliver, J. H. Evaluation of *Ixodes scapularis*, *Amblyomma americanum*, and *Dermacentor variabilis* (Acari, Ixodidae) from Georgia as Vectors of a Florida Strain of the Lyme-Disease Spirochete, *Borrelia burgdorferi*. *J. Med. Entomol.* **32**, 402–406 (1995).
- Whitmore, F. C., Jr., Emery, K. O., Cooke, H. B. & Swift, D. J. Elephant teeth from the atlantic continental shelf. *Science* **156**, 1477–1481, doi:156/3781/1477 [pii]10.1126/science.156.3781.1477 (1967).
- Spielman, A., Wilson, M. L., Levine, J. F. & Piesman, J. Ecology of *Ixodes dammini* -Borne Human Babesiosis and Lyme Disease. *Annu. Rev. Entomol.* **30**, 439–460 (1985).
- Ogden, N. H., Mechai, S. & Margos, G. Changing geographic ranges of ticks and tick-borne pathogens: drivers, mechanisms and consequences for pathogen diversity. *Front. Cell Infect. Microbiol.* **3**, doi:ARTN 46 10.3389/fcimb.2013.00046 (2013).
- Currat, M., Ruedi, M., Petit, R. J. & Excoffier, L. The hidden side of invasions: Massive introgression by local genes. *Evolution* **62**, 1908–1920 (2008).
- Wielstra, B. Historical hybrid zone movement: More pervasive than appreciated. *J. Biogeogr.* **46**, 1300–1305, <https://doi.org/10.1111/jbi.13600> (2019).
- Sonenshine, D. E. Range Expansion of Tick Disease Vectors in North America: Implications for Spread of Tick-Borne Disease. *Int. J. Env. Res. Public Health* **15**, doi:ARTN 478 10.3390/ijerph15030478 (2018).
- Bouchard, C. *et al.* Associations between *Ixodes scapularis* ticks and small mammal hosts in a newly endemic zone in southeastern Canada: Implications for *Borrelia burgdorferi* transmission. *Ticks Tick Borne Dis.* **2**, 183–190, <https://doi.org/10.1016/j.ttbdis.2011.03.005> (2011).
- Leighton, P. A., Koffi, J. K., Pelcat, Y., Lindsay, L. R. & Ogden, N. H. Predicting the speed of tick invasion: an empirical model of range expansion for the Lyme disease vector *Ixodes scapularis* in Canada. *J. Appl. Ecol.* **49**, 457–464, <https://doi.org/10.1111/j.1365-2664.2012.02112.x> (2012).
- Sonenshine, D. E. *et al.* *Borrelia burgdorferi* in Eastern Virginia - Comparison between a Coastal and Inland Locality. *Am. J. Trop. Med. Hyg.* **53**, 123–133 (1995).
- Maggi, R. G., Reichelt, S., Toliver, M. & Engber, B. *Borrelia* species in *Ixodes affinis* and *Ixodes scapularis* ticks collected from the coastal plain of North Carolina. *Ticks Tick Borne Dis.* **1**, 168–171, <https://doi.org/10.1016/j.ttbdis.2010.08.003> (2010).
- Rosen, M. E. *et al.* *Borrelia burgdorferi* Not Detected in Widespread *Ixodes scapularis* (Acari: Ixodidae) Collected From White-Tailed Deer in Tennessee. *J. Med. Entomol.* **49**, 1473–1480 (2012).

38. Hamer, S. A., Hickling, G. J., Walker, E. D. & Tsao, J. I. Increased diversity of zoonotic pathogens and *Borrelia burgdorferi* strains in established versus incipient *Ixodes scapularis* populations across the Midwestern United States. *Infect. Genet. Evol.* **27**, 531–542, <https://doi.org/10.1016/j.meegid.2014.06.003> (2014).
39. Hoen, A. G. *et al.* Phylogeography of *Borrelia burgdorferi* in the eastern United States reflects multiple independent Lyme disease emergence events. *Proc. Natl. Acad. Sci. USA* **106**, 15013–15018, <https://doi.org/10.1073/pnas.0903810106> (2009).
40. Apperson, C. S., Levine, J. F., Evans, T. L., Braswell, A. & Heller, J. Relative Utilization of Reptiles and Rodents as Hosts by Immature *Ixodes scapularis* (Acari, Ixodidae) in the Coastal Plain of North Carolina, USA. *Exp. Appl. Acarol.* **17**, 719–731 (1993).
41. Oliver, J. H. *et al.* An enzootic transmission cycle of Lyme borreliosis spirochetes in the southeastern United States. *Proc. Natl. Acad. Sci. USA* **100**, 11642–11645, <https://doi.org/10.1073/pnas.1434553100> (2003).
42. Harrison, B. A. *et al.* Recent discovery of widespread *Ixodes affinis* (Acari: Ixodidae) distribution in North Carolina with implications for Lyme disease studies. *J. Vector Ecol.* **35**, 174–179 (2010).
43. Villesen, P. FaBox: an online toolbox for FASTA sequences. *Mol. Ecol. Notes* **7**, 965–968, <https://doi.org/10.1111/j.1471-8286.2007.01821.x> (2007).
44. Swofford, D. L. Paup - a Computer-Program for Phylogenetic Inference Using Maximum Parsimony. *J. Gen. Physiol.* **102**, A9–A9 (1993).
45. Stephens, M., Smith, N. J. & Donnelly, P. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* **68**, 978–989 (2001).
46. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567, <https://doi.org/10.1111/j.1755-0998.2010.02847.x> (2010).
47. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620, <https://doi.org/10.1111/j.1365-294X.2005.02553.x> (2005).
48. Corander, J., Marttinen, P., Siren, J. & Tang, J. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9**, doi:Artn 539 Doi 10.1186/1471-2105-9-539 (2008).
49. Dennis, D. T., Nekomoto, T. S., Victor, J. C., Paul, W. S. & Piesman, J. Reported distribution of *Ixodes scapularis* and in *Ixodes pacificus* (Acari: Ixodidae) in the United States. *J. Med. Entomol.* **35**, 629–638 (1998).
50. Eisen, R. J., Eisen, L. & Beard, C. B. County-Scale Distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the Continental United States. *J. Med. Entomol.* **53**, 349–386, <https://doi.org/10.1093/jme/tjv237> (2016).
51. Guisan, A. & Thuiller, W. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* **8**, 993–1009, <https://doi.org/10.1111/j.1461-0248.2005.00792.x> (2005).
52. Peterson, A. T. Ecological niche conservatism: a time-structured review of evidence. *J. Biogeogr.* **38**, 817–827, <https://doi.org/10.1111/j.1365-2699.2010.02456.x> (2011).
53. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **190**, 231–259, <https://doi.org/10.1016/j.ecolmodel.2005.03.026> (2006).
54. Elith, J., Kearney, M. & Phillips, S. The art of modelling range-shifting species. *Methods Ecol. Evol.* **1**, 330–342, <https://doi.org/10.1111/j.2041-210X.2010.00036.x> (2010).
55. Raes, N. & ter Steege, H. A null-model for significance testing of presence-only species distribution models. *Ecography* **30**, 727–736, <https://doi.org/10.1111/j.2007.0906-7590.05041.x> (2007).
56. Warren, D. L., Glor, R. E. & Turelli, M. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33**, 607–611, <https://doi.org/10.1111/j.1600-0587.2009.06142.x> (2010).

Author contributions

G.X. conceived and conducted the experiments. B.W. analyzed ecological niche modeling data. G.X. wrote the original draft - S.M.R. and B.W. revised and edited. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-67259-0>.

Correspondence and requests for materials should be addressed to G.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020