AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

EVOLUTIONARY AND GENOMIC MICROBIOLOGY

Demographic Expansions and the Emergence of Host Specialization in Genetically Distinct Ecotypes of the Tick-Transmitted Bacterium *Anaplasma phagocytophilum*

Matthew L. Aardema,[a,b] Nina V. Bates,[a] Qiana E. Archer,[a] Friederike D. von Loewenich[c]

[a]Department of Biology, Montclair State University, Montclair, New Jersey, USA
[b]Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, New York, USA
[c]Institute of Virology, University of Mainz, Mainz, Germany

ABSTRACT In Europe, genetically distinct ecotypes of the tick-vectored bacterium *Anaplasma phagocytophilum* circulate among mammals in three discrete enzootic cycles. To date, potential ecological factors that contributed to the emergence of these divergent ecotypes have been poorly studied. Here, we show that the ecotype that predominantly infects roe deer (*Capreolus capreolus*) is evolutionarily derived. Its divergence from a host generalist ancestor occurred after the last glacial maximum as mammal populations, including roe deer, recolonized the European mainland from southern refugia. We also provide evidence that this host specialist ecotype's effective population size (N_e) has tracked changes in the population of its roe deer host. Specifically, both host and bacterium have undergone substantial increases in N_e over the past 1,500 years. In contrast, we show that while it appears to have undergone a major population expansion starting ~3,500 years ago, in the past 500 years, the contemporary host generalist ecotype has experienced a substantial reduction in genetic diversity levels, possibly as a result of reduced opportunities for transmission between competent hosts.

IMPORTANCE The findings of this study reveal specific events important for the evolution of host specialization in a naturally occurring, obligately intracellular bacterial pathogen. Specifically, they show that host range shifts and the emergence of host specialization may occur during periods of population growth in a generalist ancestor. Our results also demonstrate the close correlation between demographic patterns in host and pathogen for a specialist system. These findings have important relevance for understanding the evolution of host range diversity. They may inform future work on host range dynamics, and they provide insights for understanding the emergence of pathogens that have human and veterinary health implications.

KEYWORDS host range, enzootic cycles, arthropod vector, *Capreolus capreolus*, *Rickettsiales*, *Anaplasmataceae*

In pathogenic organisms, many distinct ecological processes have the potential to alter the phylogenetic extent of viable hosts ("host range"). These may include changes in the population densities of hosts (1–3), host geographic-range shifts (3, 4), or the emergence of competitors (5, 6). However, examples of these processes influencing the host range of specific pathogens remain few (but see reference 3). It is critical to understand the nature of host range-altering processes to reduce or mitigate the potential for emergent pathogens to infect and cause illness in humans and domestic animals. Over half of the viral and bacterial pathogens studied are host specialists (7), although this finding may be due in part to sampling biases. Nonetheless, it is particularly important to examine the processes that contribute to the evolution of restricted host ranges in pathogenic organisms.

Editor Maia Kivisaar, University of Tartu

Address correspondence to Matthew L. Aardema, aardemam@montclair.edu.
The authors declare no conflict of interest.

Received 8 April 2022
Accepted 22 June 2022
Published 11 July 2022

The Gram-negative, tick-vectored bacterium *Anaplasma phagocytophilum* is an emerging zoonotic pathogen with human health and veterinary importance (8). In Europe, different ecotypes of this obligately intracellular pathogen are genetically distinct and circulate among mammals in three discrete enzootic cycles (9–11). One of these ecotypes primarily infects nest-living hosts, such as voles and shrews (9–12). A second ecotype is a host generalist that has been found to infect members of at least five separate mammal orders (Artiodactyla, Eulipotyphla, Perissodactyla, Carnivora, and Primates). Accordingly, *A. phagocytophilum* infections in livestock, companion animals, and humans in Europe are typically caused by this ecotype (9, 10, 13, 14). The third ecotype appears to be a host specialist, being found predominantly in roe deer (*Capreolus capreolus*). Despite their substantial difference in host ranges, the host generalist and roe deer specialist ecotypes are more closely aligned genetically, while the ecotype found in burrowing mammals is more evolutionarily distinct (9, 15). This burrowing-mammal ecotype is predominately transmitted by *Ixodes trianguliceps* ticks (12, 16), whereas both the host generalist and roe deer specialist ecotypes are chiefly transmitted by *Ixodes ricinus* in Europe (17).

Given that mammalian host range appears to be the primary ecological difference between the host generalist and roe deer specialist ecotypes, the evolution of these host associations was likely a key contributor to ecotype differentiation in this bacterium (10). During the last glacial maximum in Europe (~19,000 to 26,500 years ago) (18), roe deer were geographically restricted to a few southern refugia (19). As the glaciers receded, roe deer populations were able to spread north into Europe, substantially increasing their geographic range (20). An increase in the effective population size (N_e) of the mainland European roe deer population is estimated to have occurred between 3,932 and 7,919 years ago (19), possibly reflecting this range expansion, as well as increases in the continent-wide roe deer census population size. Although relationships between estimates of N_e and census population sizes have been shown to lack consistent correlation (21), in natural systems, sustained increases in N_e generally only occur as the result of corresponding increases in census population size (22, 23). More recently, the reductions of forest cover correlating with a growth in agriculture across Europe starting roughly 1,500 years ago led to a dramatic increase in the amount of edge habit (24–26). Habitat fragmentation likely benefitted roe deer populations and could have resulted in an increase in their population density (27–29), even as these same habitat alterations reduced the populations of other large mammals (30).

It is possible that increases in roe deer population density may have facilitated the emergence of the roe deer specialist *A. phagocytophilum* ecotype if host populations were sufficiently large enough to facilitate transmission predominantly between these ungulates. In this scenario, selection acting on different strains of *A. phagocytophilum* would have favored any that harbored genetic variation for improved infection capability in roe deer. Such selection may have ultimately led to a reduced ability to infect other mammals, either through antagonistic pleiotropy, whereby adaptations for improved roe deer infection came at the cost of infection ability for other species, or else through the accumulation, via genetic drift, of mutations that reduced infection capabilities in other hosts (31).

With the availability of population-level genetic data, it is possible to examine longer-term patterns of demographic change in conjunction with taxon differentiation (22). For example, the application of coalescent theory allows estimates of temporally changing effective population sizes to be made while assessing convergent phylogenetic relationships between samples (32, 33). Of particular relevance here, the extended Bayesian skyline plot (EBSP) (34) is useful when data from multiple loci are available, as this method can model gradual changes in N_e estimates between each coalescence interval, rather than less-realistic instantaneous changes, as in other Bayesian skyline methods.

Using demographic modeling in combination with additional, complementary analyses, we examined *A. phagocytophilum* ecotype divergence in relation to changes in host population size by testing several corresponding hypotheses. The first of these was that the roe deer specialist ecotype is evolutionarily derived, having evolved from a more-

generalist bacterial population that likely resembled the contemporary host generalist *A. phagocytophilum* population. Support for this hypothesis would indicate that the two *I. ricinus*-vectored ecotypes of *A. phagocytophilum* shared an ancestor that had a relatively broad host range and that the roe deer specialist ecotype diverged from this ancestral host generalist. Alternatively, it is possible that the contemporary host generalist ecotype could have evolved from a more specialized population, especially if reduced transmission creating genetic bottlenecks occurred at some point in its evolutionary past. Such events may have selected for strains that had broader invasion capabilities (35). The emergence of competition with other microbes could also have resulted in the evolution of the generalist ecotype from an ancestral host specialist (36).

The second hypothesis we tested was that divergence of the roe deer specialist ecotype occurred in conjunction with increases in the European roe deer population, possibly as these ungulates recolonized the continent following the last glacial maximum (~3,932 to 7,919 years ago) (19), or else during reductions in forest cover brought about by increases in agriculture (starting ~1,500 years ago) (24–26). Our third hypothesis was that after the evolution of the roe deer specialist ecotype, changes in this ecotype's effective population size would be closely correlated with changes in roe deer population sizes. Given the apparently larger N_e observed in the contemporary roe deer specialist ecotype (15), evidence in support of this third hypothesis would also suggest that host specialization has allowed this ecotype to reach higher population densities than the generalist. This could be attributed to greater within-host population sizes and/or an increased frequency of transmission between viable hosts (1, 37). Such an observation would support the view that host specialization can correlate with increases in overall fitness for a pathogenic bacterial population (38).

RESULTS

Sample clustering. To test the three hypotheses described above, we leveraged previously published DNA variation found in seven partial housekeeping gene sequences amplified from 278 *A. phagocytophilum* samples isolated from either a mammalian host or the *I. ricinus* vector (9). We initially performed a nonparametric principal-component analysis (PCA) to confirm sample assignment to one of the distinct mammal-infecting ecotypes previously identified (9). As predicted, our PCA clearly delineated three discrete clusters among the samples assessed (Fig. 1). The first principal component accounted for 37.2% of the genetic variation observed and separated the burrowing-mammal ecotype from the two *I. ricinus*-vectored ecotypes. The second principal component accounted for 8.9% of the genetic variation and separated the roe deer specialist ecotype from the host generalist ecotype. All samples clearly fell into one of the three clusters, with no ambiguous intermediate samples. Of the 278 total samples used in this analysis, there were 20 representing the burrowing-mammal ecotype, 24 representing the roe deer specialist ecotype, and 234 representing the host generalist ecotype. All subsequent analyses utilized the sample ecotype designations determined from this PCA.

Evolutionary pattern of ecotype divergence. We hypothesized that the contemporary roe deer specialist ecotype evolved from an ancestral host generalist ecotype and might accordingly be more derived relative to the contemporary host generalist ecotype. When comparing two closely related contemporary populations, the one that has diverged more from the shared common ancestor should harbor the signature of this greater divergence in patterns of DNA evolution, specifically the presence of derived alleles (39, 40). During host range shifts, an excess of derived alleles could emerge and be maintained either because they were directly selectively advantageous in the environment or because they were physically linked to other adaptive variants (41, 42). It is also possible that derived alleles could be fixed in a population through genetic drift if a population has undergone one or more severe bottlenecks, as may occur during transmission events between hosts (35, 43).

Previous work (9, 10, 15) and our nonparametric analysis (Fig. 1) indicate that the roe deer specialist and host generalist ecotypes are most closely related to one another

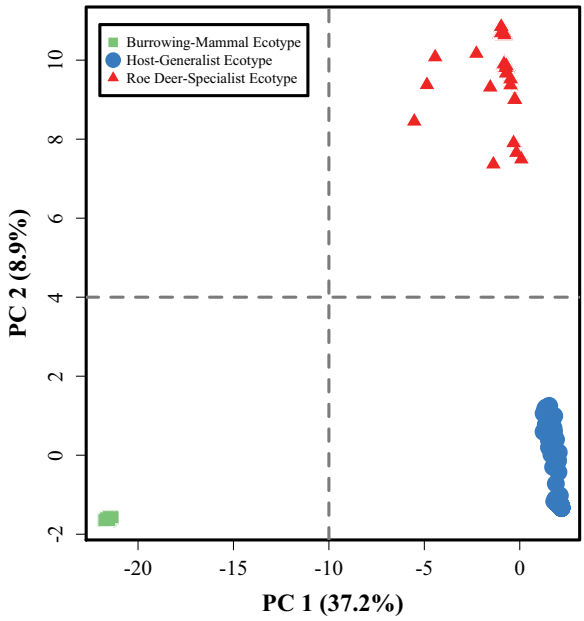

FIG 1 Results from a principal-component analysis (PCA) based on concatenated data from seven previously published genetic regions from European *A. phagocytophilum* samples (see the text). Shown are the first and second principal components (PC1 and PC2, respectively). This figure was produced using the R package Adegent version 2.1.3 (72), as implemented in R version 4.0.2 (73). Each color/shape combination represents one of the three mammal-infecting ecotypes of *A. phagocytophilum* circulating in Europe.

and the burrowing-mammal ecotype is more evolutionarily diverged. We therefore utilized the unique haplotypes observed among samples of the burrowing-mammal ecotype for outgroup comparisons, assuming that a nucleotide shared by this ecotype and one or both *I. ricinus*-vectored ecotypes is the ancestral state at this position. Accordingly, an alternative allele in one of the two focal ecotypes was assumed to be derived (44). Derived alleles most likely represent *de novo* mutations in their respective ecotype, although they may also have been introduced through recombination from unsampled populations (44). It is also possible that the same mutation arose independently in multiple ecotypes (e.g., genetic homoplasy), thus giving these variants the appearance of shared ancestry (45).

Using this conceptual framework, our pairwise comparisons of each ecotype's unique strains showed that the roe deer specialist ecotype harbored a greater number of observed derived changes across the seven genetic regions examined, considering both DNA variation that resulted in an amino acid change ("replacement" sites) (Fig. 2a) and variation that did not alter the protein sequence ("silent" sites) (Fig. 2b). For replacement sites, there was an average of 6.53 (standard deviation [SD] = 2.24) derived alleles per strain in the roe deer specialist ecotype, whereas in the host generalist ecotype, there was an average of 0.94 (SD = 0.91). These differences were statistically different from one another (t_{14079} = 303.8, $P < 0.001$). For silent substitutions, the roe deer specialist strains had an average of 35.54 (SD = 11.18) derived alleles, while the host generalist strains had an average of 32.22 (SD = 13.59) derived alleles. Again, these differences were statistically significant (t_{14079} = 54.6, $P < 0.001$).

Ecotype divergence times. We estimated approximate splitting times for the three ecotypes, with the goal of assessing how potential changes in European roe deer populations may have correlated with ecotype divergence in *A. phagocytophilum*. To make these estimates, we used a Bayesian coalescent method, implemented in BEAST and calibrated to absolute dates based on previously estimated divergence between *A. phagocytophilum* and the outgroup taxon *Anaplasma marginale* (46, 47). Our analysis recapitulated the predicted phylogenetic relationships between *A. marginale* and

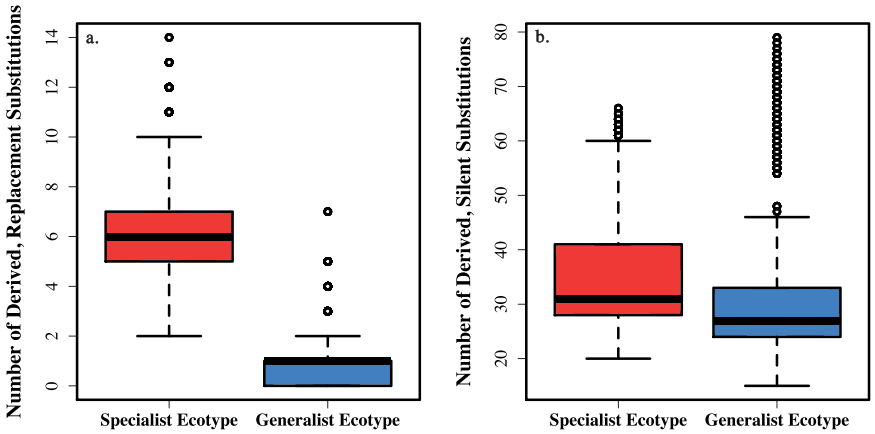

FIG 2 Box-and-whisker plots indicating the numbers of derived replacement (a) and silent (b) substitutions observed in all pairwise comparisons of each unique strain from the roe deer specialist and host generalist ecotypes. Polarization to define the ancestral state was done in comparison to the unique strains of the burrowing-mammal ecotype. The thick horizontal black bars indicate the observed median numbers of substitutions. The boxes show the second and third quartiles, while the whiskers indicate the first and fourth quartiles. Outlying observations are indicated by the open circles. Circles may represent more than one observation. Comparisons between the roe deer specialist ecotype and the generalist ecotype were both statistically significant ($P < 0.001$; see text for more details).

A. phagocytophilum, as well as those between the three ecotypes. As expected, *A. marginale* was the clear outgroup to all *A. phagocytophilum* strains (Fig. 3a). Among the *A. phagocytophilum* strains, those assigned to the burrowing-mammal ecotype clustered with one another and collectively were distinct from the *I. ricinus*-vectored ecotype strains (Fig. 3b). The strains assigned to the host generalist and roe deer specialist ecotypes, respectively, each formed distinct and monophyletic clusters and were sisters to one another. We estimated that the most recent common ancestor of all three ecotypes was present 3,330 years ago, with a range based on the 95% highest posterior density (HPD) of 529 to 8,061 years. Our estimate for the most recent common ancestor of the two *I. ricinus*-vectored ecotypes was 2,970 years ago, with a 95% HPD of 454 to 7,240 years.

Demographic changes in *I. ricinus*-vectored ecotypes and host. Given the relatively recent split observed between the two *I. ricinus*-vectored ecotypes (see above), and the apparent differences in their contemporary effective population sizes (N_e) (15), we wanted to determine if these two ecotypes have distinct demographic histories. We also wanted to examine potential correlations between demographic changes in the roe deer specialist ecotype and changes in the population density of its mammalian host. To estimate temporal changes in N_e for the two *I. ricinus*-vectored ecotypes, we applied an extended Bayesian skyline model, as implemented in BEAST, to separate subsets of the data containing only strains from each ecotype. Our results suggest that the roe deer specialist ecotype has had a relatively straightforward demographic history, undergoing a single, prolonged population expansion starting approximately 1,150 years ago (Fig. 4). In contrast, the host generalist ecotype has a more complex demographic history, with an apparent increase in N_e starting around 3,500 years ago followed by a more recent, rapid reduction starting approximately 500 years ago.

We also examined recent demographic changes in the roe deer population of Europe. To do this, we utilized previously published sequences of the mitochondrial cytochrome *b* (*cytb*) gene amplified from 46 roe deer from Poland (20). It was preferable to use samples from a single country, to reduce potential confounding effects of population structure on estimates of demographic history (48). Our results indicated a substantial increase in N_e for European roe deer starting approximately 1,500 years ago

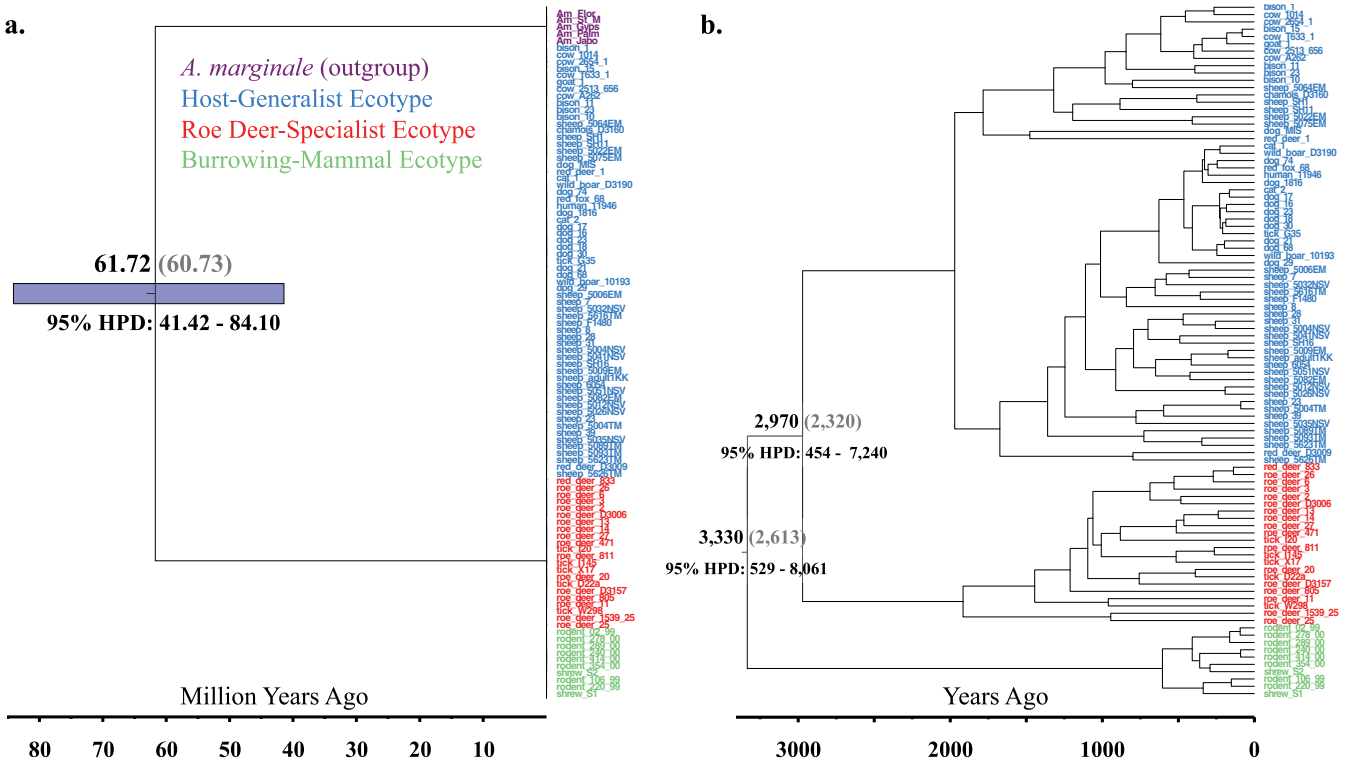

FIG 3 Divergence time estimates for the three ecotypes. (a) Tree, including the outgroup *A. marginale*. (b) Subtree with only samples of *A. phagocytophilum*. In both trees, the mean estimated divergence time is shown above the node (in black text), followed in parentheses by the median estimated divergence time (in gray text). The 95% HPD is given below the node. Only estimates for nodes representing species/ecotype divergences are shown. The scale bar at the bottom of each tree indicates time in millions of years (a) or years (b). Duplicate haplotypes were removed prior to analysis.

(Fig. 5), with a contemporary N_e estimated to be over 20 times as large as that at the start of this apparent demographic expansion.

DISCUSSION

A comparison of derived alleles in the host generalist and roe deer specialist ecotypes supports our hypothesis that the specialist evolved from a population of *A. phagocytophilum* that resembled the contemporary host generalist, with a broad host range (Fig. 2). In turn, this observation suggests that specific changes in the roe deer population of Europe facilitated the emergence and evolution of the roe deer specialist ecotype.

One such change was likely the geographic expansion of roe deer in Europe after the last glacial maximum. This recolonization event correlates with a dramatic increase in the N_e of mainland European roe deer that is estimated to have occurred between 3,932 and 7,919 years ago (19). It is also possible that an increase in roe deer population density (independent of range expansion) may have contributed to the divergence of the two *I. ricinus*-vectored ecotypes of European *A. phagocytophilum*. Such a density increase likely occurred in conjunction with more intensive agricultural practices in Europe starting around 1,500 years ago (25, 26). Our estimation of the split between the two *I. ricinus*-vectored ecotypes suggests they last shared a common ancestor ~2,970 years ago, with a range of 454 to 7,240 years (based on the 95% HPD) (Fig. 3). This estimate, combined with the earlier conclusion that the roe deer specialist is the more derived ecotype, indicates that host specialization followed the geographic range expansion of roe deer throughout Europe after the last glacial maximum. Two *I. ricinus*-vectored populations of *A. phagocytophilum* resembling the contemporary ecotypes had already evolved prior to the increases in roe deer population densities that occurred with more recent shifts in the vegetation communities of Europe.

Our analysis of patterns of demographic change in the roe deer specialist ecotype in relation to changes in the European roe deer population revealed a strong

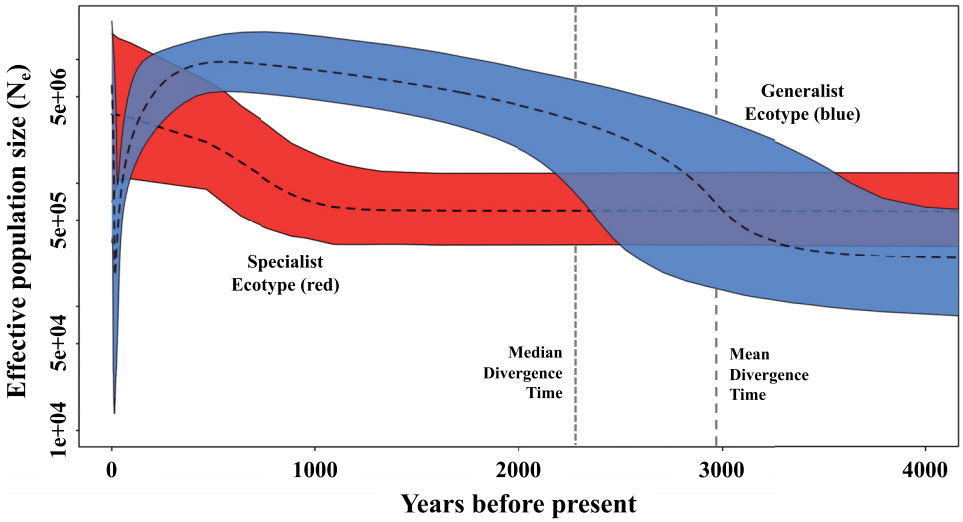

FIG 4 Extended Bayesian skyline plot (EBSP) results for the two *I. ricinus*-vectored ecotypes showing changes in effective population size (N_e) over time. Median results are shown by the black dashed lines, while the upper and lower 95% highest posterior density (HPD) intervals are indicated by the colored areas contained within the thin solid lines. Changes in the roe deer specialist ecotype's N_e are indicated in red, and changes in the host generalist ecotype's N_e are indicated in blue. The original timeline was determined in terms of number of generations, and this was converted to years assuming 100 generations per year (see Materials and Methods for more details). The timeline has been restricted to 4,000 years. Also shown are the median (finely dashed vertical gray line) and mean (coarsely dashed vertical gray line) divergence estimates for the two ecotypes from our Bayesian coalescent analysis (see the text and Fig. 3 for more details).

correlation between the two. Specifically, the N_e of both increased dramatically around 1,500 years ago (Fig. 4 and 5). These observations support our third hypothesis that changes in the population density of roe deer correlate with an increase in the N_e of the roe deer specialist ecotype. However, it is possible that these two seemingly congruent changes in N_e are coincidental and unrelated. It is also possible that additional factors not considered here were the primary drivers of the effective population size increases in both roe deer and the roe deer specialist ecotype of *A. phagocytophilum* and that the former is not directly the cause of the latter. Nonetheless, if future work does support a causal relationship between host and pathogen in this system, then the apparently strong genetic tracking of its host population by this ecotype will strengthen confidence that it is truly a host specialist, highly adapted to infecting roe deer. It will also provide evidence that host specialization can correlate with higher fitness in a pathogenic bacterial population (38).

Among possible European mammalian hosts, it is perhaps not surprising that roe deer were the species that came to support a specialist ecotype of *A. phagocytophilum*. Unlike many other mammal species across the European continent, roe deer appear to have benefitted from growth in the human population and the corresponding increased intensity of agriculture (27–29). Increases in roe deer population density meant that *A. phagocytophilum* transmission events became increasingly likely to occur from one roe deer to another. As strains of *A. phagocytophilum* progressively found themselves only in roe deer, this could have produced strong selection for improved abilities to infect this host, possibly at the expense of infection capabilities in other mammals.

The host generalist ecotype looks to have experienced a more complicated demographic history (Fig. 4). Intriguingly, it appears to have been in the midst of a substantial population increase when it last shared a common ancestor with the roe deer specialist ecotype. Range expansions of most mammal species at this time (49) may have resulted in increased transmission rates between hosts. Increased transmission and greater bacterial population sizes have been shown to correlate with genetic diversity and ecotype

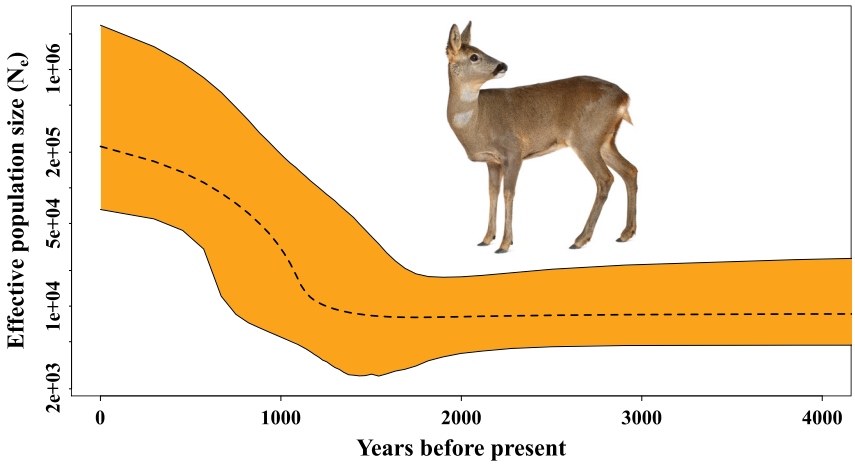

FIG 5 Extended Bayesian skyline plot (EBSP) results for roe deer (*Capreolus capreolus*), showing changes in effective population size (N_e) over time. Median results are shown by the black dashed line, while the upper and lower 95% highest posterior density (HPD) intervals are indicated by the orange areas contained within the thin solid lines. The timeline has been restricted to 4,000 years.

divergence/specialization in pathogenic bacteria (50–52). This scenario would explain both the patterns of ecotype divergence in European *A. phagocytophilum* and the major increase in the N_e of the host generalist ecotype during this period.

More contemporarily, the host generalist ecotype appears to have undergone a major reduction in N_e, possibly followed by some very recent recovery (Fig. 4). A reduction in N_e could occur if its rate of transmission decreased (53). It could also occur in a bacterial population if one or more advantageous alleles became strongly favored and this generated a selective sweep (54). For a bacterial ecotype with a wide geographic distribution and a broad host range, a selective sweep appears to be the less probable of these two scenarios. Therefore, if we postulate a decrease in the population size of the host generalist ecotype, this again points to the importance of host population dynamics in the demography of European *A. phagocytophilum*. While the host generalist ecotype can infect humans as well as a wide variety of livestock and companion animal species, its predominant natural reservoirs likely include European bison, wild boars, hedgehogs, and possibly red deer (9, 10, 55). Unlike roe deer, many of these species have experienced substantial population declines over the last several centuries in Europe (30, 56–59).

As a tick-transmitted bacterium, the emergence and maintenance of distinct *A. phagocytophilum* ecotypes could also be the result of dynamics in its arthropod vectors (60). This appears to be a contributing factor in the evolution of the burrowing-mammal ecotype, as its primary vector, *I. trianguliceps*, is found to live almost exclusively in nests and burrows, thus limiting the potential for direct transmission to larger mammals (12, 16, 61). However, in Europe, the other two mammal-infecting ecotypes of *A. phagocytophilum* are both primarily vectored by *I. ricinus*, which has one of the broadest blood-host ranges of any tick species (62). While some local population structure and host specialization may occur in *I. ricinus* (63), this is unlikely to contribute to the continued maintenance of the discrete enzootic cycle of the roe deer specialist ecotype, as it occurs throughout Europe (9–11, 64). Within *I. ricinus*, a change in its distribution or population size fluctuations could also be a major factor influencing patterns of divergence in *A. phagocytophilum* (60). Substantial changes in the vector population size would likely impact both bacterial transmission rates and host usage (65), and it is possible such changes contributed to ecotype formation and divergence in European *A. phagocytophilum*.

One important caveat of this study is the differences between effective population sizes (N_e) and census population sizes. Here, we have considered changes in N_e modeled by other researchers in addition to those determined from our own modeling

efforts. Effective population sizes are almost always smaller than census population sizes and can decrease for a variety of reasons, including genetic drift, selection, or migration (53, 66). However, with the exception of a decreased N_e modeled in the host generalist ecotype, the observations of change in N_e considered here all suggest increases over time (Fig. 4 and 5). Prolonged increases in N_e are almost always the result of population growth and, therefore, should accurately reflect increases in population sizes or rates of transmission (22, 53, 67, 68). Nonetheless, it is also possible that the observed increases in *A. phagocytophilum* reflect greater rates of gene flow between populations (68–70).

A central question that remains concerning *A. phagocytophilum* ecotype evolution is the near absence of the host generalist ecotype in roe deer. Although strains found in roe deer have been attributed to the host generalist ecotype, such infections are far less common in this host than those attributed to the specialist ecotype (10). Given the phylogenetic diversity of species the host generalist ecotype can infect (including other Artiodactyla), it seems likely that the host generalist has the capacity to establish infections in roe deer. This suggests that competition with the roe deer specialist ecotype may contribute to the generalist's absence from this host (5, 6). Additional work will be required to further elucidate this possibility.

In conclusion, we have shown that among *I. ricinus*-vectored populations of *A. phagocytophilum* in Europe, the roe deer specialist ecotype is the more derived. We have also shown that its divergence from a likely ancestral host generalist occurred after the last glacial maximum as mammal populations (including roe deer) recolonized the European mainland. Finally, we have provided evidence that the ecotype of *A. phagocytophilum* that infects roe deer in Europe is truly a specialist of this host and that changes in the population of roe deer strongly correlate with the population dynamics of this bacterium. Future comparisons of closely related *A. phagocytophilum* ecotypes may help us better understand the ways in which host specialization influences ecology, demography, and genomic evolution in pathogenic bacteria.

MATERIALS AND METHODS

Data set and sample clustering. Throughout this study, we utilized partial sequences from seven *A. phagocytophilum* housekeeping genes (*atpA*, *dnaN*, *fumC*, *glyA*, *mdh*, *pheS*, and *sucA*), representing 2,877 nucleotides in total (GenBank accession numbers KF242733 to KF245413). These data were sequenced from *A. phagocytophilum* DNA isolated from 17 different mammalian hosts and *Ixodes ricinus* ticks, collected from multiple European countries (9). These *A. phagocytophilum* samples were previously found to represent three discrete genetic groups, which are hypothesized to have independent enzootic cycles (9).

Prior to any analyses, we removed from the published data set all samples that had an ambiguous nucleotide in any of the seven genetic regions, as well as the small number of samples from the United States. This left us with 278 samples in total. With these we performed a nonparametric clustering analysis using principal components to confirm sample placement into each of three discrete groups. To do this, we first concatenated the seven gene regions into a single FASTA sequence. We then used the R package Adegenet version 2.1.3 (71, 72), as implemented in R version 4.0.2 (73), to determine the principal components from the data. The relationship between the first and second principal components was visualized in R.

Derived ecotype divergence. To examine patterns of ecotype divergence with our data set, we first removed any duplicated haplotype observed within each of the three ecotypes. This left us with 10 unique strains for the burrowing-mammal ecotype, 22 unique strains for the roe deer specialist ecotype, and 68 unique strains for the host generalist ecotype. We then performed a three-way comparison, using the burrowing-mammal ecotype strains as the outgroup, to polarize observed differences. A derived site was counted if the nucleotide of the outgroup sequence at a specific position matched the nucleotide at this same position in one of the focal ecotype strains but the second focal strain had a different nucleotide. In this case, the allele in the second strain would be counted as a derived allele. All derived changes were classified as replacement (i.e., they resulted in an amino acid change) or silent (i.e., the amino acid sequence encoded by the specific codon was not different). We used a custom Perl script to make all possible 3-way comparisons among the three ecotypes (one strain per ecotype). R version 4.0.2 (73) was used to calculate summary statistics for the resulting output. Additionally, statistical differences in the numbers of derived sites observed between the two focal ecotypes were determined with a pairwise *t* test, also implemented in R version 4.0.2 (73). Results were considered significant at a *P* value of <0.05.

Estimation of ecotype divergence times. The data for *A. phagocytophilum* consisted of the same unique strains as those used in our assessment of derived ecotype divergence (see above). To this data set, we added orthologous gene sequences from five publicly available whole genomes of the related bacterium *A. marginale* (NCBI GenBank accession numbers GCF_000020305.1, GCF_000011945.1, GCF_000495495.1, GCF_003515675.1, and GCF_003515735.1). To identify these sequences, we first located the full gene sequences for each of the seven focal genes from the annotated genome of

A. marginale strain Florida (accession number GCA_000020305.1). These full DNA sequences were converted to amino acid sequences and then aligned to the similarly converted partial *A. phagocytophilum* sequences using the Muscle algorithm (74), as implemented in SeaView version 4.6.3 (75, 76). The orthologous-sequence regions of *A. marginale* were then used as the query sequences to locate the same regions in the other four *A. marginale* genomes with a local BLAST (blastp, e value = 1e−50) (77). With the resulting genome coordinates, we used Samtools version 1.9 (78) to extract these regions and combine them with our *A. phagocytophilum* sequences. These data were then aligned based on the amino acid sequences for the seven genetic regions as described above. Sites containing missing data or gaps indicating an insertion/deletion (indels) in any of the samples were removed from all samples, as were any sites that were ambiguously aligned around an indel, as determined by manual inspection.

Next, we determined the best partitioning scheme and mutation model across this concatenated data set using PartitionFinder 2.1.1 (79) with the small-sample-size-corrected version of the Akaike information criterion (AICc). The first, second, and third codon positions for the concatenated sequences were considered separate partitions.

To estimate approximate ecotype splitting times, we used the Bayesian Evolutionary Analysis Utility tool (BEAUti version 2.3.2) to create the XML file for implementation in BEAST version 2.5.2 (80). We split our data set to apply distinct site models to the first, second, and third positions. Based on the results of our PartitionFinder analysis (see above), we used the generalized time-reversible (GTR) (81) model of mutation for each data subset, with six gamma categories and an estimated shape. A proportion of invariant sites was also applied to the first and second data subsets.

We ran two Markov chain Monte Carlo chains of 10^8 iterations in BEAST, with a log normal relaxed clock and the calibrated Yule model for our tree prior. To place our divergence estimates in absolute time, we used a prior of 59 Ma (million years ago; mean in real space, not log space) with a log normal distribution and an *S* parameter of 0.18 for the splitting time of *Anaplasma phagocytophilum* and *A. marginale*, which generated a 95% probability range that encompassed the range previously reported (46). We used LogCombiner version 1.7.5 to combine the two separate runs and Tracer version 1.5 to ensure that the effective sample size (ESS) of the parameters exceeded 200. The program TreeAnnotator version 2.5.2 was used to summarize the tree log, and FigTree version 1.4.4 (82) was used for tree visualization.

Assessment of ecotype-specific demographic changes. We looked at demographic changes in the two *I. ricinus*-vectored ecotypes using an extended Bayesian skyline plot (EBSP) (34), as implemented in BEAST version 2.5.2 (80). To do this, we used the same partitioning scheme and corresponding evolutionary models as previously determined (see above). The samples that corresponded to the host generalist and roe deer specialist ecotypes were examined separately. We used all samples in this analysis. To roughly estimate changes in effective population size, we multiplied the previously estimated substitution rate of 8.9×10^{-11} per synonymous base pair per generation determined in *Escherichia coli* (83) by the 959 bp included in this partition. This gave us a clock rate of 8.5×10^{-8} per generation, which was applied to the partition corresponding to the third codon positions of each sequence. The rates for the other two partitions were estimated. A strict molecular clock rate was applied to each of the three partitions. Although the true rates of sequence evolution in *A. phagocytophilum* are unknown, and these rates are likely different for the two ecotypes on ecological scales, across longer periods, the bacterial evolution rates appear relatively comparable (84). Furthermore, the qualitative patterns of demographic change observed here are not affected by variation in these estimates; only the timing of such events shifts depending on the mutation rate applied (results not shown).

We ran two Markov chain Monte Carlo chains of 10^8 iterations in the program BEAST, with a burn-in of 10% and sampling at every 1,000 chains. The separate runs were combined using LogCombiner version 2.5.2, and Tracer version 1.7.1 was used to confirm that the effective sample size (ESS) of the relevant parameters exceeded 200. To convert estimates of change in N_e from generations to absolute time, we first considered a previously calculated doubling time for wild *Escherichia coli* populations of ~15 h (85). This value was calculated using mutation rate estimates from laboratory observations and estimates of the number of accumulated mutations per year for *E. coli* in the wild. A doubling time of every 15 h translates to approximately 584 *E. coli* generations per year. The number of generations per year for obligately intracellular bacteria has been estimated to be nearly 6 times less than that of *E. coli* (86). This 6-fold reduction in generation time is also supported by research in *Ehrlichia canis*, an obligately intracellular, tick-transmitted pathogen closely related to *A. phagocytophilum* (87). For mathematical convenience, we rounded the applied generation time calculated from these values (584/6 = 97.3) to 100 generations per year for European *A. phagocytophilum*.

Patterns of demographic change in roe deer. To examine potential demographic changes in the European roe deer population, we utilized published sequences of the mitochondrial cytochrome *b* (*cytb*) gene amplified from 46 roe deer from Poland (20). We first partitioned this data set into the first, second, and third codon sites and used Partionfinder version 2.1.1 (79) to determine the most appropriate model of evolution to apply to each partition. We then used an extended Bayesian skyline plot (EBSP) (34), as implemented in BEAST version 2.5.2 (80), to examine changes in effective population size over time in this species. The data set was split into first, second, and third positions. Following our PartionFinder results (see above), the Hasegawa-Kishino-Yano (HKY) model of substitution (88) was applied to the first and second positions, and the first position also had a portion of invariable sites (0.1) applied to it. We applied the generalized time-reversible (GTR) (81) model of mutation to the third position, with six gamma categories and an estimated shape. To roughly estimate changes in effective population size, we applied a clock rate of 2.0×10^{-6} to the partition corresponding to the third codon positions of each sequence (representing 2% per site per million years) (89). We note here that the use of this rough rate for mammals has been shown to be problematic (90, 91) but that 2% is similar to comparable rates reported for Artiodactyla (92). The rates for the other two partitions were estimated. We applied a strict molecular clock rate to each of the three partitions.

We ran two Markov chain Monte Carlo chains of 10^8 iterations in the program BEAST, with a burn-in of 10% and sampling at every 1,000 chains. The two separate runs were combined using LogCombiner version 1.7.5, and the results assessed with Tracer version 1.5 to ensure that the effective sample size (ESS) of the parameters exceeded 200. We visualized the demographic changes by plotting the combined EBSP log files in R version 4.0.2 (73), using the plotEBSP.R script, which is available in the BEAST package.

ACKNOWLEDGMENTS

Nina V. Bates was supported in part by a fiscal year 2019 student faculty scholarship grant from Montclair State University. Qiana E. Archer was funded by an award from the Garden State-Louis Stokes Alliances for Minority Participation (LSAMP) program (NSF award number 1909824).

REFERENCES

1. Dobson A. 2004. Population dynamics of pathogens with multiple host species. Am Nat 164:S64–S78. https://doi.org/10.1086/424681.
2. Woolhouse ME, Gowtage-Sequeria S. 2005. Host range and emerging and reemerging pathogens. Emerg Infect Dis 11:1842–1847. https://doi.org/10.3201/eid1112.050997.
3. Johnson CK, Hitchens PL, Pandit PS, Rushmore J, Evans TS, Young C, Doyle MM. 2020. Global shifts in mammalian population trends reveal key predictors of virus spillover risk. Proc Biol Sci 287:20192736. https://doi.org/10.1098/rspb.2019.2736.
4. Thines M. 2019. An evolutionary framework for host shifts—jumping ships for survival. New Phytol 224:605–617. https://doi.org/10.1111/nph.16092.
5. Hibbing ME, Fuqua C, Parsek MR, Peterson SB. 2010. Bacterial competition: surviving and thriving in the microbial jungle. Nat Rev Microbiol 8:15–25. https://doi.org/10.1038/nrmicro2259.
6. Okamoto KW, Amarasekare P, Post DM, Vasseur DA, Turner PE. 2019. The interplay between host community structure and pathogen life-history constraints in driving the evolution of host-range shifts. Funct Ecol 33:2338–2353. https://doi.org/10.1111/1365-2435.13467.
7. Shaw LP, Wang AD, Dylus D, Meier M, Pogacnik G, Dessimoz C, Balloux F. 2020. The phylogenetic range of bacterial and viral pathogens of vertebrates. Mol Ecol 29:3361–3379. https://doi.org/10.1111/mec.15463.
8. Stuen S, Granquist EG, Silaghi C. 2013. *Anaplasma phagocytophilum*—a widespread multi-host pathogen with highly adaptive strategies. Front Cell Infect Microbiol 3:31. https://doi.org/10.3389/fcimb.2013.00031.
9. Huhn C, Winter C, Wolfsperger T, Wüppenhorst N, Strašek Smrdel K, Skuballa J, Pfäffle M, Petney T, Silaghi C, Dyachenko V, Pantchev N, Straubinger RK, Schaarschmidt-Kiener D, Ganter M, Aardema ML, von Loewenich FD. 2014. Analysis of the population structure of *Anaplasma phagocytophilum* using multilocus sequence typing. PLoS One 9:e93725. https://doi.org/10.1371/journal.pone.0093725.
10. Jahfari S, Coipan EC, Fonville M, van Leeuwen AD, Hengeveld P, Heylen D, Heyman P, van Maanen C, Butler CM, Földvári G, Szekeres S, van Duijvendijk G, Tack W, Rijks JM, van der Giessen J, Takke W, van Wieren SE, Takumi K, Sprong H. 2014. Circulation of four *Anaplasma phagocytophilum* ecotypes in Europe. Parasit Vectors 7:365. https://doi.org/10.1186/1756-3305-7-365.
11. Jaarsma RI, Sprong H, Takumi K, Kazimirova M, Silaghi C, Mysterud A, Rudolf I, Beck R, Földvári G, Tomassone L, Groenevelt M, Everts RR, Rijks JM, Ecke F, Hörnfeldt B, Modrý D, Majerová K, Votýpka J, Estrada-Peña A. 2019. *Anaplasma phagocytophilum* evolves in geographical and biotic niches of vertebrates and ticks. Parasit Vectors 12:328. https://doi.org/10.1186/s13071-019-3583-8.
12. Bown KJ, Lambin X, Ogden NH, Begon M, Telford G, Woldehiwet Z, Birtles RJ. 2009. Delineating *Anaplasma phagocytophilum* ecotypes in coexisting, discrete enzootic cycles. Emerg Infect Dis 15:1948–1954. https://doi.org/10.3201/eid1512.090178.
13. Langenwalder DB, Schmidt S, Gilli U, Pantchev N, Ganter M, Silaghi C, Aardema ML, von Loewenich FD. 2019. Genetic characterization of *Anaplasma phagocytophilum* strains from goats (*Capra aegagrus hircus*) and water buffalo (*Bubalus bubalis*) by 16S rRNA gene, *ankA* gene and multilocus sequence typing. Ticks Tick Borne Dis 10:101267. https://doi.org/10.1016/j.ttbdis.2019.101267.
14. Langenwalder DB, Schmidt S, Silaghi C, Skuballa J, Pantchev N, Matei IA, Mihalca AD, Gilli U, Zajkowska J, Ganter M, Hoffman T, Salaneck E, Petrovec M, von Loewenich FD. 2020. The absence of the *drhm* gene is not a marker for human-pathogenicity in European *Anaplasma phagocytophilum* strains. Parasit Vectors 13:238. https://doi.org/10.1186/s13071-020-04116-z.
15. Aardema ML, von Loewenich FD. 2015. Varying influences of selection and demography in host adapted populations of the tick-transmitted bacterium, *Anaplasma phagocytophilum*. BMC Evol Biol 15:58. https://doi.org/10.1186/s12862-015-0335-z.
16. Blaňarová L, Stanko M, Carpi G, Miklisová D, Víchová B, Mošanský L, Bona M, Derdáková M. 2014. Distinct *Anaplasma phagocytophilum* genotypes associated with *Ixodes trianguliceps* ticks and rodents in Central Europe. Ticks Tick Borne Dis 5:928–938. https://doi.org/10.1016/j.ttbdis.2014.07.012.
17. Ismail N, Bloch KC, McBride JW. 2010. Human ehrlichiosis and anaplasmosis. Clin Lab Med 30:261–292. https://doi.org/10.1016/j.cll.2009.10.004.
18. Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW, McCabe AM. 2009. The last glacial maximum. Science 325:710–714. https://doi.org/10.1126/science.1172873.
19. Baker KH, Hoelzel AR. 2014. Influence of Holocene environmental change and anthropogenic impact on the diversity and distribution of roe deer. Heredity (Edinb) 112:607–615. https://doi.org/10.1038/hdy.2013.142.
20. Matosiuk M, Borkowska A, Świsłocka M, Mirski P, Borowski Z, Krysiuk K, Danilkin AA, Zvychaynaya EY, Saveljev AP, Ratkiewicz M. 2014. Unexpected population genetic structure of European roe deer in Poland: an invasion of the mt DNA genome from Siberian roe deer. Mol Ecol 23:2559–2572. https://doi.org/10.1111/mec.12745.
21. Pierson JC, Graves TA, Banks SC, Kendall KC, Lindenmayer DB. 2018. Relationship between effective and demographic population size in continuously distributed populations. Evol Appl 11:1162–1175. https://doi.org/10.1111/eva.12636.
22. Ho SY, Shapiro B. 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences. Mol Ecol Resour 11:423–434. https://doi.org/10.1111/j.1755-0998.2011.02988.x.
23. Miller EF, Green RE, Balmford A, Delser PM, Beyer R, Somveille M, Leonardi M, Amos W, Manica A. 2021. Bayesian skyline plots disagree with range size changes based on species distribution models for holarctic birds. Mol Ecol 30:3993–4004. https://doi.org/10.1111/mec.16032.
24. Kaplan JO, Krumhardt KM, Zimmermann N. 2009. The prehistoric and preindustrial deforestation of Europe. Quat Sci Rev 28:3016–3034. https://doi.org/10.1016/j.quascirev.2009.09.028.
25. Fyfe RM, Woodbridge J, Roberts N. 2015. From forest to farmland: pollen-inferred land cover change across Europe using the pseudobiomization approach. Glob Chang Biol 21:1197–1212. https://doi.org/10.1111/gcb.12776.
26. Zanon M, Davis BAS, Marquer L, Brewer S, Kaplan JO. 2018. European forest cover during the past 12,000 years: a palynological reconstruction based on modern analogs and remote sensing. Front Plant Sci 9:253. https://doi.org/10.3389/fpls.2018.00253.
27. Saïd S, Servanty S. 2005. The influence of landscape structure on female roe deer home-range size. Landscape Ecol 20:1003–1012. https://doi.org/10.1007/s10980-005-7518-8.
28. Heurich M, Brand TT, Kaandorp MY, Šustr P, Müller J, Reineking B. 2015. Country, cover or protection: what shapes the distribution of red deer and roe deer in the Bohemian Forest Ecosystem? PLoS One 10:e0120960. https://doi.org/10.1371/journal.pone.0120960.
29. Lovari S, Serrao G, Mori E. 2017. Woodland features determining home range size of roe deer. Behav Processes 140:115–120. https://doi.org/10.1016/j.beproc.2017.04.012.
30. Crees JJ, Carbone C, Sommer RS, Benecke N, Turvey ST. 2016. Millennial-scale faunal record reveals differential resilience of European large

mammals to human impacts across the Holocene. Proc Biol Sci 283: 20152152. https://doi.org/10.1098/rspb.2015.2152.
31. Cooper VS, Lenski RE. 2000. The population genetics of ecological specialization in evolving *Escherichia coli* populations. Nature 407:736–739. https://doi.org/10.1038/35037572.
32. Kingman JFC. 1982. On the genealogy of large populations. J Appl Probab 19:27–43. https://doi.org/10.1017/S0021900200034446.
33. Kuhner MK. 2009. Coalescent genealogy samplers: windows into population history. Trends Ecol Evol 24:86–93. https://doi.org/10.1016/j.tree.2008.09.007.
34. Heled J, Drummond AJ. 2012. Calibrated tree priors for relaxed phylogenetics and divergence time estimation. Syst Biol 61:138–149. https://doi.org/10.1093/sysbio/syr087.
35. Bacigalupe R, Tormo-Mas MÁ, Penadés JR, Fitzgerald JR. 2019. A multihost bacterial pathogen overcomes continuous population bottlenecks to adapt to new host species. Sci Adv 5:eaax0063. https://doi.org/10.1126/sciadv.aax0063.
36. Johnson KP, Malenke JR, Clayton DH. 2009. Competition promotes the evolution of host generalists in obligate parasites. Proc Biol Sci 276: 3921–3926. https://doi.org/10.1098/rspb.2009.1174.
37. Toft C, Andersson SG. 2010. Evolutionary microbial genomics: insights into bacterial host adaptation. Nat Rev Genet 11:465–475. https://doi.org/10.1038/nrg2798.
38. Fry JD. 1996. The evolution of host specialization: are trade-offs overrated? Am Nat 148:S84–S107. https://doi.org/10.1086/285904.
39. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, Hansen NF, Durand EY, Malaspinas AS, Jensen JD, Marques-Bonet T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Höber B, Höffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Ž, Gušic I, Doronichev VB, Golovanova LV, Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PLF, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M, et al. 2010. A draft sequence of the Neandertal genome. Science 328: 710–722. https://doi.org/10.1126/science.1188021.
40. Lachance J, Vernot B, Elbers CC, Ferwerda B, Froment A, Bodo JM, Lema G, Fu W, Nyambo TB, Rebbeck TR, Zhang K, Akey JM, Tishkoff SA. 2012. Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse African hunter-gatherers. Cell 150:457–469. https://doi.org/10.1016/j.cell.2012.07.009.
41. Karlsson EK, Kwiatkowski DP, Sabeti PC. 2014. Natural selection and infectious disease in human populations. Nat Rev Genet 15:379–393. https://doi.org/10.1038/nrg3734.
42. Didelot X, Walker AS, Peto TE, Crook DW, Wilson DJ. 2016. Within-host evolution of bacterial pathogens. Nat Rev Microbiol 14:150–162. https://doi.org/10.1038/nrmicro.2015.13.
43. Vera-Ponce de Leon A, Schneider MG, Jahnes BC, Sadowski V, Camuy-Vélez LA, Duan J, Sabree ZL. 2022. Genetic drift and host-adaptive features likely underlie cladogenesis of insect-associated Lachnospiraceae. Genome Biol Evol 14:evac086. https://doi.org/10.1093/gbe/evac086.
44. Arnold B, Sohail M, Wadsworth C, Corander J, Hanage WP, Sunyaev S, Grad YH. 2020. Fine-scale haplotype structure reveals strong signatures of positive selection in a recombining bacterial pathogen. Mol Biol Evol 37:417–428. https://doi.org/10.1093/molbev/msz225.
45. Sakoparnig T, Field C, van Nimwegen E. 2021. Whole genome phylogenies reflect the distributions of recombination rates for many bacterial species. Elife 10:e65366. https://doi.org/10.7554/eLife.65366.
46. Foley J, Nieto NC, Foley P, Teglas MB. 2008. Co-phylogenetic analysis of *Anaplasma phagocytophilum* and its vectors, *Ixodes* spp. ticks. Exp Appl Acarol 45:155–170. https://doi.org/10.1007/s10493-008-9173-7.
47. de la Fuente J, Estrada-Peña A, Cabezas-Cruz A, Brey R. 2015. Flying ticks: anciently evolved associations that constitute a risk of infectious disease spread. Parasit Vectors 8:538. https://doi.org/10.1186/s13071-015-1154-1.
48. Heller R, Chikhi L, Siegismund HR. 2013. The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. PLoS One 8:e62992. https://doi.org/10.1371/journal.pone.0062992.
49. Hewitt GM. 1999. Post-glacial re-colonization of European biota. Biol J Linn Soc 68:87–112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x.
50. Ueti MW, Tan Y, Broschat SL, Castañeda Ortiz EJ, Camacho-Nuez M, Mosqueda JJ, Scoles GA, Grimes M, Brayton KA, Palmer GH. 2012. Expansion of variant diversity associated with a high prevalence of pathogen strain superinfection under conditions of natural transmission. Infect Immun 80:2354–2360. https://doi.org/10.1128/IAI.00341-12.
51. Pollitt EJ, West SA, Crusz SA, Burton-Chellew MN, Diggle SP. 2014. Cooperation, quorum sensing, and evolution of virulence in *Staphylococcus aureus*. Infect Immun 82:1045–1051. https://doi.org/10.1128/IAI.01216-13.
52. Chavhan Y, Malusare S, Dey S. 2020. Larger bacterial populations evolve heavier fitness trade-offs and undergo greater ecological specialization. Heredity (Edinb) 124:726–736. https://doi.org/10.1038/s41437-020-0308-x.
53. Bobay LM, Ochman H. 2018. Factors driving effective population size and pan-genome evolution in bacteria. BMC Evol Biol 18:153. https://doi.org/10.1186/s12862-018-1272-4.
54. Bendall ML, Stevens SL, Chan LK, Malfatti S, Schwientek P, Tremblay J, Schackwitz W, Martin J, Pati A, Bushnell B, Froula J, Kang D, Tringe SG, Bertilsson S, Moran MA, Shade A, Newton RJ, McMahon KD, Malmstrom RR. 2016. Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. ISME J 10:1589–1601. https://doi.org/10.1038/ismej.2015.241.
55. Dugat T, Lagrée AC, Maillard R, Boulouis HJ, Haddad N. 2015. Opening the black box of *Anaplasma phagocytophilum* diversity: current situation and future perspectives. Front Cell Infect Microbiol 5:61. https://doi.org/10.3389/fcimb.2015.00061.
56. Scandura M, Iacolina L, Crestanello B, Pecchioli E, Di Benedetto MF, Russo V, Davoli R, Apollonio M, Bertorelle G. 2008. Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable? Mol Ecol 17:1745–1762. https://doi.org/10.1111/j.1365-294X.2008.03703.x.
57. Kuemmerle T, Hickler T, Olofsson J, Schurgers G, Radeloff VC. 2012. Reconstructing range dynamics and range fragmentation of European bison for the last 8000 years. Divers Distrib 18:47–59. https://doi.org/10.1111/j.1472-4642.2011.00849.x.
58. Rosvold J, Røed KH, Hufthammer AK, Andersen R, Stenøien HK. 2012. Reconstructing the history of a fragmented and heavily exploited red deer population using ancient and contemporary DNA. BMC Evol Biol 12: 191. https://doi.org/10.1186/1471-2148-12-191.
59. Rasmussen SL, Nielsen JL, Jones OR, Berg TB, Pertoldi C. 2020. Genetic structure of the European hedgehog (*Erinaceus europaeus*) in Denmark. PLoS One 15:e0227205. https://doi.org/10.1371/journal.pone.0227205.
60. Ogden NH, Mechai S, Margos G. 2013. Changing geographic ranges of ticks and tick-borne pathogens: drivers, mechanisms and consequences for pathogen diversity. Front Cell Infect Microbiol 3:46. https://doi.org/10.3389/fcimb.2013.00046.
61. Mysterud A, Byrkjeland R, Qviller L, Viljugrein H. 2015. The generalist tick *Ixodes ricinus* and the specialist tick *Ixodes trianguliceps* on shrews and rodents in a northern forest ecosystem—a role of body size even among small hosts. Parasit Vectors 8:639. https://doi.org/10.1186/s13071-015-1258-7.
62. Estrada-Peña A, de la Fuente J. 2017. Host distribution does not limit the range of the tick *Ixodes ricinus* but impacts the circulation of transmitted pathogens. Front Cell Infect Microbiol 7:405. https://doi.org/10.3389/fcimb.2017.00405.
63. McCoy KD, Léger E, Dietrich M. 2013. Host specialization in ticks and transmission of tick-borne diseases: a review. Front Cell Infect Microbiol 3: 57. https://doi.org/10.3389/fcimb.2013.00057.
64. Chastagner A, Dugat T, Vourc'h G, Verheyden H, Legrand L, Bachy V, Chabanne L, Joncour G, Maillard R, Boulouis H-J, Haddad N, Bailly X, Leblond A. 2014. Multilocus sequence analysis of *Anaplasma phagocytophilum* reveals three distinct lineages with different host ranges in clinically ill French cattle. Vet Res 45:114. https://doi.org/10.1186/s13567-014-0114-7.
65. Barrett LG, Thrall PH, Burdon JJ, Linde CC. 2008. Life history determines genetic structure and evolutionary potential of host-parasite interactions. Trends Ecol Evol 23:678–685. https://doi.org/10.1016/j.tree.2008.06.017.
66. McInerney JO, McNally A, O'Connell MJ. 2017. Why prokaryotes have pangenomes. Nat Microbiol 2:17040. https://doi.org/10.1038/nmicrobiol.2017.40.
67. Gillespie JH. 2004. Population genetics: a concise guide. Johns Hopkins University Press, Baltimore, MD.
68. Hague MT, Routman EJ. 2016. Does population size affect genetic diversity? A test with sympatric lizard species. Heredity (Edinb) 116:92–98. https://doi.org/10.1038/hdy.2015.76.
69. Charlesworth B. 2009. Effective population size and patterns of molecular evolution and variation. Nat Rev Genet 10:195–205. https://doi.org/10.1038/nrg2526.
70. Willi Y, Fracassetti M, Bachmann O, Van Buskirk J. 2020. Demographic processes linked to genetic diversity and positive selection across a species' range. Plant Commun 1:100111. https://doi.org/10.1016/j.xplc.2020.100111.

71. Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. https://doi.org/10.1093/bioinformatics/btn129.
72. Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinform 27:3070–3071. https://doi.org/10.1093/bioinformatics/btr521.
73. R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
74. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.
75. Galtier N, Gouy M, Gautier C. 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci 12:543–548. https://doi.org/10.1093/bioinformatics/12.6.543.
76. Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27:221–224. https://doi.org/10.1093/molbev/msp259.
77. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
78. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
79. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol 34:772–773. https://doi.org/10.1093/molbev/msw260.
80. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, Du Plessis L, Popinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. PLoS Comput Biol 15:e1006650. https://doi.org/10.1371/journal.pcbi.1006650.
81. Tavaré S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lect Math Life Sci 17:57–86.
82. Rambaut A. 2018. FigTree v.1.4.4. http://tree.bio.ed.ac.uk/software/figtree/. Accessed 28 December 2018.
83. Wielgoss S, Barrick JE, Tenaillon O, Cruveiller S, Chane-Woon-Ming B, Médigue C, Lenski RE, Schneider D. 2011. Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*. G3 (Bethesda) 1:183–186. https://doi.org/10.1534/g3.111.000406.
84. Ochman H, Elwyn S, Moran NA. 1999. Calibrating bacterial evolution. Proc Natl Acad Sci U S A 96:12638–12643. https://doi.org/10.1073/pnas.96.22.12638.
85. Gibson B, Wilson DJ, Feil E, Eyre-Walker A. 2018. The distribution of bacterial doubling times in the wild. Proc R Soc B 285:20180789. https://doi.org/10.1098/rspb.2018.0789.
86. Clark MA, Moran NA, Baumann P. 1999. Sequence evolution in bacterial endosymbionts having extreme base compositions. Mol Biol Evol 16:1586–1598. https://doi.org/10.1093/oxfordjournals.molbev.a026071.
87. Mavromatis K, Doyle CK, Lykidis A, Ivanova N, Francino MP, Chain P, Shin M, Malfatti S, Larimer F, Copeland A, Detter JC, Land M, Richardson PM, Yu XJ, Walker DH, McBride JW, Kyrpides NC. 2006. The genome of the obligately intracellular bacterium *Ehrlichia canis* reveals themes of complex membrane structure and immune evasion strategies. J Bacteriol 188:4015–4023. https://doi.org/10.1128/JB.01837-05.
88. Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174. https://doi.org/10.1007/BF02101694.
89. Brown WM, George M, Wilson A. 1979. Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A 76:1967–1971. https://doi.org/10.1073/pnas.76.4.1967.
90. Nabholz B, Glémin S, Galtier N. 2008. Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. Mol Biol Evol 25:120–130. https://doi.org/10.1093/molbev/msm248.
91. Nabholz B, Glémin S, Galtier N. 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. BMC Evol Biol 9:54. https://doi.org/10.1186/1471-2148-9-54.
92. Bulmer M, Wolfe KH, Sharp PM. 1991. Synonymous nucleotide substitution rates in mammalian genes: implications for the molecular clock and the relationship of mammalian orders. Proc Natl Acad Sci U S A 88:5974–5978. https://doi.org/10.1073/pnas.88.14.5974.