Demographic history influences spatial patterns of genetic diversity in recently expanded coyote (*Canis latrans*) populations

Elizabeth Heppenheimer[1*], Daniela S. Cosio[1], Kristin E. Brzeski[1], Danny Caudill[2,3], Kyle Van Why[4], Michael J. Chamberlain[5], Joseph W. Hinton[5], Bridgett vonHoldt[1]

1. Department of Ecology & Evolutionary Biology, Princeton University, 106A Guyot Hall, Princeton, NJ, 08544
2. Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, 1105 SW Williston Road, Gainesville FL, 32601
3. Alaska Department of Fish Game, 1300 College Road, Fairbanks, AK, 99701
4. United States Department of Agriculture, Animal Plant Health Inspection Service, Wildlife Services, PO Box 60827, Harrisburg, PA, 17106
5. Warnell School of Forestry and Natural Resources, University of Georgia, 180 E Green Street, Athens, GA, 30621

*Corresponding author: Elizabeth Heppenheimer, eh7@princeton.edu, T: 609-258-3571

Keywords: Range expansion, microsatellite survey, geographic cline, population genetics, coyote, *Canis latrans*

Running title: Genetic structure in eastern coyotes

Word count: 6,269

Abstract

Human-mediated range expansions have increased in recent decades and represent unique opportunities to evaluate genetic outcomes of establishing peripheral populations across broad expansion fronts. Over the past century, coyotes (*Canis latrans*) have undergone a pervasive range expansion and now inhabit every state in the continental United States. Coyote expansion into eastern North America was facilitated by anthropogenic landscape changes and followed two broad expansion fronts. The northern expansion extended through the Great Lakes region and southern Canada, where hybridization with remnant wolf populations was common. The southern and more recent expansion front occurred approximately 40 years later and across territory where gray wolves have been historically absent and remnant red wolves were extirpated in the 1970s. We conducted a genetic survey at 10 microsatellite loci of 482 coyotes originating from 11 eastern U.S. states to address how divergent demographic histories influence geographic patterns of genetic diversity. We found that population structure corresponded to a north-south divide, which is consistent with the two known expansion routes. Additionally, we observed extremely high genetic diversity, which is atypical of recently expanded populations and is likely the result of multiple complex demographic processes, in addition to hybridization with other *Canis* species. Finally, we considered the transition of allele frequencies across geographic space and suggest the mid-Atlantic states of North Carolina and Virginia as an emerging contact zone between these two distinct coyote expansion fronts.

Introduction

Range expansions are a common aspect of the natural history of most species (Excoffier *et al.,* 2009). In recent decades, however, substantial range expansions have become increasingly frequent as a result of anthropogenic changes to the climate and landscape (Prugh *et al.*, 2009; Chen *et al.*, 2011). For example, barred owls (Livezey, 2009), northern flying squirrels (Garroway *et al*., 2011), as well as several species of sea turtle (Pike, 2013; Maffucci *et al.,* 2016), have all undergone contemporary range expansions that are at least partially attributed to climate warming or additional human mediated environmental factors. However, despite the ubiquity of range expansions across taxa, empirical studies of the genetic consequences of recent expansions remain relativity rare. As these expansions present species with novel ecological and evolutionary pressures, understanding patterns of genetic diversity and consequently, adaptive potential, in recently expanded populations is of great interest.

Population genetic theory posits that expansion fronts are populated by a small number of individuals, resulting in reduced genetic diversity and the founder effect (Nei *et al.,* 1975; Excoffier *et al.,* 2009). In some cases, dispersers can form “pocket populations” at the expansion front that are subject to high rates of inbreeding, further reducing genetic diversity (Ibrahim *et al.,* 1996). Additionally, a phenomenon known as “allele surfing” may occur during range expansions, in which rare or even deleterious alleles can reach high frequencies at the front of the expansion axis (Edmonds *et al*., 2004; Klopfstein *et al*., 2006). Though there are few empirical examples, allele surfing

69 has been suggested in a range of taxa, including microbes (Gralka *et al.,* 2016), coral
70 snakes (Streicher *et al.,* 2016), and humans (Hofer *et al.,* 2008). More generally,
71 demographic factors such as long distance dispersal events from the source population
72 (Bialozyt *et al.,* 2006) and gene flow (Pfennig *et al.,* 2016) may play an important role in
73 shaping both genetic diversity and structure in recently expanded populations. Long
74 distance dispersal events, in which the expansion front experiences moderate gene flow
75 with the source population, is well documented in plants (e.g. Davies *et al*., 2004; Kremer
76 *et al*., 2012) and has been described in a population of recently introduced European
77 starlings in South Africa (Berthouly-Salazar *et al*., 2013). Additionally, gene flow, with
78 either a closely related species or additional fronts of expansion, can increase genetic
79 diversity through the introduction of novel alleles as well as those that have been lost to
80 drift.

81 Here, we focus on the coyote (*Canis latrans*), which has recently undergone a
82 substantial range expansion from west to east and therefore provides a tractable system to
83 evaluate the demographic and genetic effects of a contemporary range expansion.
84 Coyotes were historically absent from eastern North America but have expanded their
85 range over the last century and now occupy every state in the continental United States.
86 This expansion from the west into eastern North America followed large scale wolf
87 control efforts on the east coast in the 1890s (e.g. Laliberte and Ripple 2004), and it has
88 been suggested that the empty niche space was subsequently filled by northeasterly
89 dispersing coyotes (e.g. Thornton and Murray, 2014). However, this time period also
90 corresponds to the transformation of dense forests into open agricultural land across the

91 eastern landscape, and a combination of these factors likely influenced range expansion
92 (Parker, 1995; Nowacki and Abrams, 2008). This eastward expansion occurred along two
93 spatiotemporally isolated expansion fronts (Figure 1). The first major wave of coyote
94 expansion from the northern Great Plains began in the early twentieth century and
95 consisted of two routes: 1) across the northern Great Lakes region and southern Canada
96 into New England, and 2) along the southern Great Lakes eastward to Pennsylvania
97 (Parker, 1995; Kays *et al.,* 2010). These two fronts likely converged in New York and
98 Pennsylvania during the late 1940s, and now operate as a single front of expansion (Kays
99 *et al.,* 2010; Bozarth *et al.,* 2011). The second major coyote expansion began in the mid-
100 20th century and followed a southeastern route from Texas to the Carolinas by the 1980s
101 (Figure 1; DeBow *et al.,* 1998). These two distinct expansion fronts have experienced
102 different rates of gene flow with other *Canis* species. The northeastern colonization
103 experienced gene flow with remnant wolf populations in the Great Lakes region and
104 Ontario (*C. lupus* and/or *C. lycaon*), confirmed in a genome-wide scan of ancestry
105 (vonHoldt *et al.,* 2011, 2016a). Along the southern expansion front, similar gene flow
106 was documented with remnant red wolves (*C. rufus*; McCarley, 1962; Nowak, 2002;
107 Miller *et al.,* 2003; Hinton *et al.,* 2013; Bohling *et al.,* 2016). Red wolves were later
108 extirpated from the southeastern U.S. in the 1970s, reducing the opportunity for
109 subsequent gene flow.
110 The two expansion fronts are suspected to meet along the mid-Atlantic coast,
111 resulting in evolutionary and ecological consequences. First, overall genetic diversity
112 among populations may increase in a geographically restricted region as a result of

113 increased population connectivity (Hagen *et al.,* 2015). Further, wolf genes have likely
114 entered southern coyote populations through this recent connectivity (vonHoldt *et al.,*
115 2011, 2016b), which alter phenotypic characters and influence adaptive traits.
116 Northeastern coyotes exhibit a phenotype that is distinct from their western counterparts,
117 most notably in overall larger body size and craniodental morphology (Silver and Silver,
118 1969; Kays *et al.,* 2010). This unique phenotype has been attributed to the selective
119 introgression of wolf genes (vonHoldt *et al.,* 2016b) and likely contributed to adaptive
120 differences of northeastern coyotes from other populations (Kays *et al.,* 2010; Thornton
121 and Murray, 2014). For instance, this phenotype is presumed to have enabled
122 northeastern coyotes to hunt larger prey (Benson *et al.,* 2017). Although reports of adult
123 white tailed deer predation are fairly common (e.g. Patterson and Messier, 2003), studies
124 found that coyote diets are highly variable across habitat types (Tremblay *et al.,* 1998)
125 and seasons (Dumond *et al.,* 2001); therefore the full ecological and behavioral
126 implications of the northeastern coyote phenotype are unclear. In contrast, the
127 population-level phenotype of southeastern coyotes has not been extensively quantified,
128 although regional studies suggest southeastern coyotes are smaller in size (Hinton and
129 Chamberlain, 2014) with only a few documented instances of adult deer predation
130 (Chitwood *et al.,* 2014). While much remains unknown about the ecology of eastern
131 coyotes, these findings suggest divergence between the southeastern and northeastern
132 coyotes.
133 Overall, eastern coyote populations represent an opportunity to examine the
134 molecular consequences of how range expansion and secondary contact shape genetic

diversity. We conducted a survey of 10 microsatellite loci of 482 eastern coyotes to evaluate the correspondence between population structure, genetic diversity, and the expansion routes. We predict that genetic structure will be consistent with the two known expansion routes in the northern and southern U.S., respectively. Additionally, though theory predicts low genetic diversity and strong structuring in recent expansion fronts (Excoffier *et al.*, 2009), we hypothesize that the northern expansion front will harbor higher genetic diversity as a result of interspecific breeding with other *Canis* species. Finally, we assess the degree of secondary contact between northern and southern expansion fronts, a likely source of genetic diversity that breaks down population structure. While numerous microsatellite studies have been conducted on northeastern (e.g. Kays *et al.*, 2008; Rutledge *et al.*, 2010) and southeastern coyote populations (e.g. Damm, 2015), few studies have evaluated genetic structure and diversity among the two groups (e.g. Way *et al.*, 2010; Bozarth *et al.*, 2011).

Methods

Study area and sample collection

Coyote whole blood and tissue (e.g. liver, tongue, kidney, etc.) samples (n=482) were obtained from 129 counties in 11 states in the eastern U.S. (Figure 1; Table S1) between 2001 and 2015. In a minority of cases, sampling year was unknown, but believed to be approximately within this timeframe. The majority of samples were collected within a three-year period (2012-2015), which is consistent with the 2-3 year generation time for coyotes (Bekoff and Wells, 1986). Removal of samples collected two

or more years outside of this period, as well as samples with unknown collections years, produced qualitatively identical results in downstream analyses. Most samples were archived by government organizations including Florida Fish and Wildlife, Ohio Department of Natural Resources, and US Department of Agriculture. Additional samples were obtained from state management programs (IACUC # 1961A-13) as well as from the New York State Museum. In most cases, body size, age (e.g. adult, juvenile), and sex, as well as the date and location of capture were recorded. All but two samples from New York had a known county of origin (Table S1), and these samples were therefore excluded from spatially explicit analyses at the county level.

Microsatellite genotyping

DNA was extracted from all samples using the DNeasy Blood and Tissue kit (Qiagen, Louisville, KY) following the instructions provided by the manufacturer and quantified by Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Carlsbad, CA). Water controls were occasionally included to control for contamination. Each sample was then genotyped at 10 highly polymorphic microsatellite loci: FH2001, FH2004, FH2010, FH2137 (Francisco *et al.*, 1996), FH2611, FH2658, FH3399 (Guyon *et al.*, 2003), Pez11, Pez16, and Pez17 (Neff *et al.*, 1999). Polymerase chain reactions were performed using a forward primer with a 5' 16 bp-M13F sequence tag, a fluorescently dye-labeled (6-FAM; Applied Biosystems, Foster City, CA) complement to the M13F tag (Boutin-Ganache *et al.*, 2001) and an unlabeled reverse primer. Reactions were a total volume of 10 µl, and contained 1.5 µl (6 ng) DNA, 1.0 µl primer mix, 0.4 µl 10 mg/ml BSA (New England

Biolabs, Ipswich, MA), 5.0 µl Type-It master mix (Qiagen, Louisville, KY), and 2.1 µl ddH_2O. Cycling conditions consisted of an initial denaturation at 95 °C for 15 min, followed by 25 cycles at 94 °C for 30 s, 59 °C for 90 s, and 72 °C for 60 s, then 15 cycles at 94 °C for 30 s, 53 °C for 90 s, and 72 °C for 60 s, with a final extension at 60 °C for 30 min. To ensure consistent genotyping across reactions, 22 randomly selected samples were amplified ≥3 times. To confirm the absence of contamination, negative controls were also included with each reaction. PCR products were denatured with Hi-Di formamide (Applied Biosystems, Foster City, CA) and LIZ GeneScan 500 size standard (Applied Biosystems, Foster City, CA). PCR fragments were then analyzed on an ABI 3730XL capillary sequencer and genotypes called using `GENEIOUS` v6.1.6 (Kearse *et al.*, 2012). Samples with more than 30% missing data were excluded from the analysis.

Genetic diversity

Observed and expected heterozygosity, pairwise linkage disequilibrium (LD), as well as deviations from Hardy-Weinberg equilibrium (HWE) at each sampling location were evaluated with `ARLECORE` v3.5.2, the console version of `ARLEQUIN` (Excoffier and Lischer, 2010). The exact tests to evaluate LD and HWE were conducted with 1 000 000 steps following 100 000 dememorisation. We calculated additional metrics of genetic distance, including allelic richness (A_R; rarified for eight), with the `R` package `HIERFSTAT` (Goudet, 2005). To assess genetic distance among all sampling locations, we calculated pairwise F_{ST} in `ARLECORE` and evaluated significance using 10 000 permutations and applying a Bonferroni correction for multiple comparisons. Inbreeding

coefficients (F_{IS}) were also estimated in ARLECORE, with significance evaluated using 1000 permutations. We then evaluated the genetic distance among northern and southern sampling locations with a hierarchical locus-by-locus analysis of molecular variance (AMOVA) in ARLECORE. Samples were grouped according to collection location, with Florida, Alabama, Georgia, Louisiana, and South Carolina designated "southern" and New York, Pennsylvania, Ohio, Maryland, Virginia, and North Carolina considered "northern." Within each group, populations were defined as the state of origin for each sample. Additionally, we identified private alleles within the northern and southern groups with GenAlEx v.6.5 (Peakall and Smouse, 2012).

Population structure

We conducted both spatial independent and spatial dependent analyses of eastern coyote population structure. Our spatial independent analysis was implemented in the Bayesian-clustering program STRUCTURE v2.3.4 (Pritchard *et al.,* 2000). With no prior populations assumed, we conducted 10 independent runs for each *K* value with the admixture model for *K*= 1–10, using 500 000 repetitions after a burn-in of 250 000. Output from each independent run was then combined using CLUMPP v64.1.1.2 (Jakobsson and Rosenberg, 2007). The most likely number of genetic clusters represented by the data was estimated by considering both the log-likelihood (LnProbability) values inferred directly from STRUCTURE (Pritchard *et al.,* 2000), as well as Δ*K* (Evanno *et al.,* 2005), which was calculated with STRUCTURE HARVESTER v0.6.94 (Earl and vonHoldt, 2012). While it has been suggested that Δ*K* is a superior indicator of the "true" number of

223 clusters represented by the data (Evanno *et al.*, 2005), this statistic is limited in that it
224 cannot provide support for *K*=1 or the highest *K* value (i.e. *K*=10), as it is based on the
225 rate of change in log-likelihood between successive *K* values. To account for biases
226 induced by uneven sampling, we first subsampled all sampling locations to the smallest n
227 (i.e. five) and reran STRUCTURE. However, this consistently resulted in an optimal K of
228 1 and given that allelic diversity was extremely high, it is likely that five individuals per
229 sampling location does not provide enough power to detect subtle differences in allele
230 frequencies. We therefore addressed this potential bias by removing locations with small
231 sample sizes (n < 15) and conducting an additional STRUCTURE run using the parameters
232 described above.
233 The spatially explicit analysis of population structure was conducted in TESS
234 v.2.3 (Chen *et al.*, 2007) using the BYM admixture model (Durand *et al.*, 2009) for *K*=2–
235 10, with 1 000 000 total sweeps, a burn-in of 250 000, and 10 independent runs per *K*.
236 Geographic information was included as the latitude and longitude of the county centroid
237 from which each coyote was sampled. Though similar to STRUCTURE, TESS additionally
238 incorporates spatial information into clustering assignments and has been suggested to
239 outperform STRUCTURE when populations are weakly differentiated (Chen *et al.*, 2007).
240 To evaluate the optimal *K* value for the TESS analysis, the deviance information criterion
241 (DIC) value was averaged over each independent run and plotted against *K*. Generally,
242 the optimal *K* value corresponds to the plateau of the DIC curve (Durand *et al.*, 2009);
243 however, clustering patterns at successive *K* values were also taken into account when

244 selecting the optimal *K* (Yamashiro *et al.,* 2016). Output over each independent run was
245 combined with CLUMPP prior to graphic representation.
246 For both the STRUCTURE and TESS analyses, we considered individuals to have
247 high assignments to a given inferred cluster if the ancestry proportion (i.e. Q-value) was
248 greater than or equal to 0.8. Further, individuals were considered "admixed" if Q was less
249 than 0.8 for any single inferred cluster (e.g. Rutledge *et al.,* 2010). To evaluate
250 substructure within the northern and southern sampling locations, we reran STRUCTURE
251 including only individuals sampled from northern or southern locations, with identical
252 parameters as described above. Additionally, we evaluated the association of pairwise
253 genetic and geographic distances, that is, the extent of isolation-by-distance (IBD), within
254 northern and southern sampling locations, with a series of Mantel tests implemented in
255 the R package ade4 (Dray and Dufour, 2007). Pairwise genetic distances between
256 sampling locations were calculated as $F_{ST}/(1- F_{ST})$ following Rousset (1997) and
257 geographic distances were calculated as the shortest straight-line distance between state
258 centroids using the Advanced Google Maps Distance Calculator
259 (https://www.daftlogic.com/projects-advanced-google-maps-distance-calculator.htm).
260 Finally, to further visualize clustering in our data, we conducted a centered, unscaled,
261 principal component analysis (PCA) with the R package ADEGENET (Jombart, 2008).
262
263 *Geographic cline analysis*
264 We conducted a geographic cline analysis to describe the transition between
265 divergent groups across the landscape by considering the frequency of genotypes along a

one dimensional geographic transect. Clines are modeled as sigmoidal shaped curves with exponential decay curves on either end (i.e. tails) that can be described mathematically (Szymura and Barton, 1986, 1991). Two of the key cline parameters include cline center, the inflection point of the curve, which indicates the location along the geographic transect where change in trait frequency is most rapid, and cline width, the inverse of the maximum slope, which describes the geographic distance over which this rapid change occurs. Two additional parameters, pMin and pMax, describe the frequency of the focal trait at each end of cline, which indicates the level of trait fixation at each end of the geographic transect.

Locations along a north-south transect were calculated as the shortest straight line distance between each sampling location (i.e. county) and the southernmost site (Collier County, Florida), again using the Advanced Google Maps Distance Calculator. In cases where the shortest straight-line distance encompassed habitat that is obviously unsuitable for coyotes (e.g. the Atlantic Ocean), a pivot point was created near the edge of the suitable habitat range, such that the total number of pivot points needed to avoid the unsuitable habitat was minimized. The distance between sampling locations was then calculated as the sum of the straight line distances passing through the pivot points (Baldassarre *et al.,* 2014). These distances were not intended to simulate animal movement, but to standardize transect distances across land for geographic analyses.

As cline theory assumes all samples were collected along a one dimensional transect (Barton and Hewitt, 1985), we excluded samples from Louisiana, which is approximately 900 km west of the major north-south axis formed by the other sampling

288 locations, and could induce biases as a result of excessive perpendicular sampling. To
289 ensure that the inclusion of Ohio, which is approximately 788 km west of New York, did
290 not induce similar biases, we conducted a second cline analysis excluding all samples
291 originating from Ohio.

292 Lastly, to evaluate additional potential biases in the cline induced by the pivot
293 points, we conducted a third cline analysis in which locations along a north-south transect
294 were calculated as the shortest straight line distance between sampling location and
295 Thomas County, Georgia, which did not require pivot points to avoid major bodies of
296 water. That is, the shortest straight-line distance between Thomas County, Georgia and
297 all other sampled counties extended over land only. This analysis excluded samples
298 collected from Florida, as well as samples from Louisiana.

299 Geographic clines in average ancestry proportion were evaluated using the `R`
300 package `HZAR` v0.2-5, which implements a Metropolis-Hastings Markov chain Monte
301 Carlo algorithm (Derryberry *et al.*, 2014). We chose to examine clinal variation in
302 average ancestry proportions, rather than allele frequency directly, as the loci surveyed in
303 this study were highly polymorphic, and reducing the analysis to the frequency of a
304 single allele per locus did not capture the complexity represented by the data. We fit a
305 total of 15 possible models to the data in addition to a null model of no clinal variation, in
306 which ancestry portions do not vary across the landscape. All models estimated cline
307 center and width, but incorporated all possible combinations of scaling (pMin and pMax
308 fixed at 0 and 1, fixed to observed values, free parameters) and tail parameters (no tails,
309 right tail only, left tail only, both tails mirrored about cline center, both tails estimated

independently). We further estimated the two log-likelihood support limits (analogous to 95% confidence intervals) around both the cline center and width.

Model selection was based on Akaike's Information Criterion (Akaike, 1973), corrected for sample size (AICc), with the lowest AICc score indicative of the best model (Burnham and Anderson, 2002; Derryberry *et al.*, 2014). The estimations for cline center and width given by strongly competing models (ΔAICc < 2) were also considered in the identification of the contact zone between the northern and southern populations, with clines considered coincident and concordant if the two log-likelihood support limits around the center and width, respectively, were overlapping.

Results

Genotyping and genetic diversity

We genotyped 482 coyotes from 11 states and 129 counties at 10 microsatellite loci. All loci were highly polymorphic, with the number of alleles per locus ranging 7–35 with an average of 16.7 (Table S2). Allelic richness (A_R), a metric of allelic diversity that accounts for differences in sample size, ranged 3.78–6.33 (average = 4.95; Table 1, S2). Observed heterozygosity values were similarly high across all sampling locations, ranging 0.780–0.864 (average H_O = 0.838; Table 1). We did not observe any significant deviations from HWE after applying a Bonferroni correction for multiple tests (α =0.05, adjusted $p > 4.5x10^{-4}$). However, following Bonferroni correction, two loci significantly deviated from linkage equilibrium (FH2001 and FH2137) in Ohio samples ($p = 2 X 10^{-5}$). Removal of these loci from Ohio samples did not impact population level trends of

genetic structure (results not shown). F_{IS} values were low across all sampling locations (F_{IS} average=-0.0495, range =-0.0962–0.04) and none were significantly different from zero (adjusted $p > 0.0045$). Overall, pairwise F_{ST} values varied between all locations (F_{ST} average = 0.020, range = -0.004–0.057). Following Bonferroni correction, 15 of these 55 pairwise F_{ST} values were significantly different from zero, 12 of which were north-south comparisons (e.g. PA and SC: $F_{ST} = 0.041$, $p < 10^{-5}$), two were south-south comparisons (e.g. SC and FL: $F_{ST} = 0.019$, $p = 6.9 \times 10^{-4}$), and one was a north-north comparison (PA and OH: $F_{ST} = 0.006$, $p < 10^{-6}$; Table 2). The AMOVA indicated significant genetic distance between northern and southern sampling locations ($F_{CT} = 0.017$, $p < 10^{-5}$). Further, variation among populations (i.e. states) within groups was also significant ($F_{SC} = 0.010$, $p < 10^{-5}$), as was variation within populations ($F_{ST} = 0.027$, $p < 10^{-5}$). We identified 23 and 20 private alleles in the northern and southern groups, respectively, all of which were relatively at low frequency (average: 0.016, range: 0.002-0.088; Table S3).

Population structure

Our spatial independent analysis of population structure provided support for two distinct populations, as both ΔK and the mean LnProbability converged on K=2 (Figure S1). These two inferred clusters corresponded to an approximate north-south divide, with the majority of samples originating from New York, Pennsylvania, Ohio, Maryland, Virginia, and North Carolina composing one genetic cluster, and the majority of samples originating from South Carolina, Louisiana, Georgia, Alabama, and Florida in the second genetic cluster (Figure 2A). However, we identified 48 and 22 admixed individuals in the

354 northern and southern populations, respectively. We further identified 13 individuals that
355 were sampled from the north, but exhibited higher membership to the southern
356 population, and three individuals sampled from the south that clustered with the north
357 (Figure 2). Finally, removing locations with small sample size (n < 15) produced
358 qualitatively identical results with regard to a north-south divide (Figure S3).
359 Overall, the spatially explicit analysis in TESS yielded similar results to the
360 STRUCTURE analysis (Figure 2B). The plateau of the DIC curve occurred at *K*=5;
361 however, clustering patterns above *K*=2 did not reveal a new distinct population, but
362 rather suggested admixture with an unsampled population in the mid-Atlantic states of
363 Pennsylvania, Ohio, Maryland, and Virginia (Figure S2). These results suggest that two
364 clusters are optimal, but that these two populations likely experience different rates of
365 gene flow with a neighboring unsampled population. Average individual level ancestry
366 proportions within these two clusters were similar to STRUCTURE and showed a clear
367 geographic north-south divide (Figure 2B). While in the STRUCTURE analysis the
368 majority of samples originating from North Carolina clustered with the northern
369 population (eight out of 19) and a minority of samples were either admixed (six out of
370 19) or clustered with the southern population (five out of 19), TESS revealed eight North
371 Carolina coyotes clustered with the north, one with the south, and the remaining ten were
372 admixed. Further, 23 and 42 additional admixed coyotes were identified in the northern
373 and southern populations, respectively, and one coyote from Ohio clustered with the
374 south (Figure 2B).

In our evaluation of substructure within sampling locations, we found support for an optimal *K* of 1 for individuals collected from southern sampling locations (Figure S4A) and no evidence of IBD between southern sampling locations (Mantel: r=0.564, p=0.124; Figure S5A). However, for the individuals sampled from the north, support for an optimal *K* of 1 was comparable to support for an optimal *K* of 2 (Figure S4B). At *K*=2, clustering patterns suggested admixture with a neighboring unsampled population, particularly in Ohio (Figure S4C), a pattern similar to what was observed at *K*=3 in the TESS analysis (Figure S2). Despite this potential weak substructuring, we found no evidence for a correlation between genetic and geographic distance between northern sampling locations (Mantel test: r=0.358, p=0.146; Figure S5B). Lastly, the PCA clearly separated coyotes by the expansion fronts and by geography (PC1, 3.57% variation explained; Figure 3). However, the pattern of separation along PC2 (3.43% variation) did not follow a geographic pattern, suggesting that within each population, coyotes in neighboring states are not necessarily the most genetically similar, further indicating a lack of substantial substructure.

Geographic cline analysis

Our data clearly showed clinal variation in average ancestry proportions per sampling location along a 2072 km north-south transect extending from Collier County, Florida to Hamilton County, New York (Figure 4A). The best-fit cline model estimated pMin and pMax as free parameters (0.075 and 0.924, respectively) and did not fit decay tails about the cline center. The cline center was estimated at 1218 km and cline width

was 579 km. To determine the approximate geographic location of the cline center, we identified eleven counties with transect distances within the two log-likelihood support limits of the maximum likelihood estimated (MLE) cline center (2LLc=1106–1322 km), seven of which were located in southwest Virginia and four were in northern North Carolina (Figure 4B). Of these ten counties, Beaufort County, North Carolina had a transect distance most similar to the MLE cline center, at 1219 km from Collier County, Florida. In addition to this model, four models were within two ΔAICc units of the lowest AICc score (Table S4). These models provided similar estimates of cline center (range: 1218-1228 km, average: 1218 km) and larger estimates of cline width (range: 778-844 km, average: 822 km). However, all estimates of cline center and width for these alternative models were within the two log-likelihood support limits calculated for the model with the lowest AICc score (2LLc: 1106-1322 km; 2LLw: 296 -907; Table S4) indicating that these clines are both coincident and concordant and identify approximate the same geographic area as the selected model. Additionally, the removal of sampling locations in Ohio did not appreciably alter estimates of cline center (1224 km; 2LLc: 1122-1302 km) or width (775 km; 2LLw: 330-993 km).

Our third cline analysis, which extended from Thomas County, Georgia to Hamilton County, New York (and excluded Florida and Louisiana), also showed similar results (Figure S6). The best fit model for the cline in average Q-value fixed pMin and pMax at the observed values (0.022 and 0.986, respectively), and did not fit exponential decay tails. The cline center was estimated at 708 km from Thomas County, Georgia with a cline width of 882 km. One alternative model was within 2 ΔAICc units of the lowest

AICc score, however this model provided similar estimates of both cline center (700 km) and width (956 km), with overlapping two log-likelihood support limits (Table S5). For the selected model, the same seven counties in southwest Virginia were within two log-likelihood units of the MLE cline center (599 – 796 km), with Carroll County closest to the cline center (711 km). However, the four counties in North Carolina identified by the first cline had transect distances slightly outside of this range (825, 872, 875, 908 km). These results suggest that the first cline was not substantially biased by the use of pivot points to avoid unsuitable coyote habitat, as despite minor differences, approximately the same geographic area was identified in both analyses.

Discussion

Over the past century, coyotes colonized eastern North America along two discrete expansion fronts that occurred during distinct periods of the 20th century and differed in the frequency of hybridization events with other *Canis* species. Our analyses provide evidence for two genetically distinct regional populations of coyotes that correspond to the two known historic colonization routes. These findings are consistent with our expectations that divergent demographic histories result in observable genetic differences among groups of conspecifics at fine temporal scales.

Despite this clear geographic separation of coyote populations, it is likely that coyotes originating from distinct expansion fronts have begun to overlap in range, forming a contact zone between the two previously isolated populations. While other studies have suggested intraspecific gene flow between northern and southern coyote

populations (Bozarth *et al.*, 2011; vonHoldt *et al.*, 2011, 2016b), our results provide the first estimate of the precise geographic space over which this contact zone occurs. We observed a latitudinal gradient of ancestry proportions in eastern coyotes, where the most rapid change in ancestry proportion occurred in the mid-Atlantic region. There are two primary explanations for this clinal distribution: selection against admixture among the two populations, or alternatively neutral demographic processes. Selection typically results in a steep change in frequency over a short geographic distance and narrow estimates of cline width, while neutral processes would form an initially steep clinal transition that gradually widens over time as populations homogenize (Barton and Hewitt, 1985). In this study, our estimated cline width was sizable (MLE: 579 km), suggesting that this cline is driven by demography and recent contact among the two groups, rather than selective processes. However, we surveyed neutral loci, which may introgress readily even in the presence of selection against admixture (e.g. Gompert *et al.*, 2012). It is therefore unclear if this cline will persist over generations due to selection or if we have simply captured the initial stages of homogenization between the two groups. Future studies should address the change in cline over time as well as investigate clines in the frequency of functionally relevant alleles to address the role of selection in shaping our observed cline. Finally, it is also theoretically possible to observe clines in allele frequencies under a pure isolation-by-distance scenario within a single linage (Wright, 1943). In this case, however, the known spatiotemporal isolation of the coyote expansion fronts suggests that eastern coyotes are unlikely to represent a single lineage, and that our observed cline is attributable to secondary contact.

While cline analyses are useful for addressing the change in allele frequencies across a landscape, they are limited in that multidimensional sampling locations are collapsed into a one-dimensional transect. In this study, the 129 counties sampled across the eastern U.S. did not follow a perfectly one dimensional transect and we acknowledge that multidimensional sampling may bias cline analyses (e.g. Dufkov, 2011). However, we found no evidence for substructure in the southern cluster and only weak evidence for substructure in the northern population, perhaps as a result of unequal gene flow with a neighboring mid-western population. Though we did detect significant genetic distance between Ohio and Pennsylvania, which may be attributed to a high rate of gene flow in Ohio, removal of Ohio from the cline analysis did not appreciably change the MLE of cline center or width. These results suggest variation in allele frequencies perpendicular to the major north-south transect in the geographic regions surveyed was not substantial enough to have markedly biased our results.

Generally, this lack of substructure within each inferred cluster was interesting, as fine scale population structure has been documented in California coyotes (Sacks *et al*., 2004). However, population structure in these western coyotes corresponds to habitat breaks across the landscape (Sacks *et al*., 2004; Sacks *et al*., 2008), whereas the eastern United States represents a more homogeneous ecoregion (U.S. Environmental Protection Agency, 2017). That is, major landscape changes occur over much larger geographic distances on the east coast than in central California, increasing the scale at which habitat breaks could influence population structure.

Interestingly, we observed extremely high heterozygosity and allelic diversity in both expansion fronts, which is atypical of recently expanded populations (Excoffier *et al.,* 2009) and contrary to our expectations that the northern population would be more diverse as a result of more extensive gene flow with wolf populations. Though we acknowledge that the microsatellite markers used in this study have a high mutation rate (Irion et al., 2003) and are ascertained based on diversity, heterozygosity levels in eastern coyotes are approximately 10% higher than those reported for a subset of the same markers in California coyote populations (H_E= 0.76; Sacks *et al.,* 2004), suggesting that this pattern is not entirely a methodological artifact. Although it is not immediately clear from these results how coyote populations along both expansion fronts were able to maintain such high genetic diversity, several demographic processes could mitigate these expected patterns of reduced genetic diversity and increase the adaptive potential of an expanding population. For example, if the expansion front is subject to even a modest rate of gene flow from the source population as a result of long distance dispersal, genetic diversity at the periphery could be maintained (e.g. Alleaume-Benharira *et al.,* 2006; Berthouly-Salazar *et al.,* 2013). Coyotes are known to be highly mobile, with male and female coyotes observed to disperse up to 102.5 km from their natal range, likely playing an important role in the maintenance of genetic diversity (Harrison, 1992; Mastro, 2011; Hinton *et al.,* 2012, 2015).

It is also possible that in the early stages of expansion, coyotes along the range periphery did experience decreased genetic diversity as predicted by population genetic theory. Over the past 50 years, however, rapid increases in population size and

connectivity, in combination with the high mutation rate of microsatellites, may have reintroduced genetic diversity and counteracted the impact of a historic bottleneck on contemporary populations. This phenomenon has been documented in an introduced population of rabbits in Australia (Zenger *et al.,* 2003), as well as observed in as few as 1.5 generations in European brown bears (Hagen *et al.,* 2015). As coyotes have a generation time of 2-3 years (Bekoff and Wells, 1986), it is conceivable that a marked increase in genetic diversity could be observed after 50 years.

Finally, coyotes along both expansion fronts are known to have interbred with other *Canis* species (Bohling *et al.,* 2016; Kays *et al.,* 2010). Though interspecific hybridization is often synonymous with outbreeding depression (e.g. Muhlfeld 2009), this process may play a more beneficial role for closely related species through the introduction of novel or advantageous variation (Hedrick, 2013). In the case of range expansion, hybridization could increase overall genetic diversity as well as introduce locally adapted genes to populations at the expansion front (Pfennig *et al*., 2016). The introgression of locally adapted genes through interspecific gene flow has previously been suggested to have facilitated range expansion not only in northeastern coyotes (Kays *et al*. 2010), but also in *Anopheles* mosquitos by transferring genes critical to adaptation to arid environments (Besansky *et al*. 2003; Pfennig *et al*., 2016). While the markers utilized in this study are not sensitive enough to detect signatures of interspecific hybridization, the impact of interspecific hybridization events on population level genetic diversity remains an important area for future research.

Overall, it is important to note that these three demographic processes, long-range dispersal, recent increases in population size and connectivity, and interspecific hybridization, are not mutually exclusive and a combination of factors likely contributed to the observed genetic diversity in eastern coyote populations. For instance, heterozygosity was similarly high in both the northern and southern populations, despite interspecific hybridization occurring at higher frequency in the north and therefore, other demographic processes likely contributed to the observed pattern across all sampling locations. Future studies should address genome-wide trends in heterozygosity to elucidate the contribution of interspecific hybridization, in addition to other demographic processes, to genetic diversity across eastern coyote populations.

Acknowledgements

We gratefully acknowledge the Princeton Department of Ecology and Evolutionary Biology Leslie K. Johnson Senior Thesis Fund, the Princeton Environmental Institute, and the Florida Fish and Wildlife Conservation Commission for partial funding. We thank Suzanne Prange, Scott Woodruff, Roland Kays, Jon Donaldson, Eric Wilhelm, Chad Fox, Tom Elliot, Harold McDaniel, Robert K. Wayne, the US Department of Agriculture, Ohio Department of Natural Resources, and numerous hunters and trappers for their generous donation of samples. We also acknowledge Linda Rutledge and Rebecca Kartzinel for their analytical advice.

Data Accessibility

Sampling locations and microsatellite genotypes for all individuals will be available on Dryad Digital Repository (datadryad.org) upon acceptance.

Conflict of Interest

The authors declare no conflicts of interest.

Supplemental

Supplementary information is available through *Heredity's* website (http:/www.nature.com/hdy).

References

Akaike H (1973). Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki BF (eds) *Second International Symposium on Information Theory*. Akademiai Kiado: Budapest, Hungary. pp. 267–281.

Alleaume-Benharira M, Pen IR, Ronce O (2006). Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *J Evolution Biol* **19**: 203–215.

Baldassarre DT, White TA, Karubian J, Webster MS (2014). Genomic and morphological analysis of a semipermeable avian hybrid zone suggests asymmetrical introgression of a sexual signal. *Evolution* **68**: 2644–2657.

Barton NH, Hewitt GM (1985). Analysis of hybrid zones. *Annu Rev Ecol Evol S* **16**: 113–148.

Bekoff M, Wells MC (1986). Social ecology and behavior of coyotes. *Adv Stud*

572 *Behav* **16**: 251–338.
573 Benson JF, Loveless KM, Rutledge LY, Patterson BR (2017). Ungulate predation and
574 ecological roles of woles and coyotes in eastern North America. *Ecol Appl*
575 doi:10.1002/eap.1499.
576 Berthouly-Salazar C, Hui C, Blackburn TM, Gaboriaud C, Van Rensburg BJ, Van
577 Vuuren BJ, *et al.* (2013). Long-distance dispersal maximizes evolutionary
578 potential during rapid geographic range expansion. *Mol Ecol* **22**: 5793–5804.
579 Besansky NJ, Krzywinski J, Lehmann T, Simard F, Kern M, Mukabayire O *et al.*
580 (2003). Semipermeable species boundaries between *Anopheles gambiae* and
581 *Anopheles arabiensis*: evidence from multilocus DNA sequence variation.
582 *PNAS* **100**:10818–10823.
583 Bialozyt R, Ziegenhagen B, Petit RJ (2006). Contrasting effects of long distance seed
584 dispersal on genetic diversity during range expansion. *J Evolution Biol* **19**:
585 12–20.
586 Bohling JH, Dellinger J, McVey JM, Cobb DT, Moorman CE, Waits LP (2016).
587 Describing a developing hybrid zone between red wolves and coyotes in
588 eastern North Carolina. *Evol Appl* **9**: 791–804.
589 Boutin-Ganache I, Raposo M, Raymond M, Deschepper CF (2001). M13-tailed
590 primers improve the readability and usability of microsatellite analyses
591 performed with two different allele-sizing methods. *BioTechniques* **31**: 24–28.
592 Bozarth CA, Hailer F, Rockwood LL, Edwards CW, Maldonado JE (2011). Coyote
593 colonization of northern Virginia and admixture with Great Lakes wolves. *J*

Mammal **92**: 1070–1080.

Burnham KP, Anderson DR (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer-Verlag: New York, USA.

Chen C, Durand E, Forbes F, Francois O (2007). Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol Ecol Notes* **7**: 747–756.

Chen I, Hill JK, Ohlemuller R, Roy DB, Thomas CD (2011). Rapid range shifts of species associated with high levels of climate warming. *Science* **333**: 1024-1026.

Chitwood MC, Lashley MA, Moorman CE, DePerno CS (2014). Confirmation of coyote predation on adult female white-tailed deer in the southeastern United States. *Southeast Nat* **3**: N30–N32.

Damm DL, Armstrong JB, Arjo WM, Piaggio AJ (2015). Assessment of population structure of coyotes in east-central Alabama using microsatellite DNA. *Southeast Nat* **14**: 106–122.

Davies S, White A, Lowe A (2004). An investigation into effects of long-distance seed dispersal on organelle population genetic structure and colonization rate: a model analysis. *Heredity* **93**:566–576.

DeBow TM, Webster WMD, Sumner PW (1998). Range expansion of the coyote, *Canis latrans* (Carnivora: Canidae), into North Carolina; with comments on some management implications. *J Elisha Mitch Sci S* **114**: 113–118.

Derryberry EP, Derryberry GE, Maley JM, Brumfield RT (2014). Hzar: Hybrid zone

616 analysis using an R software package. *Mol Ecol Resour* **14**: 652–663.
617 Dray S, Dufour AB (2007): The ade4 package: implementing the duality diagram for
618 ecologists. *J Stat Softw* **22**: 1-20.
619 Dufkov P, Macholan M, Pialek J (2011). Inference of selection and stochastic effects
620 in the house mouse hybrid zone. *Evolution* **65**: 993–1010.
621 Dumond M, Villard M, Tremblay E (2001). Does coyote diet vary seasonally
622 between a protected and an unprotected forest landscape? *Ecoscience* **8**: 301–
623 310.
624 Durand E, Jay F, Gaggiotti OE, Francois O (2009). Spatial inference of admixture
625 proportions and secondary contact zones. *Mol Biol Evol* **26**: 1963–1973.
626 Earl DA, vonHoldt BM (2012). STRUCTURE HARVESTER: a website and program
627 for visualizing STRUCTURE output and implementing the Evanno method.
628 *Conserv Genet Resour* **4**: 359–361.
629 Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004). Mutations arising in the wave
630 front of an expanding population. *PNAS* **101**:975-979.
631 Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of
632 individuals using the software structure: a simulation study. *Mol Ecol* **14**:
633 2611–2620.
634 Excoffier L, Foll M, Petit RJ (2009). Genetic consequences of range expansions.
635 *Annu Rev Ecol Evol S* **40**: 481–501.
636 Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: a new series of programs to
637 perform population genetics analyses under Linux and Windows. *Mol Ecol*

638 *Resour* **10**: 564–567.
639 Francisco LV, Langston AA, Mellersh CS, Neal CL, Ostrander EA (1996). A class of
640 highly polymorphic tetranucleotide repeats for canine genetic mapping.
641 *Mamm Genome* **7**: 359–362.
642 Garroway CJ, Bowman J, Holloway GL, Malcolm JR, Wilson PJ (2011). The genetic
643 signature of rapid range expansion by flying squirrels in response to
644 contemporty climate warming. *Global Change Biol* **17**: 1760-1769.
645 Gompert Z, Parchman TL, Buerkle CA (2012). Genomics of isolation in hybrids. *Phil*
646 *Trans R Soc B* **367**: 439-450.
647 Goudet J (2005). Hierfstat, a package for R to compute and test hierarchical F-
648 statistics. *Mol Ecol Notes* **5**: 184–186.
649 Gralka M, Stiewe F, Farrell F, Mobius W, Waclaw, B, Hallatschek, O (2016). Allele
650 surfing promotes microbial adaption from standing variation. *Ecology Letters*
651 **19:** 889-898.
652 Guyon R, Lorentzen TD, Hitte C, Kim L, Cadieu E, Paker HG *et al.* (2003). A 1-Mb
653 resolution radiation hybrid map of the canine genome. *PNAS* **100**: 5296–5301.
654 Harrison DJ (1992). Social ecology of coyotes in northeastern North America:
655 relationships to dispersal, food resources, and human exploitation. In: Boer
656 AH (ed) *Ecology and Management of the Eastern Coyote*, Wildlife Research
657 Unit, University of New Brunswick: Fredericton, New Brunswick, Canada.
658 pp. 53-72.
659 Hagen SB, Kopatz A, Aspi J, Kojola I, Eiken HG (2015). Evidence of rapid change in

genetic structure and diversity during range expansion in a recovering large terrestrial carnivore. *Proc R Soc B* **282**: 20150092.

Hedrick PW (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol Ecol* **22**: 4606–4618.

Hinton JW, Chamberlain MJ, van Manen FT (2012). Long-distance movements of transient coyotes in eastern North Carolina. *Am Midl Nat* **168**: 281–288.

Hinton JW, Chamberlain MJ, Rabon DR (2013). Red wolf (*Canis rufus*) recovery: a review with suggestions for future research. *Animals* **3**: 722–744.

Hinton JW, Chamberlain MJ (2014). Morphometrics of *Canis* taxa in eastern North Carolina. *J Mammal* **95**: 855–861.

Hinton JW, van Manen FT, Chamberlain MJ (2015). Space use and habitat selection by resident and transient coyotes (*Canis latrans*). *PLoS ONE* **10**: e0132203.

Hofer T, Ray N, Wegmann, D, Excoffier L (2008). Large allele frequency differences between human contental groups are more likely to have occurred by drift during range expansions than by selection. *Ann Hum Genet* **73**: 95-108.

Ibrahim KM, Nichols RA, Hewitt GM. Spatial patters of genetic variation by different forms of dispersal during range expansion. *Heredity* **77**:282-291.

Irion DN, Schaffer AL, Famula, TR, Eggleston ML, Hughes SS, Pedersen NC (2003). Analysis of genetic variation in 28 dog breed populations with 100 microsatellite markers. *J Hered* **94**: 81-87.

Jakobsson M, Rosenberg NA (2007). CLUMPP: a cluster matching and permutation

682 program for dealing with label switching and multimodality in analysis of
683 population structure. *Bioinformatics* **23**: 1801–1806.
684 Jombart T (2008). Adegenet: a R package for the multivariate analysis of genetic
685 markers. *Bioinformatics* **24**: 1403–1405.
686 Kays R, Curtis A, Kirchman JJ (2010). Rapid adaptive evolution of northeastern
687 coyotes via hybridization with wolves. *Biol Letters* **6**: 89–93.
688 Kays RW, Gompper ME, Ray JC (2008). Landscape ecology of eastern coyotes based
689 on large-scale estimates of abundance. *Ecological Applications* **18**: 1014–
690 1027.
691 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S *et al*. (2012).
692 Geneious Basic: an integrated and extendable desktop software platform for
693 the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–
694 1649.
695 Klopfstein S, Currat M, Excoffier L (2006). The fate of mutations surfing on the
696 wave of a range expansion. *Mol Biol Evol* **23**:482-490.
697 Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R *et al.*
698 (2012). Long-distance gene flow and adaptation of forest trees to rapid climate
699 change. *Ecology Letters* **15**: 378–392.
700 Laliberte AS, Ripple WJ (2004). Range contractions of North American carnivores
701 and ungulates. *BioScience* **54**: 123-138.
702 Livezey KB (2009). Range expansion of barred owls, part II: facilitating ecological
703 changes. *Am Midl Nat* **161**: 323-349.

Maffucci F, Corrado R, Palatella L, Borra M, Marullo S, Hochscheid S *et al*. (2016). Season heterogeneity of ocean warming: a mortality sink for ectotherm colonizers. *Sci Rep* **6:** 23983.

Mastro LL (2011). Life history and ecology of coyotes in the mid-Atlantic states: a summary of the scientific literature. *Southeast Nat* **10**: 721–730.

McCarley H (1962). The Taxonomic status of wild *Canis* (Canidae) in the south central United States. *Southwest Nat* **7**: 227–235.

Miller CR, Adams JR, Waits LP (2003). Pedigree-based assignment tests for reversing coyote (*Canis latrans*) introgression into the wild red wolf (*Canis rufus*) population. *Mol Ecol* **12**: 3287-3301.

Muhlfeld CC, Kalinowski ST, McMahon TE, Taper ML, Painter S, Leary RF *et al.* (2009). Hybridization rapidly reduces fitness of a native trout in the wild. *Biol Letters* **5**: 328–331.

Neff MW, Broman KW, Mellersh CS, Ray K, Acland GM, Aguirre GD *et al.* (1999). A second-generation genetic linkage map of the domestic dog, *Canis familiaris*. *Genetics* **151**: 803–820.

Nei M, Maruyama T, Chakraborty R (1975). The Bottleneck effect and genetic variability in populations. *Evolution* **29**: 1–10.

Nowacki GJ, Abrams MD (2008). The demise of fire and “mesophication” of forests in the eastern United States. *BioScience* **58**: 123–138.

Nowak RM (2002). The original status of wolves in eastern North America. *Southeast Nat* **1**: 95–130.

726 Parker G (1995). *Eastern coyote: the story of its success*. Nimbus Publishing: Halifax,
727 Nova Scotia, Canada.
728 Patterson BR, Messier F (2003). Age and condition of deer killed by coyotes in Nova
729 Scotia. *Can J Zoolog* **81**:1894–1898.
730 Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population
731 genetic software for teaching and research-an update. *Bioinformatics* **28**:2537-
732 2539.
733 Pfennig KS, Kelly AL, Pierce AA (2016). Hybridization as a facilitator of species
734 range expansion. *Proc R Soc B* **283**: 20161329.
735 Pike D (2013). Forecasting range expansion into ecological traps:climate mediated
736 shifts in sea turtle nesting beaches and human development. *Glob Chang Biol*
737 **19**:3082-3092.
738 Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using
739 multilocus genotype data. *Genetics* **155**: 945–959.
740 Prugh LR, Stoner CJ, Epps CW, Bean WT, Ripple WJ, Laliberte AS *et al*. (2009).
741 The rise of the mesopredator. *BioScience* **59**: 779–791.
742 Rousset F (1997). Genetic differentiation and estimation of gene flow from F-
743 statisitics under isolation by distance. *Genetics* **145**: 1219-1228.
744 Rutledge LY, Garroway CJ, Loveless KM, Patterson BR (2010). Genetic
745 differentiation of eastern wolves in Algonquin Park despite bridging gene
746 flow between coyotes and grey wolves. *Heredity* **105**: 520–531.
747 Sacks BN, Brown SK, Ernest HB (2004). Population strcuture of California coyotes

correponds to habitat-specific breaks and illuminates species history. *Molec Ecol* **13**:1265-1275.

Sacks BN, Bannasch DL, Chomel BB, Ernest HB (2008). Coyotes demonstrate how habitat specialization by individuals of a generalist species can diversify populations in heterogeneous ecoregion. *Mol Biol Evol* **25**:1384–1394.

Silver H, Silver WT (1969). Growth and behavior of the coyote-like Canid of northern New England and observations on Canid hybrids. *Wildlife Monographs* **17**:1-41

Streicher JW, McEntee JP, Drzich LC, Card DC, Scield DR, Smart U (2016). Genetic surfing, not allopatric divergence, expalins spatial sorting of mitochondrial haplotypes in venomous coralsnakes. *Evolution* **70**: 1435-1449.

Szymura JM, Barton NH (1986). Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B.variegata*, near Cracow in southern Poland. *Evolution* **40**: 1141–1159.

Szymura JM, Barton NH (1991). The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and B. *variegata*: comparisons between transects and between loci. *Evolution* **45**: 237–261.

Thornton DH, Murray DL (2014). Influence of hybridization on niche shifts in expanding coyote populations. *Divers Distrib* **20**: 1355–1364.

Tremblay JP, Crête M, Huot J (1998). Summer foraging behaviour of eastern coyotes in rural versus forest landscape: a possible mechanism of source-sink dynamics. *Ecoscience* **5**: 172–182.

U.S. Environmental Protection Agency. 2017. Ecoregions of North America. November 2016. https://www.epa.gov/eco-research/ecoregions. Accessed 15 February 2017.

vonHoldt BM, Pollinger JP, Earl DA, Knowles JC, Boyko AR, Parker H *et al.* (2011). A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Res* **21**: 1294–1305.

vonHoldt BM, Cahill JA, Fan Z, Gronau I, Robinson J, Pollinger JP *et al.* (2016a). Whole-genome sequence analysis shows that two endemic species of North American wolf are admixtures of the coyote and gray wolf. *Sci Adv* **2**: e1501714.

vonHoldt BM, Kays R, Pollinger JP, Wayne RK (2016b). Admixture mapping identifies introgressed genomic regions in North American canids. *Mol Ecol* **25**: 2443-2453.

Way JG, Rutledge L, Wheeldon T, White BN (2010). Genetic characterization of eastern “coyotes ” in eastern Massachusetts. *Northeast Nat* **17**: 189–204.

Wright S (1943). Isolation by distance. *Genetics* **28**: 28: 114-138.

Yamashiro T, Yamashiro A, Inoue M, Maki M (2016). Genetic diversity and divergence in populations of the threatened grassland perennial *Vincetoxicum atratum* (Apocynaceae-Asclepiadoideae) in Japan. *J Hered* **107**: 455–462.

Zenger KR, Richardson BJ, Vachot-Griffin AM (2003). A rapid population expansion retains genetic diversity within European rabbits in Australia. *Mol Ecol* **12**: 789–794.

Table 1. Diversity statistics across sampling locations and average over all locations. State abbreviations given in in parentheses. (Abbreviations: sample size, n; average number of alleles per locus, A; allelic richness, A_R; observed heterozygosity, H_O; expected heterozygosity, H_E; inbreeding coefficient, F_{IS}).

	Location	n	A	A_R	H_O	H_E	F_{IS}
North	New York (NY)	5	4.9	4.47	0.78	0.809	0.040
	Pennsylvania (PA)	163	13.8	4.89	0.821	0.833	-0.096
	Ohio (OH)	85	13.2	5.08	0.839	0.848	-0.075
	Maryland (MD)	7	6.3	4.78	0.838	0.819	-0.079
	Virginia (VA)	18	9.3	5.15	0.827	0.856	-0.093
	North Carolina (NC)	19	10.2	5.16	0.864	0.852	-0.053
	OVERALL NORTH	297	9.62	4.92	0.828	0.836	
South	South Carolina (SC)	10	7.5	4.98	0.849	0.846	-0.014
	Georgia (GA)	22	10.2	5.09	0.830	0.830	-0.025
	Alabama (AL)	15	9.2	5.09	0.854	0.834	-0.055
	Louisiana (LA)	5	5.2	4.70	0.86	0.838	-0.050
	Florida (FL)	133	13.4	5.09	0.854	0.849	-0.045
	OVERALL SOUTH	185	9.10	4.99	0.849	0.839	

799

800 **Table 2.** Pairwise F_{ST} values among all sampling locations.

	NY	PA	OH	MD	VA	NC	SC	LA	GA	FL	AL
NY	-	-	-	-	-	-	-	-	-	-	-
PA	0.002	-	-	-	-	-	-	-	-	-	-
OH	0.031	0.006*	-	-	-	-	-	-	-	-	-
MD	0.038	0.006	0.013	-	-	-	-	-	-	-	-
VA	0.014	0.008	-0.004	0.011	-	-	-	-	-	-	-
NC	0.009	0.003	0.010	0.010	-0.001	-	-	-	-	-	-
SC	0.057	0.041*	0.038*	0.014	0.013	0.017	-	-	-	-	-
LA	0.051	0.038	0.029	0.019	0.014	0.037	0.034	-	-	-	-
GA	0.018	0.029*	0.033*	0.028*	0.007	0.012	0.011	0.029	-	-	-
FL	0.031	0.023*	0.024*	0.027*	0.003	0.015*	0.019*	0.024	0.017*	-	-
AL	0.030	0.028*	0.025*	0.037*	0.011	0.013	0.017	0.031	0.007	0.004	-

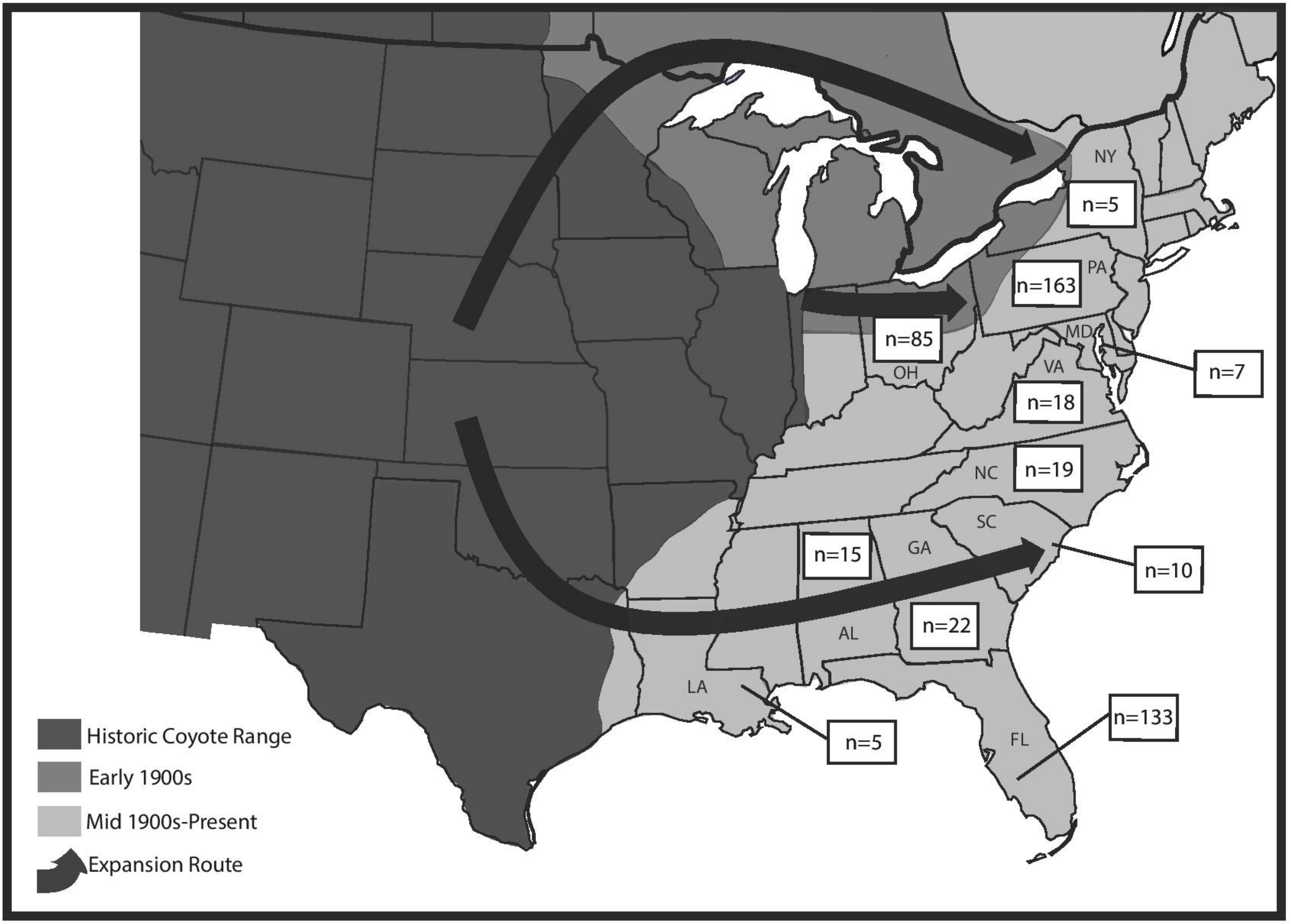
801 * significant at α =0.05 after Bonferroni correction (adjusted $p < 0.0009$).

Figure 1. Map of eastern coyote range expansion and sampling locations. Historic range and expansion routes are approximate and modified from Parker (1995), Nowak (2002), and Kays *et al.* (2010). (Abbreviations: sample size, n).

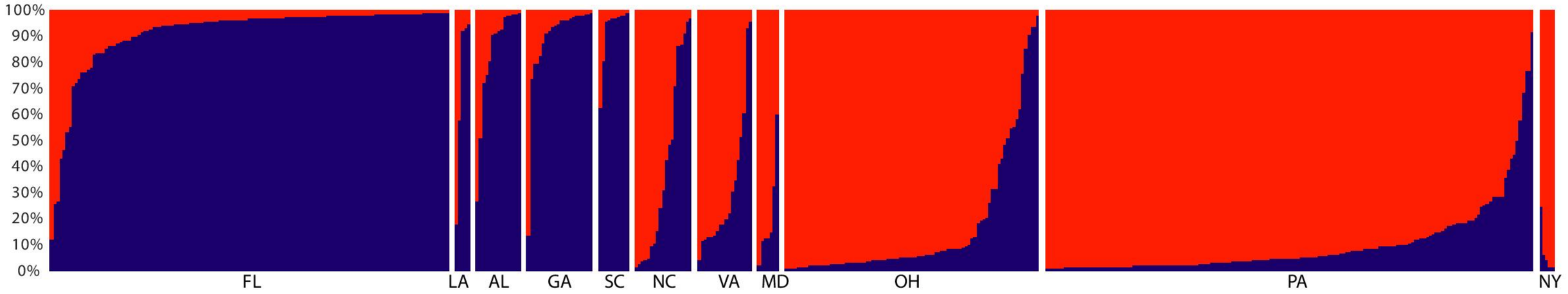
Figure 2. Genetic structure inferred by Bayesian clustering in `STRUCTURE` (A) and `TESS` (B) at K=2 with sampling locations indicated on the X-axis. (C) Average Q-values per state inferred via `STRUCTURE`.

Figure 3. Principal component analysis (PCA) of all 482 coyote samples using 10 microsatellite loci, with colors corresponding to northern and southern sampling locations. Labels for Ohio and Maryland are overlapping. The variation explained by PC1 and PC2 was 3.57% and 3.43%, respectively.

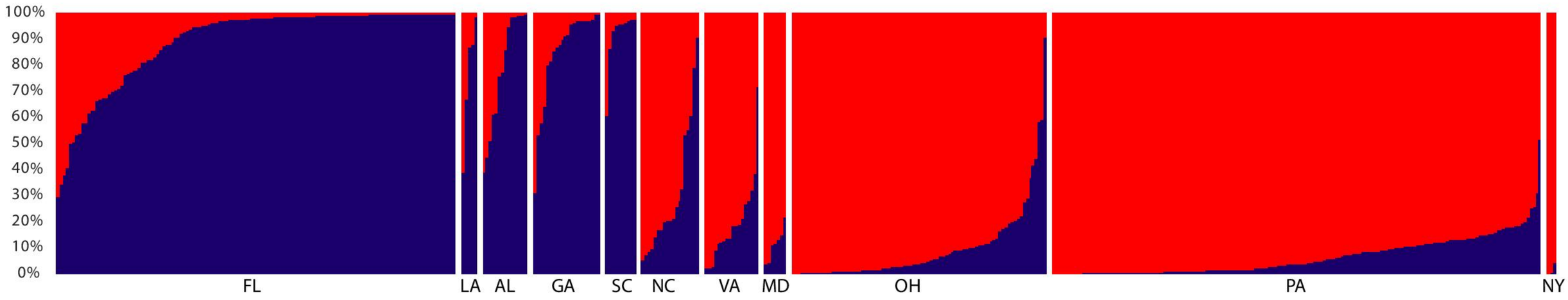
Figure 4. (A) Geographic cline in average Q_{North} frequency along a 2072 km north-south sampling transect connecting Collier County, Florida and Hamilton County, New York. Crosses represent sampled counties. (B) Approximate location of cline center along the sampling transect, highlighting sampling locations within the two log-likelihood support limits of the cline center.


NY
n=5
PA
n=163
n=85
OH
MD
VA
n=7
n=18
NC
n=19
SC
n=15
GA
n=10
AL
n=22
LA
n=5
FL
n=133
Historic Coyote Range
Early 1900s
Mid 1900s-Present
Expansion Route

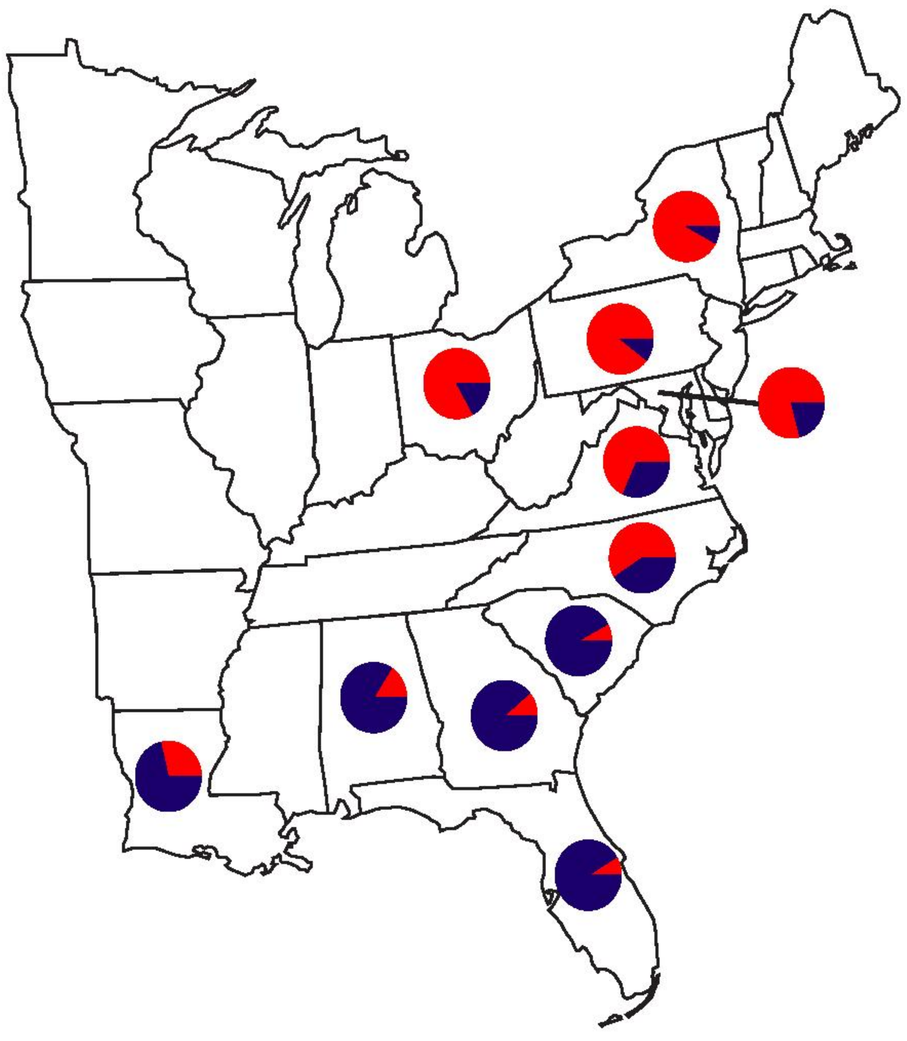
A

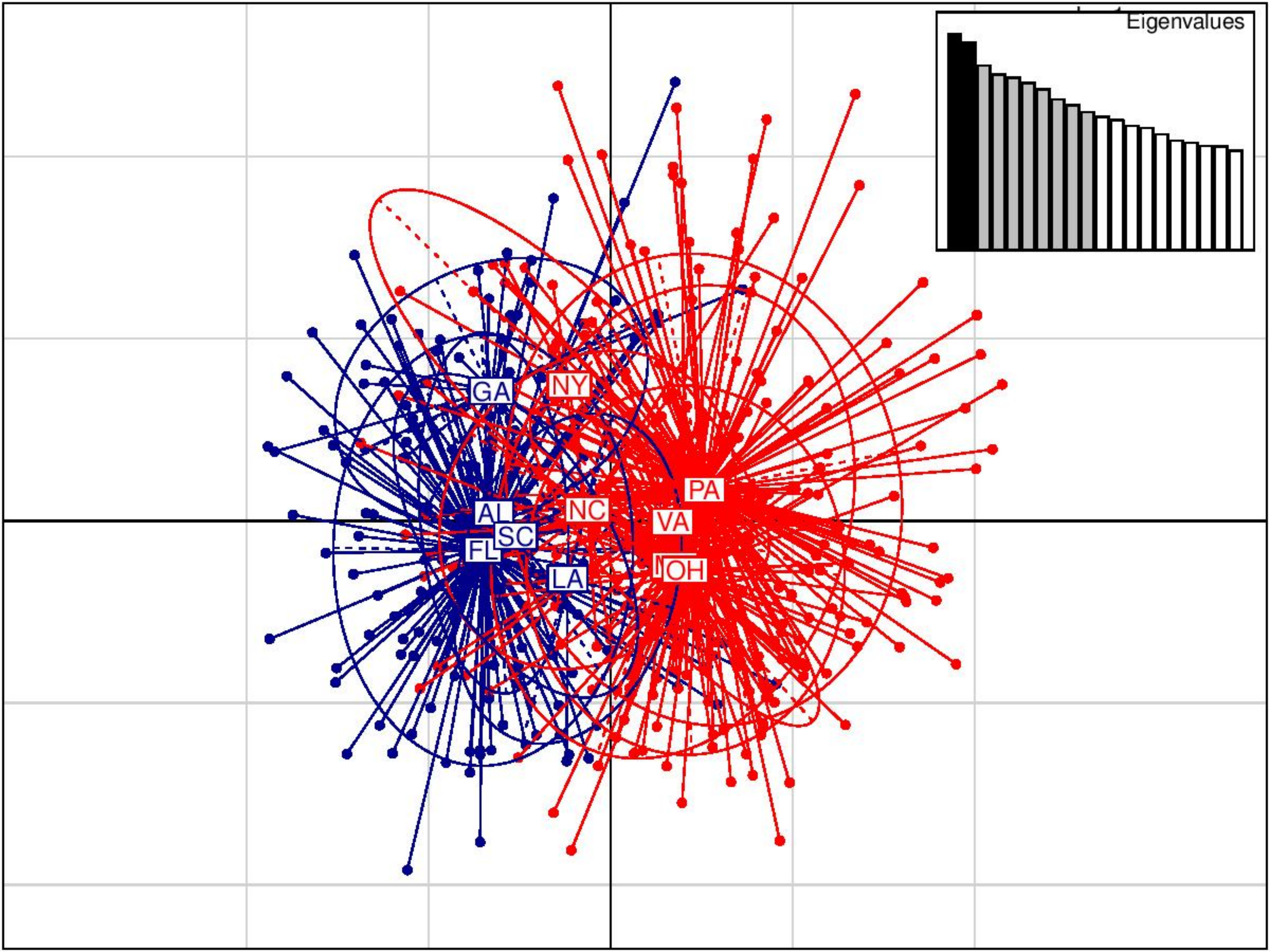

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
FL
LA
AL
GA
SC
NC
VA
MD
OH
PA
NY

B

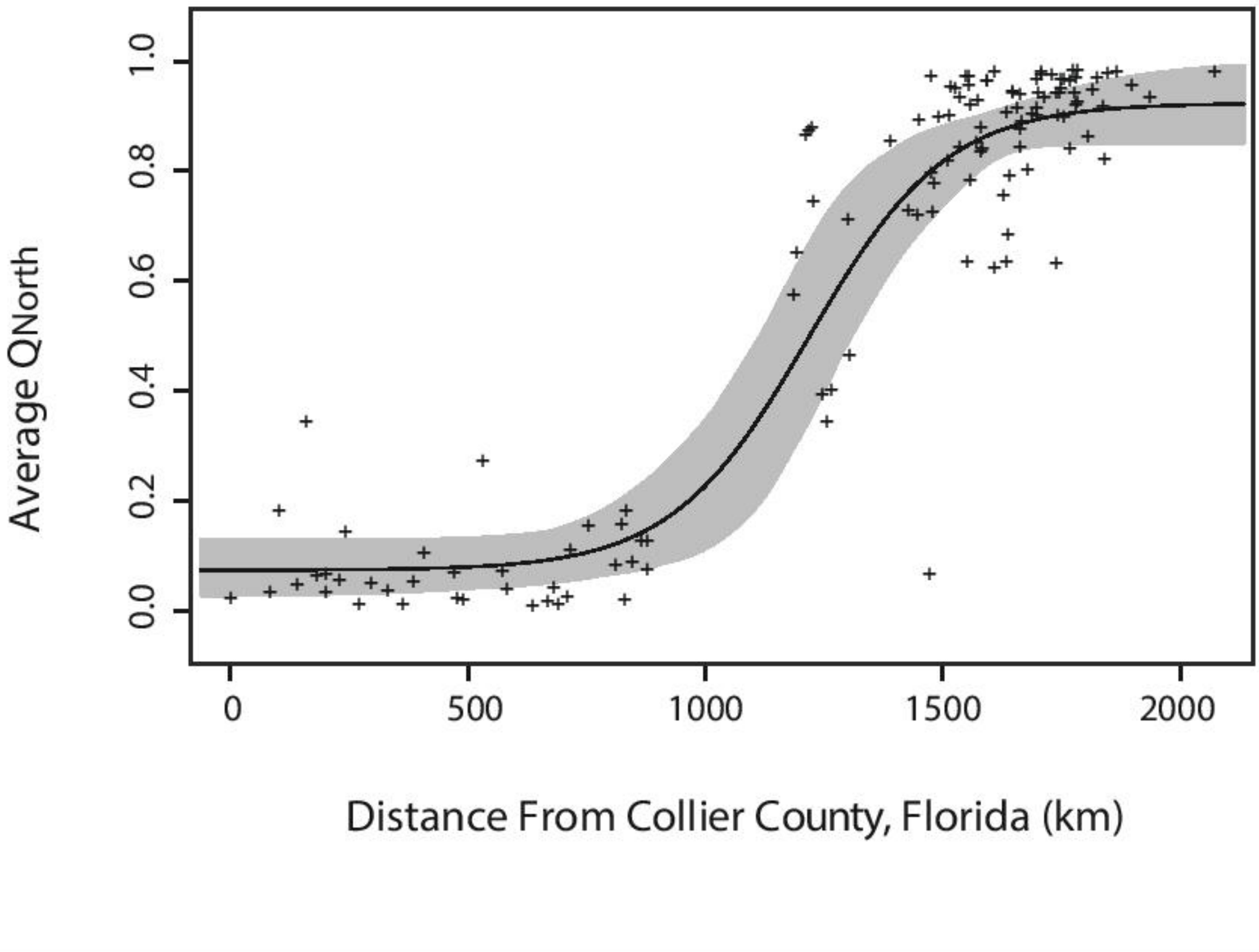

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
FL
LA
AL
GA
SC
NC
VA
MD
OH
PA
NY

C

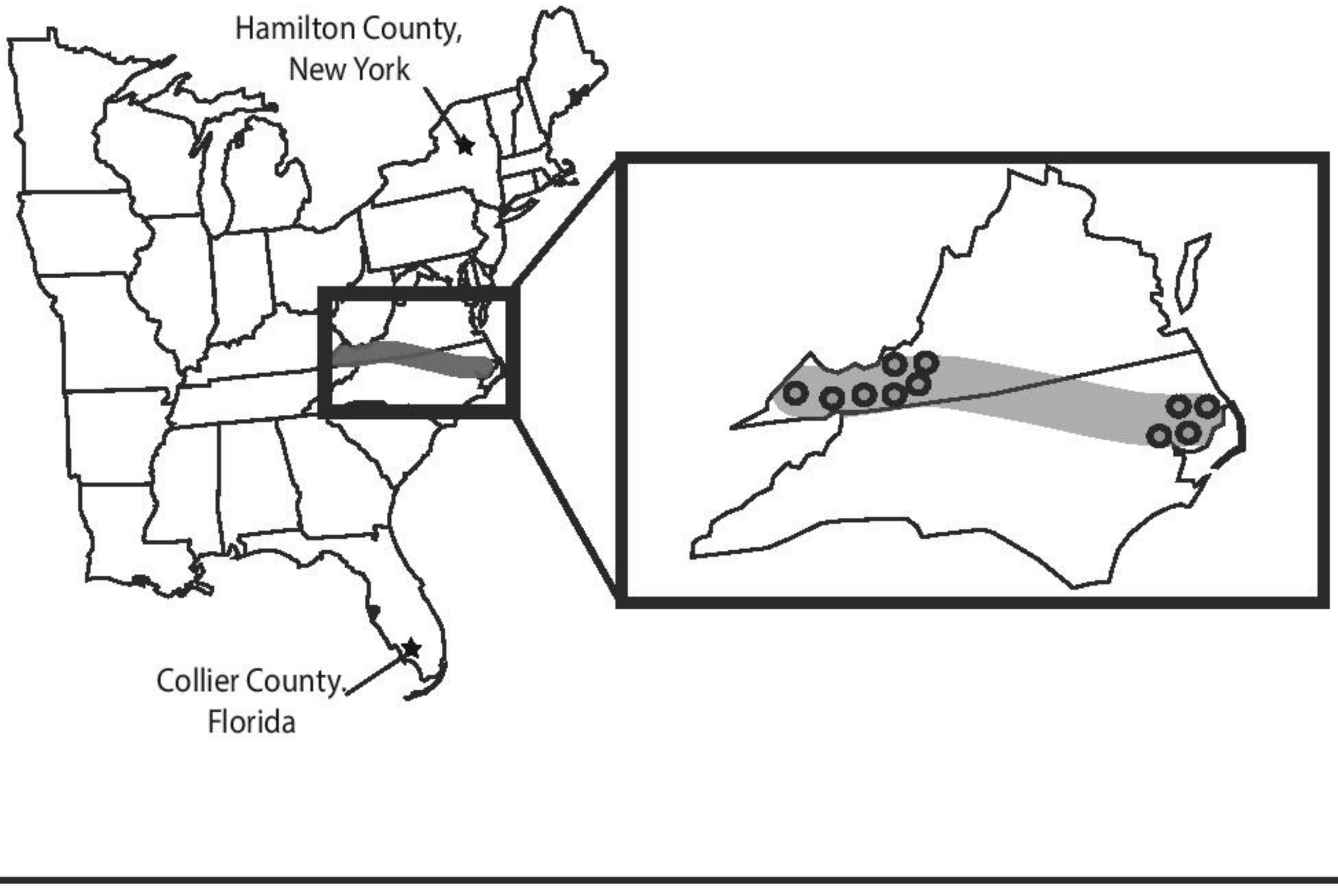



Eigenvalues
GA
NY
PA
AL
NC
VA
SC
FL
OH
LA

A


Average QNorth
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
2000
Distance From Collier County, Florida (km)

B


Hamilton County,
New York
Collier County,
Florida