



## RESEARCH ARTICLE

# First regional evaluation of nuclear genetic diversity and population structure in northeastern coyotes (*Canis latrans*) [v1; ref status: indexed, <http://f1000r.es/2y3>]

Javier Monzón

Departments of Ecology &amp; Evolution and Molecular Genetics &amp; Microbiology, Stony Brook University, Stony Brook, NY, 11794, USA

**v1** First published: 03 Mar 2014, 3:66 (doi: [10.12688/f1000research.3567.1](https://doi.org/10.12688/f1000research.3567.1))  
 Latest published: 03 Mar 2014, 3:66 (doi: [10.12688/f1000research.3567.1](https://doi.org/10.12688/f1000research.3567.1))

**Abstract**

Previous genetic studies of eastern coyotes (*Canis latrans*) are based on one of two strategies: sampling many individuals using one or very few molecular markers, or sampling very few individuals using many genomic markers. Thus, a regional analysis of genetic diversity and population structure in eastern coyotes using many samples and several molecular markers is lacking. I evaluated genetic diversity and population structure in 385 northeastern coyotes using 16 common single nucleotide polymorphisms (SNPs). A region-wide analysis of population structure revealed three primary genetic populations, but these do not correspond to the same three subdivisions inferred in a previous analysis of mitochondrial DNA sequences. More focused geographic analyses of population structure indicated that ample genetic structure occurs in coyotes from an intermediate contact zone where two range expansion fronts meet. These results demonstrate that genotyping several highly heterozygous SNPs in a large, geographically dense sample is an effective way to detect cryptic population genetic structure. The importance of SNPs in studies of population and wildlife genomics is rapidly increasing; this study adds to the growing body of recent literature that demonstrates the utility of SNPs ascertained from a model organism for evolutionary inference in closely related species.

**Article Status Summary****Referee Responses**

Referees	1	2	3
v1 published 03 Mar 2014	 report	 report	 report

- Astrid Stronen**, Aalborg University  
Denmark
- Linda Rutledge**, Trent University Canada
- Mark Weckel**, American Museum of  
Natural History USA

**Latest Comments**

No Comments Yet

**Corresponding author:** Javier Monzón ([javier.monzon@stonybrook.edu](mailto:javier.monzon@stonybrook.edu))

**How to cite this article:** Monzón J. First regional evaluation of nuclear genetic diversity and population structure in northeastern coyotes (*Canis latrans*) [v1; ref status: indexed, <http://f1000r.es/2y3>] F1000Research 2014, 3:66 (doi: [10.12688/f1000research.3567.1](https://doi.org/10.12688/f1000research.3567.1))

**Copyright:** © 2014 Monzón J. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

**Grant information:** Research reported in this publication was supported by a Stony Brook University Turner Fellowship research grant to Javier Monzón, and by the National Institute of General Medical Science of the National Institutes of Health IRACDA grant K12GM102778 to Jorge Benach.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Competing interests:** No competing interests were disclosed.

**First published:** 03 Mar 2014, 3:66 (doi: [10.12688/f1000research.3567.1](https://doi.org/10.12688/f1000research.3567.1))

**First indexed:** 01 May 2014, 3:66 (doi: [10.12688/f1000research.3567.1](https://doi.org/10.12688/f1000research.3567.1))

## Introduction

Historically restricted to the open deserts and plains of central and western North America, the coyote (*Canis latrans*) has colonized almost the whole continent in the last 100 years, with few exceptions (e.g. Long Island)<sup>1,2</sup>. The eastward range expansion was likely facilitated by widespread deforestation associated with agricultural development in the early 20<sup>th</sup> century and by the near extirpation of eastern wolves (*Canis lupus lycaon* or *Canis lycaon*) and red wolves (*Canis rufus*)<sup>3-5</sup>. The range expansion into northeastern North America advanced as two primary colonization fronts: the northern route through Ontario, where coyotes hybridized with resident eastern wolves, and the southern route through Ohio, where wolves were eradicated prior to coyote expansion<sup>2,5,6</sup>. Along with its geographic range expansion, the coyote also experienced a niche expansion by rapidly colonizing whole new biomes, including eastern temperate and boreal forests<sup>1</sup>. This complex scenario of colonization provides an interesting opportunity to explore the swift formation of population genetic structure following a rapid expansion in geographic distribution and ecological niche.

Recent analyses of population structure in northeastern coyotes have described a general lack of genetic differentiation among sampling localities, except at the coarsest geographic scales. Way *et al.*<sup>7</sup> examined genetic variation and structure in a sample of coyotes from eastern Massachusetts using mitochondrial DNA (mtDNA) and eight microsatellite loci. They found no genetic structure in coyotes within Massachusetts or even within the broader region of northeastern North America. Instead, northeastern coyotes seemed to constitute one uniform population slightly differentiated from western coyotes. In another analysis of genetic variation in northeastern coyotes, Kays *et al.*<sup>5</sup> identified three coarse phylogeographic areas: Ohio, the northeast zone, and a contact zone in western Pennsylvania and New York where the colonization front from Ohio has spread into the northeastern population. Although Kays *et al.* surveyed genetic variation in a dense geographic sample of 687 coyotes, they only used one genetic marker, the hypervariable mtDNA control region. vonHoldt *et al.*<sup>8</sup> conducted a genome-wide analysis of North American *Canis* species and detected population structure in *C. latrans*, but only at the broadest continental scale. Although vonHoldt *et al.* genotyped tens of thousands of loci, they only sampled 14 northeastern coyotes, making detection of finer levels of population structure in the region very improbable. Thus, all the previous studies of population structure in northeastern coyotes have adopted one of two strategies: sampling many individuals using one (mtDNA) or very few molecular markers, or sampling very few individuals using thousands of genomic markers. A regional analysis of genetic diversity and population structure in northeastern coyotes using many samples and many nuclear molecular markers is currently lacking.

Genetic structure is a ubiquitous property of natural, domesticated, and human populations. Population genetic structure plays considerable roles in evolution, as both the basis and the consequence of local adaptation, the splitting of one species into two if the environments are markedly different, and the adaptability of a species as a whole across its range (i.e., transformation rather than speciation)<sup>9</sup>. The detection of genetic structure largely depends on the

type and number of molecular markers examined, their variability in the target population, the number of individuals sampled, and the spatial sampling scheme<sup>10,11</sup>. Single nucleotide polymorphisms (SNPs) have become a popular and inexpensive tool in the field of molecular population genetics. SNPs have properties that make them a superior alternative to other widely used genetic markers, such as microsatellites and mtDNA sequences, in evaluating genetic diversity and population structure<sup>12-14</sup>. Furthermore, the sampling scheme can greatly influence the location and composition of genetic clusters, especially in species that are continuously distributed across a landscape<sup>11</sup>, as are northeastern coyotes. Thus the conclusion of no population structure from previous analyses based on limited individual or genomic sampling may be imprecise.

The objective of this study was to test the hypothesis that fine-scale population structure in northeastern coyotes exists, but remains undetected due to the small number of individuals or the low resolution of the genetic markers previously analyzed. I hypothesized that population structure would be detectable, at finer levels than in previous analyses, by using an array of 16 high-heterozygosity nuclear SNPs ascertained from the dog genome and a spatially dense sample of 385 coyotes. This is the first comprehensive regional survey of genetic diversity and population structure in northeastern coyotes that uses a dense geographic sampling scheme and several SNPs. This regional analysis reveals a cryptic population structure and a geographic pattern of nuclear genetic diversity that is discrepant with previous mtDNA- and microsatellite-based surveys. More generally, this study adds to the growing body of recent literature that demonstrates the utility of SNPs discovered in a model organism for evolutionary inference in wild relatives, as long as ascertainment bias is explicitly evaluated.

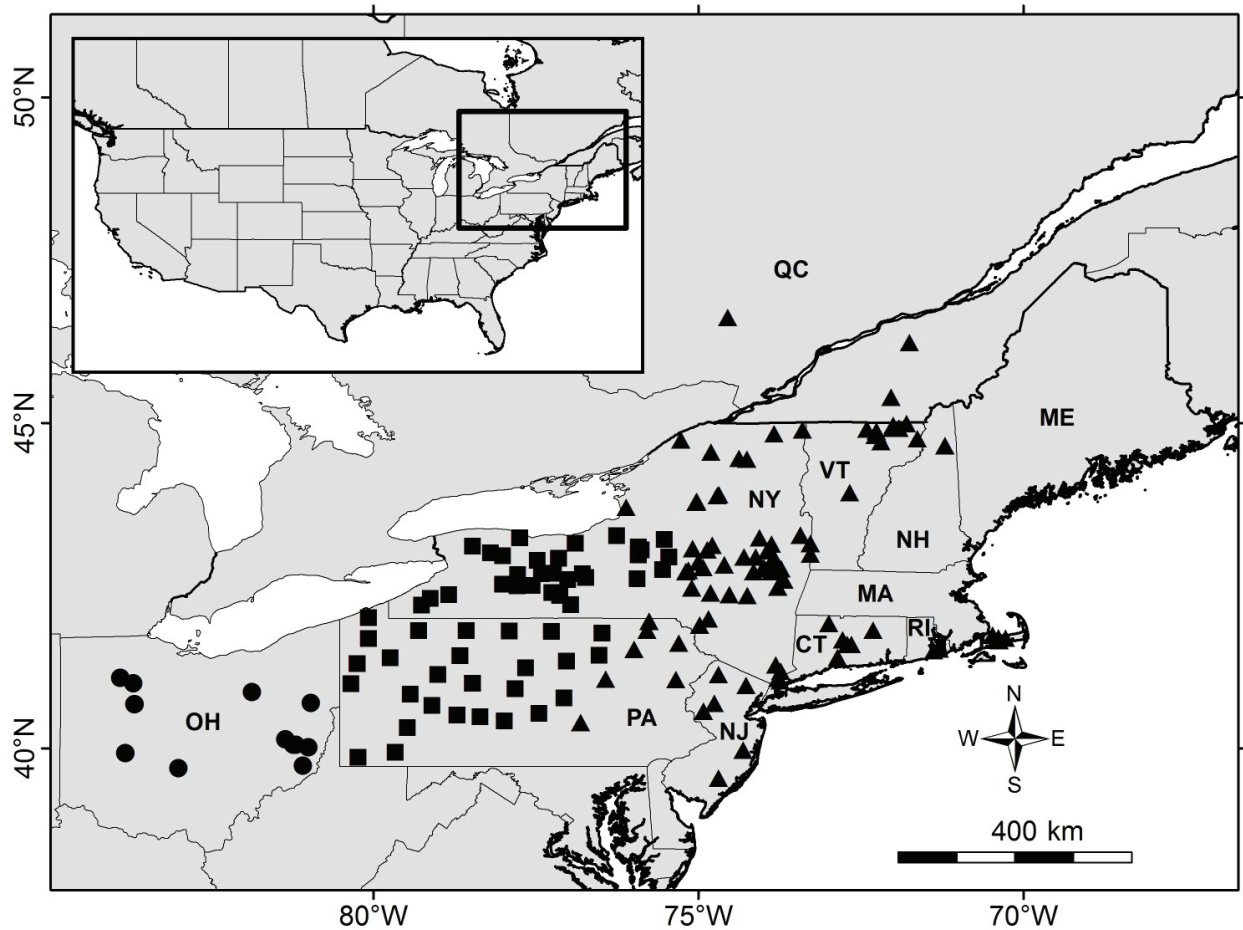
## Materials and methods

### Study area and sampling

The study area was located in northeastern North America. Coyotes were sampled from New York (N = 174), Pennsylvania (N = 103), Vermont (N = 34), Ohio (N = 30), New Jersey (N = 14), New Hampshire (N = 11), Connecticut (N = 8), Massachusetts (N = 5), southern Quebec (N = 4), and Rhode Island (N = 2) (Figure 1). All samples (Total N = 385) used in this study are archived and vouchered in the New York State Museum, Albany, NY, where they were kept at -80°C. Specimens (tissues or combination of skin, skull, skeleton) were obtained primarily through donations from licensed local hunters and trappers since 1999. Six samples came from previous scat surveys in New York<sup>15,16</sup> (Data File). No Institutional Animal Care and Use Committee (IACUC) review was required for this study because the DNA samples came from scat or animals killed for reasons other than research.

### Marker selection and laboratory methods

I selected molecular markers based on a genomics study that used the Affymetrix Canine Mapping Array to genotype 61,435 SNP loci in hundreds of wild and domestic canids, including 14 northeastern and 3 Ohio coyotes<sup>8,17</sup>. I used the program PLINK<sup>18</sup> to compute observed and expected heterozygosity per locus in the subset of 17 northeastern and Ohio coyotes. I selected 16 unlinked SNPs, each on a different autosomal chromosome, with the highest



**Figure 1. Study area and sampling localities of 385 coyotes from northeastern United States and southeastern Canada.** Symbol shapes represent phylogeographic zones as in Kays *et al.*<sup>5</sup>: circle, Ohio; square, contact zone; triangle, northeast zone.

observed heterozygosity, a measure of genetic variability in a population (Table 1). Ascertainment schemes that select highly variable SNP loci increase power to detect population structure<sup>13</sup>; accordingly, these 16 informative SNPs with high heterozygosity and high minor allele frequencies allowed me to assess genetic variation and population structure in the larger target sample. I designed primers using the Primer3 software<sup>19</sup> and tested them in silico against the dog CanFam2 genome assembly (GenBank Assembly ID: GCA\_000002285.1)<sup>20</sup> using the University of California, Santa Cruz In-Silico PCR and BLAT web tools (<http://genome.ucsc.edu>).

I extracted total genomic DNA from muscle or hide using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, USA) according to the manufacturer's instructions, and also used DNA samples that were extracted in the New York State Museum as described in Kays *et al.*<sup>5</sup>. I quantified DNA concentrations using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, USA) and diluted the samples in water to attain concentrations of 5–30 ng/ $\mu$ l. I prepared four 96-well plates with template genomic DNA of 378 eastern coyotes, seven of which served as controls because they were already genotyped with the canine SNP microarray<sup>8</sup>. Two

western coyotes from Washington and two western gray wolves (*Canis lupus*) from Yellowstone National Park served as additional controls; two wells containing only water served as negative controls.

SNP genotyping was performed at the GenoSeq Core laboratory in the University of California, Los Angeles, using a high resolution melting curve quantitative PCR assay on a LightCycler 480 thermal cycler (Roche, Indianapolis, USA). Each of the 16 SNPs was amplified and genotyped separately. DNA was amplified in a total volume of 10  $\mu$ l, including 1  $\mu$ l (5–30 ng) of genomic DNA, 0.2  $\mu$ M of each primer, 4.2 mM MgCl<sub>2</sub>, and 1 $\times$  Roche High Resolution Melting Master kit mix. The latter contains FastStart Taq DNA polymerase, dNTP mix, and ResoLight, a high resolution melting dye that fluoresces when DNA is double-stranded. During the melt curve analysis, the temperature increases very slowly to denature double-stranded DNA. Samples with variations in DNA sequence, even in one base pair, are distinguished by discrepancies in the shape of the melt curve, thus discriminating each of the two homozygous and the heterozygous genotypes. I processed the raw data using the Gene Scanning module of the Roche LightCycler

**Table 1. Sixteen single nucleotide polymorphisms (SNPs) and their primer pairs for high-resolution melt curve genotyping assay.** SNP ID chromosomal coordinates and dbSNP rs numbers correspond to the *CanFam2* dog genome assembly. Measures of genetic variability correspond to the initial seventeen northeastern and Ohio coyotes (*Canis latrans*) that formed the ascertainment panel.

SNP ID	dbSNP rs#	Forward primer	Reverse primer	PCR product size (bp)	MAF	H <sub>o</sub>	H <sub>e</sub>
chr1:92426160	rs21988674	GGGGTTTCTGAAGTGCTGAC	TGTGATAGCCACAGAAAAGCA	92	0.382	0.765	0.472
chr3:60134962	rs8795212	CACTGAGGAATGCTGGGAAG	TCAGGAAGTCTACTCCAGTGTCTG	80	0.441	0.765	0.493
chr4:33800600	rs24071674	ATCACCTCCAGAAAGCCAAG	TAAGGATCATCCCCTCGTTC	95	0.382	0.765	0.472
chr5:65740765	rs24241051	GGACCTCCATAGGACATCCA	TGTGTGGGGAGATGCAAAT	97	0.500	0.765	0.500
chr6:17110138	rs24352476	CAGTCACAATGGGGTGTCAG	AAGCGGGAGGTAGTATTACTGGT	97	0.441	0.765	0.493
chr8:69650155	rs24514604	GCTCCTGGCTATTGTATTTCC	TTCATTCTGCATGGTTGGT	99	0.441	0.765	0.493
chr10:45343436	rs22055760	TCTTTGAGGACATGGAACGA	TCACTCTGGAGACCAAGACG	94	0.500	0.765	0.500
chr11:66863044	rs8946304	TGGGTAATTTAATCAACGAGGAA	AAAAGCAAGAGGAGGGAACC	92	0.441	0.765	0.493
chr12:17166054	rs22184574	CAACGGCTGGATTCTGACTA	GCACACTGGTGTAGCAGAGC	118	0.441	0.765	0.493
chr16:9533917	rs22444520	TTGATAAATCAAAACCTGGGATG	GATCTGGCCCACAGCTCA	96	0.441	0.765	0.493
chr17:31508687	rs22603056	CAAAAATCAGGGATACAGACAAG	GCCAGAGAATGCCATCTTTA	100	0.471	0.824	0.498
chr19:50618604	rs22758397	TTTTTCCCTGCCTGATTTTT	TTGGAAAGAGATGTCAAGATGG	92	0.441	0.765	0.493
chr22:57259397	rs23051971	GTAGAGGACACCCTTAGATGTGG	TGCTGGAGGGAGTTCAACA	95	0.500	0.765	0.500
chr25:44793770	rs23209441	TGACTACCCAAGGTGATATG	CAGCTCTGATCATGCCAAAT	100	0.471	0.824	0.498
chr27:5811313	rs23365246	AATCACACACGAGCAACACC	CTGCTTGTCTGGGATGAA	96	0.471	0.824	0.498
chr37:26421162	rs9205317	GGCTCCCAGCTAACTGTTC	AGCTATCCAGAAGCCCAAGAG	93	0.471	0.824	0.498

**Note:** MAF: minor allele frequency; H<sub>o</sub> and H<sub>e</sub>: observed and expected heterozygosity. Several SNPs have the same MAF, H<sub>o</sub>, and H<sub>e</sub> due to the relatively small number of coyotes in the ascertainment panel. For example, H<sub>o</sub> = 0.765 if 13 of 17 individuals were heterozygous.

software and followed Roche's recommendations for evaluating data quality. The software automatically generates genotypes from the raw melt curve data. All SNPs were biallelic, so there are three possible genotypes per locus—AA, AB and BB—each forming three distinct clusters when signal intensity is plotted against temperature. I visually inspected all software-generated genotype calls and manually removed the ambiguous ones that did not conform to any of the three possible genotype clusters,

#### Analyses of genetic diversity and population structure

To the 378 samples genotyped in this study, I added 7 more northeastern coyote samples genotyped by vonHoldt *et al.*<sup>8</sup> making a total sample size of 385. I used PLINK to calculate average observed and expected heterozygosity, measures of genetic diversity, and deviations from Hardy-Weinberg equilibrium in the overall sample and in each of the three phylogeographic zones inferred by Kays *et al.*<sup>5</sup> (Figure 1). In order to assess the ascertainment bias of using dog genome SNPs to study genetic diversity and population structure in coyotes, I compared genetic variation in five subsets of canids: northeastern coyotes, western coyotes, western gray wolves, Great Lakes wolves, and dogs (*Canis familiaris*)<sup>8,17</sup>. I calculated average observed and expected heterozygosity using all 61,435 SNPs from the microarray because the vast majority of the loci in the microarray were ascertained by dog-dog comparisons<sup>17</sup>, and again using only the 16 high-heterozygosity SNPs genotyped in this study.

To assess population genetic structure, I excluded all individuals with more than five missing genotypes, resulting in 247 coyotes (16 in Ohio, 118 in contact zone, 113 in northeast zone). I used the program STRUCTURE 2.3<sup>21</sup> to infer the most likely number of genetic populations. STRUCTURE implements a Bayesian algorithm to assign multilocus genotypes to genetic clusters by calculating the likelihood that a group of individuals constitutes a population. I tested whether finer population structure was detectable with the high-heterozygosity SNPs relative to the coarse structure detected with hypervariable mtDNA. I analyzed all 247 coyotes together to test whether more than three populations were detectable at the regional level, and then analyzed each zone separately to test whether more than one population was detectable within each zone. For all analyses, I used three replicate runs of 20,000 burn-in and 100,000 Markov chain Monte Carlo iterations, used the admixture ancestry model with correlated allele frequencies<sup>22</sup>, and set the number of populations from K = 1 to K = 8. I verified that alpha and likelihood statistics reached convergence during the burn-in period for each number of populations analyzed. I used STRUCTURE HARVESTER<sup>23</sup> to evaluate the relative support for each value of K by plotting Ln P(D), the mean posterior probability of the data<sup>21</sup>, and ΔK, a quantity related to the second-order rate of change of the likelihood function with respect to K<sup>10</sup>. I used CLUMPP<sup>24</sup> to align and average the three replicate cluster membership coefficient matrices, and ArcMap 10 (Esri, Redlands, USA) to visualize



the spatial distribution of genetic structure. I considered individuals with ancestry coefficients  $q > 0.8$  as belonging to a specific cluster and individuals with all  $q < 0.8$  as being admixed, consistent with recent genetic investigations of *Canis*<sup>25,26</sup>. In order to corroborate inferences from the STRUCTURE analysis with a model-free approach, I conducted a principal components analysis (PCA) using the adegenet 1.3–4 package<sup>27</sup> in R<sup>28</sup>. I also computed pairwise  $F_{ST}$ , the inbreeding coefficient within populations relative to the total, among the populations inferred by STRUCTURE and tested the significance of the differentiation by analysis of molecular variance (AMOVA) using 999 permutations in GenAlEx 6.5<sup>29,30</sup>.

**Results**

I interrogated 16 SNP loci in 378 coyotes for a total of 6,048 expected genotypes (Data File). Genotyping efficacy varied by source of DNA, with fecal samples amplifying less effectively than tissue samples (Mann-Whitney test:  $U = 1771$ ,  $P = 0.017$ ). That is, the fecal samples had, on average, more missing or ambiguous genotypes than the tissue samples.

The overall sample of 385 eastern coyotes had lower genetic diversity than expected (Table 2), even though observed heterozygosity generally exceeded expected heterozygosity in the ascertainment panel of 17 northeastern and Ohio coyotes (Table 1). Ohio coyotes were the most genetically diverse in the region, but the eastward decay in genetic diversity observed with mtDNA was not replicated with nuclear SNPs. Although coyotes from the northeast zone had

the lowest levels of mitochondrial genetic diversity, these same individuals had a level of nuclear genetic diversity comparable to Ohio coyotes (Table 2). The most pronounced differences between observed and expected heterozygosity occurred in the contact zone and in the overall regional analysis (Table 2). In the overall sample of 385 eastern coyotes, five loci were in Hardy-Weinberg equilibrium. The number of loci in Hardy-Weinberg equilibrium increased when each phylogeographic zone was analyzed separately: 13 in Ohio, nine in the contact zone, and nine in the northeast zone (Table 2).

Dogs appeared to be the most genetically diverse when the diversity of the five different canid groups was estimated using all 61,435 SNPs from the canine microarray. The genome-wide ascertainment bias was towards dogs: the expected heterozygosity of dogs was almost twice that of western coyotes. But the ascertainment bias reversed when heterozygosity was measured using only the 16 selected SNPs: coyotes appeared to be the most genetically diverse, with northeastern coyotes having a very high expected heterozygosity, whereas dogs appeared the least genetically diverse (Table 3).

In the region-wide population structure analysis of 247 individuals with little missing data, the values of  $K$  with the strongest statistical support were  $K = 2$  and  $K = 3$  (Figure 2A, B). Pairwise  $F_{ST}$  among the three genetic clusters varied from 0.08 to 0.10 and all were significant in the AMOVA framework ( $P = 0.001$ ). This indicates that there are three primary genetic subdivisions in the broad sampling area. The three groups did not correspond to the Ohio, contact, and

**Table 2. Genetic diversity of eastern coyotes (*Canis latrans*) measured with mtDNA sequences and 16 nuclear SNP genotypes.** Most individuals genotyped at 16 nuclear SNPs represent a subset of those individuals sequenced.

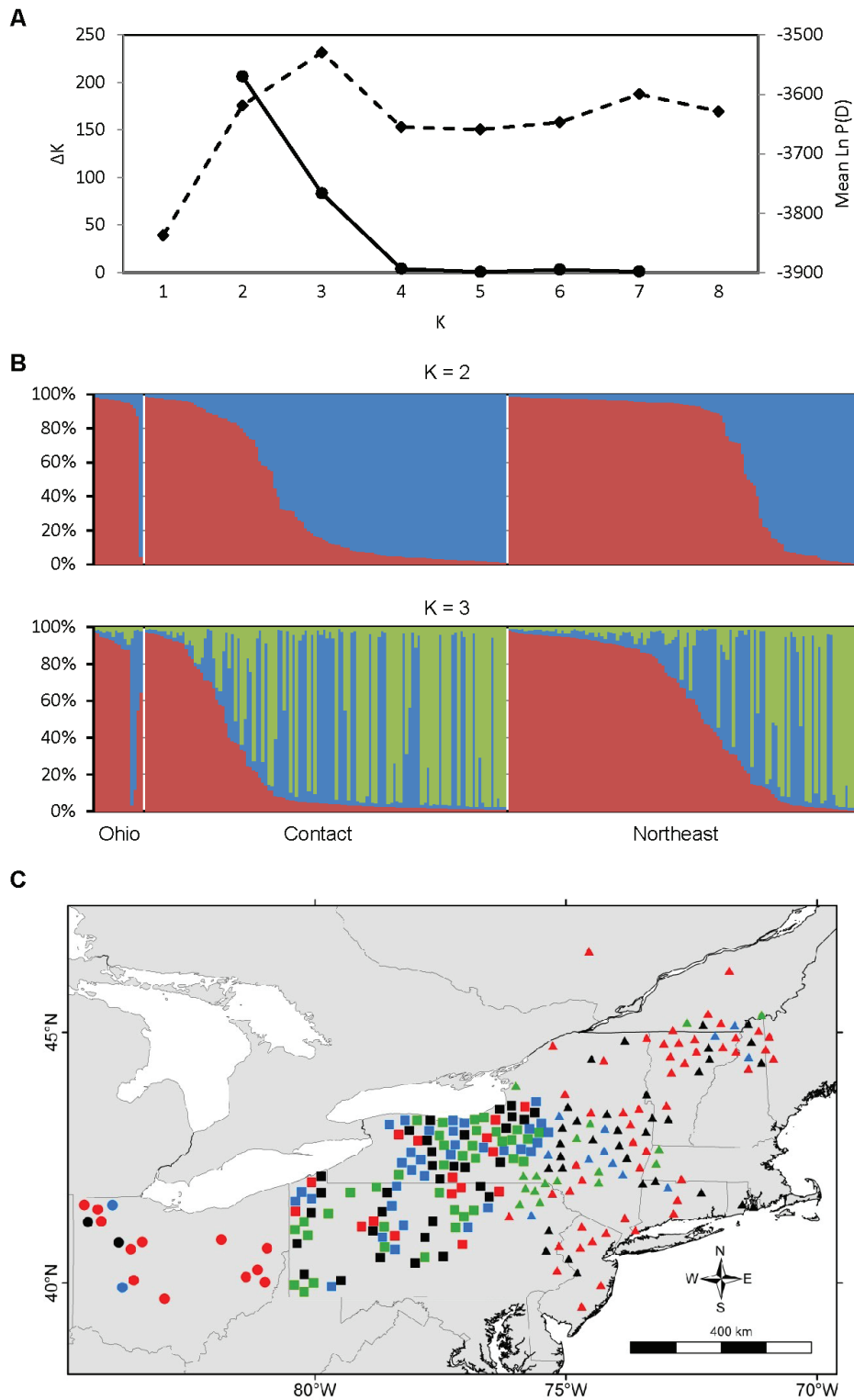
Zone	N	mtDNA control region		High-heterozygosity SNPs			
		Haplotype diversity	$\theta$ (per site)	N	$H_o$	$H_E$	HWE
Ohio	30	0.844	0.018	30	0.465	0.435	13
Contact	207	0.721	0.014	177	0.312	0.411	9
Northeast	450	0.664	0.008	178	0.442	0.457	9
Total	687	0.708	0.013	385	0.388	0.444	5

**Note:** mtDNA data and zone designations from Kays *et al.*<sup>5</sup>. N: sample size;  $H_o$  and  $H_E$ : observed and expected heterozygosity; HWE: number of loci in Hardy-Weinberg equilibrium.

**Table 3. Ascertainment bias of surveying genetic diversity in different groups of canids using SNPs discovered after completion of the dog genome assembly.**

Canid group	N	61,435 SNPs		16 high-heterozygosity SNPs	
		$H_o$	$H_E$	$H_o$	$H_E$
Northeastern coyotes	14	0.190	0.202	0.763	0.493
Western coyotes	45	0.147	0.182	0.387	0.399
Great Lakes wolves	19	0.187	0.217	0.278	0.290
Western gray wolves	32	0.203	0.238	0.271	0.319
Dogs	50	0.234	0.359	0.270	0.387

**Note:** Data for 61,435 SNPs from vonHoldt *et al.*<sup>8</sup>. N: sample size;  $H_o$  and  $H_E$ : observed and expected heterozygosity.



**Figure 2. Regional population genetic structure of eastern coyotes. A)** Estimation of the number of genetic populations using two criteria:  $\Delta K$  (solid line) and  $\text{Ln P(D)}$  (dashed line). The most probable number of populations is the value of  $K$  with the maximal value of  $\Delta K$  and  $\text{Ln P(D)}$ . The data strongly support  $K = 2$  and  $K = 3$ . **B)** STRUCTURE bar plots of  $N = 247$  coyotes subdivided into  $K = 2$  and  $K = 3$  genetic populations; each individual is represented by a vertical bar partitioned into two or three colored segments indicating that individual's proportional membership in each of two or three genetic clusters. **C)** Spatial distribution of  $K = 3$  genetic populations. Symbol colors represent the genetic cluster with  $> 80\%$  assignment from panel B, or black if highly admixed (*i.e.*, no assignment  $> 80\%$ ). Symbol shapes represent phylogeographic zones as in Figure 1. Symbol locations are slightly jittered to display each individual and to reduce clutter.

northeast zones previously inferred by mtDNA (Figure 2C). There is some geographic structuring, but the three groups overlap extensively in space. Although the red cluster in Figure 2B includes most of the Ohio coyotes, it is more cosmopolitan, also including many coyotes from the contact and northeast zones. The green cluster is mostly restricted to the contact zone, but extends slightly into eastern New York and Vermont. The PCA corroborated the results from STRUCTURE (Figure 3). The first two PCA axes explained 24.2% of the total variance and clearly separated the three STRUCTURE-inferred clusters. Most admixed individuals with no clear membership in any STRUCTURE-inferred cluster also showed no clear association with any PCA cluster.

Ample population genetic structure was detected in the contact zone; the values of  $K$  with the strongest statistical support were  $K = 3$  and  $K = 5$  (Figure 4A). However, the  $K = 5$  structure seems biologically unrealistic, characterized by highly admixed individuals of the various “populations” (Figure 4B). Alternatively, the  $K = 3$  structure identified for the contact zone was virtually identical to

the regional  $K = 3$  structure (Figure 4C). In contrast to the contact zone, no fine-scale genetic structure was detected in Ohio or in the northeast zone. In Ohio, the value of  $K$  with the highest explanatory power was  $K = 1$  (Figure 5A); in the northeast zone, the value of  $\ln P(D)$  does not increase beyond  $K = 2$  and the change in  $\ln P(D)$  between  $K = 1$  and  $K = 2$  is minimal, indicating weak support for genetic structure (Figure 5B).

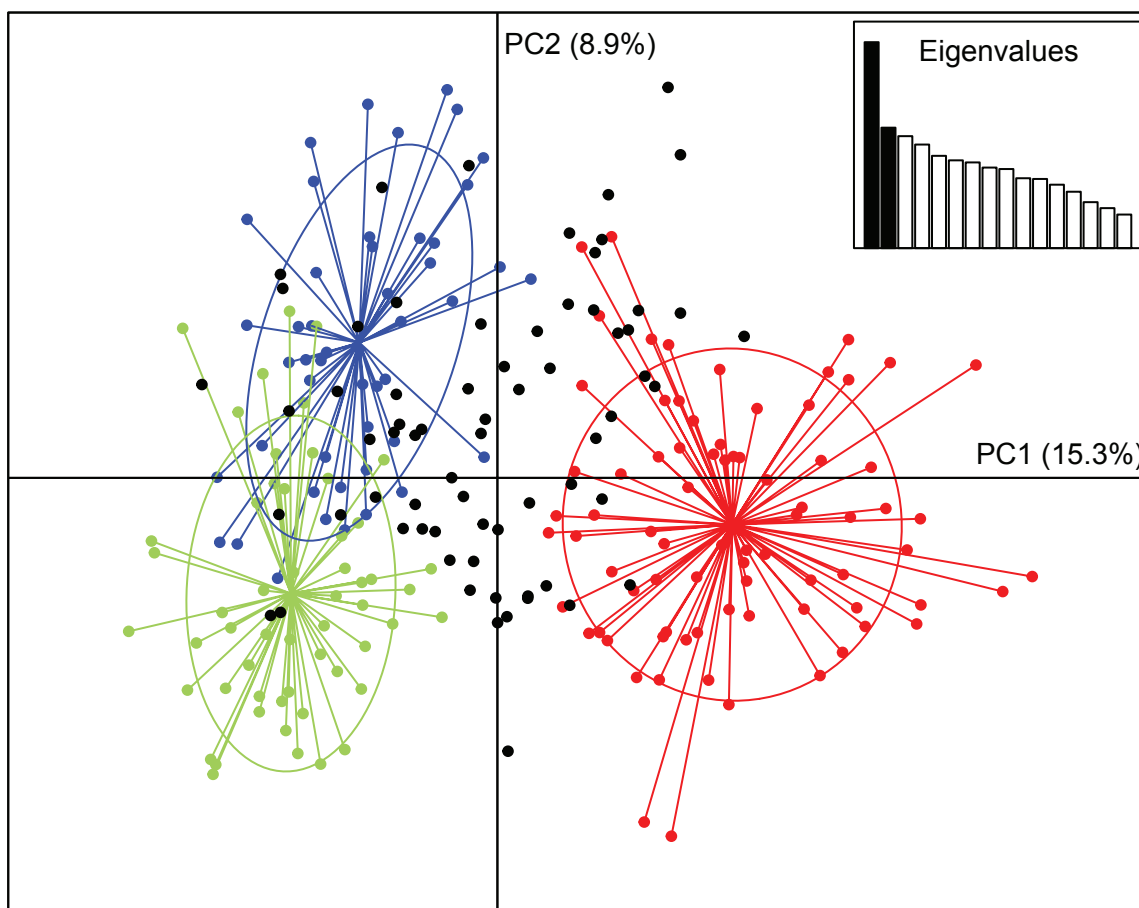
#### Data of genetic diversity in northeastern coyotes

2 Data Files

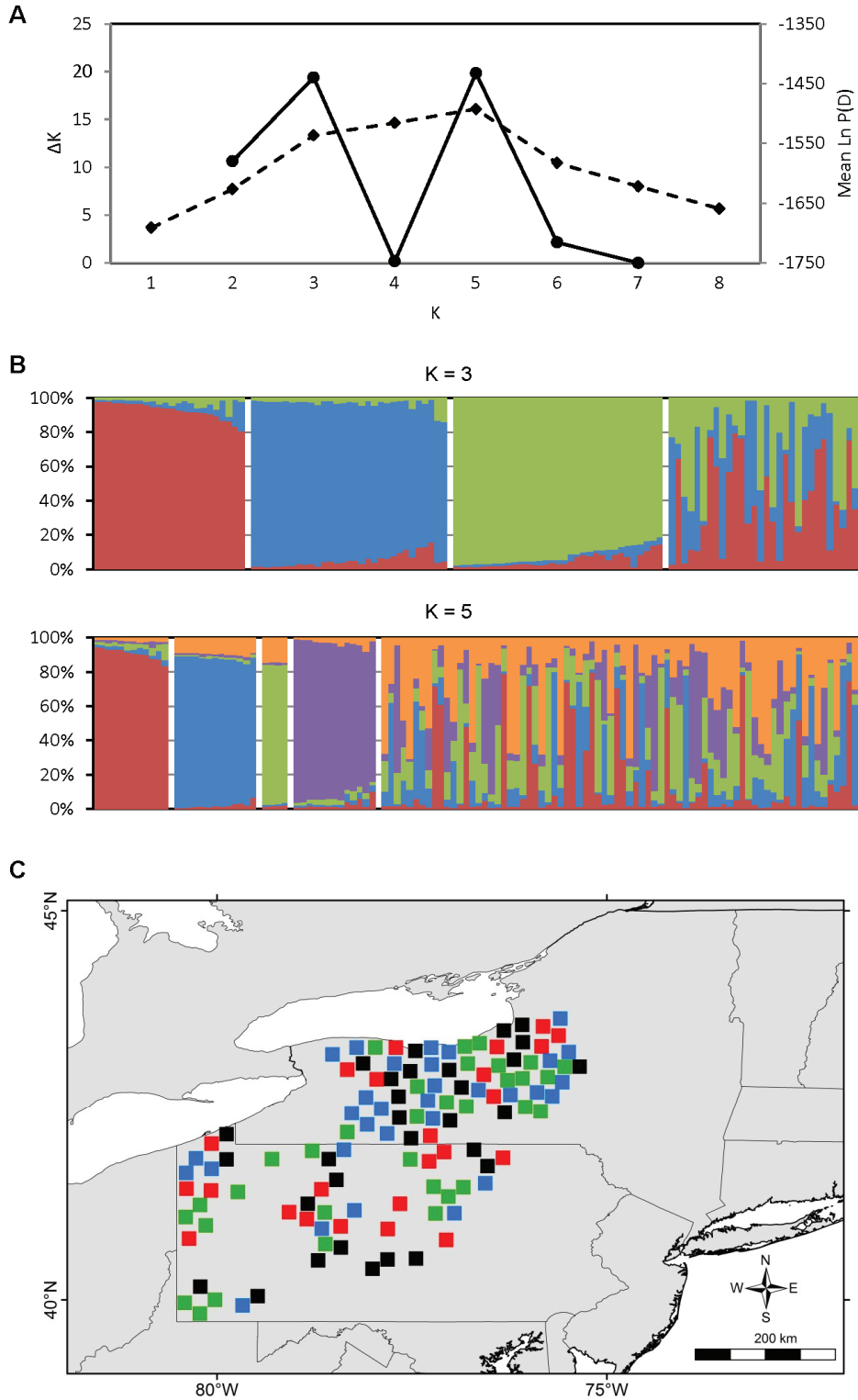
<http://dx.doi.org/10.6084/m9.figshare.943483>

#### Discussion

This study documents the presence of fine-scaled population genetic structure in eastern coyotes. Specifically, the contact zone exhibits a strong signal of population structure, the same signal detected in the regional analysis. This pattern may reflect the recent merging of

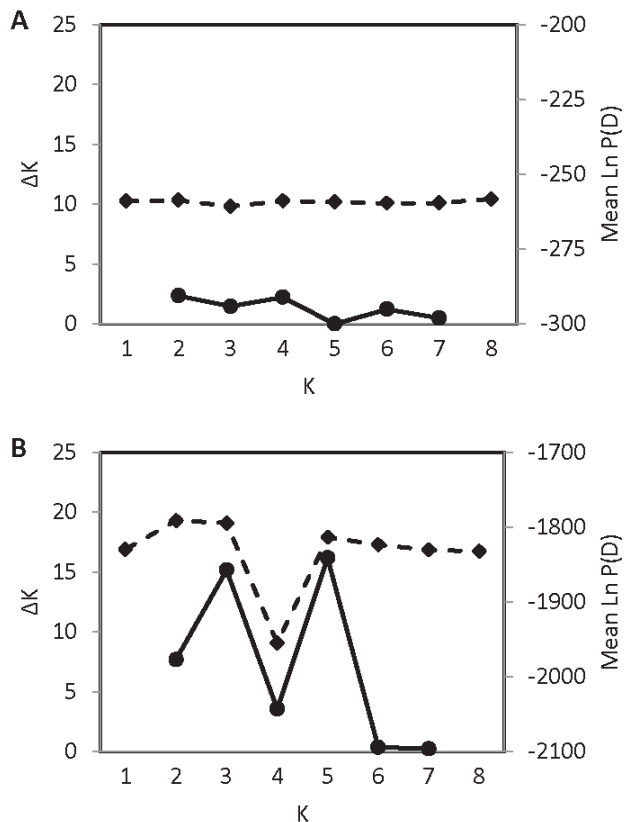


**Figure 3.** Principal components analysis of 16 autosomal SNPs genotyped in 247 eastern coyotes. Symbol colors correspond to  $K = 3$  STRUCTURE-inferred clusters with  $> 80\%$  assignment, or black if highly admixed (*i.e.*, no assignment  $> 80\%$ ), as in Figure 2C. Inset shows scree plot of eigenvalues.



**Figure 4. Population genetic structure of eastern coyotes in the contact zone. A)** Estimation of the number of genetic populations using two criteria:  $\Delta K$  (solid line) and  $\text{Ln P(D)}$  (dashed line). The data strongly support  $K = 3$  and  $K = 5$ . **B)** STRUCTURE bar plots of  $N = 118$  coyotes subdivided into  $K = 3$  and  $K = 5$  genetic clusters. **C)** Spatial distribution of  $K = 3$  genetic clusters. Symbol colors represent the genetic cluster with  $> 80\%$  assignment from panel **B**, or black if highly admixed (*i.e.*, no assignment  $> 80\%$ ).





**Figure 5.** Absence of fine-scale population genetic structure in **A)** N = 16 coyotes from Ohio, and **B)** N = 113 coyotes from the northeast zone. See [Figure 1](#) for zone designations.

two colonization fronts and the highly heterogeneous landscape of New York and Pennsylvania. A separate, spatially-explicit assessment with an independent set of SNPs demonstrated that northeastern coyotes exhibit a strong signal of population structure in Central New York<sup>1</sup>. Furthermore, the same investigation showed that population structure is partially explained by ecological factors, such as deer density and human land use<sup>1</sup>. Indeed, eastern coyotes living in areas of high deer density are genetically more wolf-like than those living in areas of low deer density<sup>31</sup>.

In contrast to the contact zone, coyotes in Ohio appear to make up a single panmictic population, as do coyotes in the northeast zone. Given the high mtDNA and nuclear diversity observed in Ohio by Kays *et al.*<sup>5</sup> and in this study, it is unlikely that the lack of genetic structure stems from a founder effect. The failure to uncover more than one genetic population in Ohio may be due instead to its landscape homogeneity or to the lack of resolution afforded by 16 individuals to detect finer levels of structure. On the other hand, the failure to uncover more than one genetic population in the northeast zone is surprising given the vast geographic area with its ecological heterogeneity and the large number of individuals sampled. Coyotes in the northeast zone are likely the descendants of a few founders and therefore do not exhibit a strong signal of population genetic structure.

The perception of population structure, even in highly vagile animals where it was least expected, has been refined by steadily

improving molecular data and geographic sampling. Initially, using mtDNA restriction site polymorphisms and nuclear microsatellites, no evidence of population structure or isolation by distance was found in coyotes, even at the continental scale<sup>32,33</sup>. Various behavioral and historical explanations have been invoked to explain these early genetic patterns. But a more likely explanation is that the patterns of weak differentiation were artifacts of sparse geographic sampling or poor resolution due to the use of few molecular markers. More recent studies employing advanced analyses of spatial and genetic data have revealed strong differentiation among parapatric populations of coyotes and wolves, even in the absence of physical barriers to movement<sup>1,34–38</sup>. These investigations used multiple loci and dense geographic sampling to uncover cryptic genetic subdivisions. Strong genetic differentiation between adjacent populations of coastal and inland wolves in British Columbia was shown with mtDNA<sup>39</sup>, demonstrating that fine-scale genetic differentiation can be detected with denser sampling alone, even using a single molecular mtDNA marker. Similar cryptic subdivisions have been discovered in several highly mobile groups, such as *Lynx*<sup>40,41</sup>, ungulates<sup>42–44</sup>, cetaceans<sup>45,46</sup>, and hawks<sup>47</sup>. In all these cases, genetic subdivisions appear to emerge from ecological factors and local foraging adaptations. Future studies should focus on the ecological mechanisms underlying the cryptic genetic structure in northeastern coyotes, especially because they have only inhabited the region for the last 30–80 years<sup>2</sup>. Confirming ecological determinants of population structure in the absence of obvious physical dispersal barriers would provide an interesting example of rapid ecological differentiation.

There are some important similarities and discrepancies between mtDNA and autosomal SNP patterns. The data indicate that coyotes in Ohio are the most genetically diverse in the region when surveyed with nuclear SNPs, as with mtDNA<sup>5</sup>. However, the gradual eastward decay in genetic diversity observed with mtDNA is not replicated with nuclear SNPs. The marked reduction of heterozygosity in the contact zone and in the overall region is very likely caused by population structure, *i.e.*, the Wahlund effect. This interpretation is supported by the increase in the number of loci in Hardy-Weinberg equilibrium at smaller geographic scales and by the congruent signal of population structure in the overall region and in the contact zone. In addition, the three primary populations detected in this study do not correspond to the three subdivisions inferred with mtDNA<sup>5</sup>. Together, these results suggest that studies based solely on mtDNA should be interpreted cautiously. For example, mtDNA sequence similarity suggested that a small population of Scandinavian wolves was founded by individuals released from Swedish zoos, but nuclear polymorphic markers falsified the release hypothesis and instead supported a hypothesis of natural immigration or expansion from an unknown relict wolf pack<sup>48</sup>. Discrepancies between patterns observed with mtDNA and nuclear DNA may be caused by true organismal processes, such as sex-biased dispersal<sup>49</sup>. However, there is no evidence for sex-specific dispersal behaviors in eastern coyotes, consistent with their monogamous breeding system<sup>50,51</sup>. Alternatively, discrepancies may be caused by marker-specific phenomena such as effective population size, lineage sorting, mutation rate, and coalescent times<sup>52</sup>, or the violation of certain assumptions of mtDNA inheritance, such as recombination, paternal leakage, and heteroplasmy<sup>53</sup>. Future studies should further evaluate these sources of discrepancies.

The present study underscores three related methodological issues that are of broad interest, especially as SNPs continue to be in vogue in population and wildlife genomics. First, these data confirm that SNPs discovered in a model organism are an appropriate tool to address various questions regarding the ecology and evolution of non-model relatives. The sequencing of the dog genome<sup>20</sup> quickly enabled SNP-based investigations into wild members of the family Canidae, including coyotes, wolves, jackals, and foxes<sup>8,54–57</sup>. Second, this study highlights the importance of evaluating the ascertainment bias of markers employed in a survey of genetic variation, especially in multi-species comparisons. Many SNP-based studies are not addressing the issue of ascertainment bias<sup>14</sup>. In the present study, the genome-wide analysis of variation is dog-biased because SNPs were ascertained primarily from comparisons of boxer and poodle genomes after the completion of the dog genome project<sup>8</sup>. On the other hand, the high-heterozygosity SNP analysis is coyote-biased because the 16 SNPs were chosen from an ascertainment panel of northeastern coyotes. The bias in the latter case is not problematic because the ascertainment panel from which the loci were selected is representative of the whole population of northeastern coyotes. However, the diversity measures reported in this study should not be compared to similar measures from other populations, unless the comparison corrects for ascertainment bias<sup>58</sup>. Lastly, this study emphasizes the necessity of selecting SNPs very carefully to match the research question of interest. Here, a set of high-heterozygosity SNPs was interrogated in order to examine geographic patterns of genetic diversity and population structure. But other research questions may require polymorphic markers with other properties. For example, in order to better understand the complex hybrid ancestry of the northeastern coyote, Monzón *et al.*<sup>31</sup> used species-diagnostic SNPs to quantify the relative contributions of its parental populations. In addition, recent advances in the molecular genetics of

phenotypic traits in dogs allow the use of SNPs linked to genes of known function to address long-standing questions about morphological, physiological, and behavioral adaptations in northeastern coyotes and other wild canids<sup>1,59</sup>.

### Data availability

*figshare*: Data of genetic diversity in northeastern coyotes. <http://dx.doi.org/10.6084/m9.figshare.943483><sup>60</sup>

### Competing interests

No competing interests were disclosed.

### Grant information

Research reported in this publication was supported by a Stony Brook University Turner Fellowship research grant to Javier Monzón, and by the National Institute of General Medical Science of the National Institutes of Health IRACDA grant K12GM102778 to Jorge Benach.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

### Acknowledgments

Robert Wayne, John Pollinger, and Bridgett vonHoldt provided technical support to initiate this project and shared the 61K-SNP dataset. John True generously provided access to lab resources. Roland Kays kindly provided access to the coyote collection of the New York State Museum. Shian-Ren Liou and Nashwa Khalil assisted with laboratory procedures. Daniel Dykhuizen, Catherine Graham and four anonymous reviewers provided constructive comments to earlier drafts of this manuscript.

## References

- Monzón J: **Rapid evolution of northeastern coyotes**. PhD thesis, Stony Brook University, 2012. [Reference Source](#)
- Parker G: **Eastern Coyote: the Story of its Success**. (Nimbus Publishing, 1995). [Reference Source](#)
- Gompper ME: **The ecology of northeast coyotes: current knowledge and priorities for future research**. (Wildlife Conservation Society, Bronx, New York, 2002). [Reference Source](#)
- Fener H, Ginsberg J, Sanderson E, *et al.*: **Chronology of range expansion of the coyote, *Canis latrans*, in New York**. *Can Field Nat.* 2005; **119**: 1–5. [Reference Source](#)
- Kays R, Curtis A, Kirchman J: **Rapid adaptive evolution of northeastern coyotes via hybridization with wolves**. *Biol Lett.* 2010; **6**(1): 89–93. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kays R, Curtis A, Kirchman J: **Reply to Wheeldon *et al.* ‘Colonization history and ancestry of northeastern coyotes’**. *Biol Lett.* 2010; **6**(2): 248–249. [Publisher Full Text](#)
- Way JG, Rutledge L, Wheeldon T, *et al.*: **Genetic characterization of eastern “coyotes” in eastern Massachusetts**. *Northeastern Nat.* 2010; **17**(2): 189–204. [Publisher Full Text](#)
- vonHoldt BM, Pollinger JP, Earl DA, *et al.*: **A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids**. *Genome Res.* 2011; **21**(8): 1294–1305. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Wright S: **The genetical structure of populations**. *Ann Hum Genet.* 1949; **15**(1): 323–354. [Publisher Full Text](#)
- Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study**. *Mol Ecol.* 2005; **14**(8): 2611–2620. [PubMed Abstract](#) | [Publisher Full Text](#)
- Schwartz MK, McKelvey KS: **Why sampling scheme matters: the effect of sampling scheme on landscape genetic results**. *Conserv Genet.* 2009; **10**(2): 441–452. [Publisher Full Text](#)
- Morin PA, Luikart G, Wayne RK: **SNPs in ecology, evolution and conservation**. *Trends Ecol Evol.* 2004; **19**(4): 208–216. [Publisher Full Text](#)
- Haasl RJ, Payseur BA: **Multi-locus inference of population structure: a comparison between single nucleotide polymorphisms and microsatellites**. *Heredity.* 2011; **106**(1): 158–171. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Helyar S, Hemmer-Hansen J, Bekkevold D, *et al.*: **Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges**. *Mol Ecol Resour.* 2011; **11**(suppl 1): 123–136. [PubMed Abstract](#) | [Publisher Full Text](#)
- Kays RW, Gompper ME, Ray JC: **Landscape ecology of eastern coyotes based on large-scale estimates of abundance**. *Ecol Appl.* 2008; **18**(4): 1014–1027. [PubMed Abstract](#) | [Publisher Full Text](#)
- Gompper ME, Kays RW, Ray JC, *et al.*: **A comparison of noninvasive techniques to survey carnivore communities in northeastern North America**. *Wildlife Soc*

- Bull. 2006; **34**(4): 1142–1151.  
[Publisher Full Text](#)
17. vonHoldt BM, Pollinger JP, Lohmueller KE, *et al.*: **Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication.** *Nature*. 2010; **464**(7290): 898–902.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  18. Purcell S, Neale B, Todd-Brown K, *et al.*: **PLINK: a tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet*. 2007; **81**(3): 559–575.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  19. Rozen S, Skaletsky H: **Primer3 on the WWW for general users and for biologist programmers.** In *Bioinformatics methods and protocols: methods in molecular biology* (eds Krawetz, S. & Misener, S.), Humana Press. 2000; **132**: 365–386.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  20. Lindblad-Toh K, Wade CM, Mikkelsen TS, *et al.*: **Genome sequence, comparative analysis and haplotype structure of the domestic dog.** *Nature*. 2005; **438**(7069): 803–819.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  21. Pritchard J, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data.** *Genetics*. 2000; **155**(2): 945–959.  
[PubMed Abstract](#) | [Free Full Text](#)
  22. Falush D, Stephens M, Pritchard JK: **Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies.** *Genetics*. 2003; **164**(4): 1567–1587.  
[PubMed Abstract](#) | [Free Full Text](#)
  23. Earl D, vonHoldt B: **STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.** *Conserv Genet Resour*. 2012; **4**: 359–361.  
[Publisher Full Text](#)
  24. Jakobsson M, Rosenberg NA: **CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure.** *Bioinformatics*. 2007; **23**(14): 1801–1806.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  25. Benson JF, Patterson BR, Wheelodon TJ: **Spatial genetic and morphologic structure of wolves and coyotes in relation to environmental heterogeneity in a *Canis* hybrid zone.** *Mol Ecol*. 2012; **21**(24): 5934–5954.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  26. Rutledge L, Garraway C, Loveless K, *et al.*: **Genetic differentiation of eastern wolves in Algonquin Park despite bridging gene flow between coyotes and grey wolves.** *Heredity*. 2010; **105**(6): 520–531.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  27. Jombart T: **ade4: a R package for the multivariate analysis of genetic markers.** *Bioinformatics*. 2008; **24**(11): 1403–1405.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  28. **R: A Language and Environment for Statistical Computing.** (R Core Team, Vienna, Austria, 2012).  
[Reference Source](#)
  29. Peakall R, Smouse PE: **GenAIEx 6: genetic analysis in Excel. Population genetic software for teaching and research.** *Mol Ecol Notes*. 2006; **6**(1): 288–295.  
[Publisher Full Text](#)
  30. Peakall R, Smouse PE: **GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update.** *Bioinformatics*. 2012; **28**(19): 2537–2539.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  31. Monzón J, Kays R, Dykhuizen D: **Assessment of coyote–wolf–dog admixture using ancestry-informative diagnostic SNPs.** *Mol Ecol*. 2014; **23**(1): 182–197.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  32. Lehman N, Wayne RK: **Analysis of coyote mitochondrial DNA genotype frequencies: estimation of the effective number of alleles.** *Genetics*. 1991; **128**(2): 405–416.  
[PubMed Abstract](#) | [Free Full Text](#)
  33. Roy MS, Geffen E, Smith D, *et al.*: **Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci.** *Mol Biol Evol*. 1994; **11**(4): 553–570.  
[PubMed Abstract](#)
  34. Sacks BN, Brown SK, Ernest HB: **Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history.** *Mol Ecol*. 2004; **13**(5): 1265–1275.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  35. Sacks BN, Mitchell BR, Williams CL, *et al.*: **Coyote movements and social structure along a cryptic population genetic subdivision.** *Mol Ecol*. 2005; **14**(4): 1241–1249.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  36. Sacks BN, Bannasch DL, Chomel BB, *et al.*: **Coyotes demonstrate how habitat specialization by individuals of a generalist species can diversify populations in a heterogeneous ecoregion.** *Mol Biol Evol*. 2008; **25**(7): 1384–1394.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  37. Pilot M, Jedrzejewski W, Branicki W, *et al.*: **Ecological factors influence population genetic structure of European grey wolves.** *Mol Ecol*. 2006; **15**(14): 4533–4553.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  38. Musiani M, Leonard JA, Cluff HD, *et al.*: **Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat colour and association with migratory caribou.** *Mol Ecol*. 2007; **16**(19): 4149–4170.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  39. Muñoz-Fuentes V, Darimont CT, Wayne RK, *et al.*: **Ecological factors drive differentiation in wolves from British Columbia.** *J Biogeogr*. 2009; **36**(8): 1516–1531.  
[Publisher Full Text](#)
  40. Rueness E, Jorde PE, Hellborg L, *et al.*: **Cryptic population structure in a large, mobile mammalian predator: the Scandinavian lynx.** *Mol Ecol*. 2003; **12**(10): 2623–2633.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  41. Rueness E, Stenseth NC, O'Donoghue M, *et al.*: **Ecological and genetic spatial structuring in the Canadian lynx.** *Nature*. 2003; **425**(6953): 69–72.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  42. Pease K, Freedman AH, Pollinger JP, *et al.*: **Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation.** *Mol Ecol*. 2009; **18**(9): 1848–1862.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  43. Brown D, Brenneman RA, Koepfli KP, *et al.*: **Extensive population genetic structure in the giraffe.** *BMC Biol*. 2007; **5**: 57.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  44. Courtisot R, Bernatchez L, Ouellet J, *et al.*: **Significance of caribou (*Rangifer tarandus*) ecotypes from a molecular genetics viewpoint.** *Conserv Genet*. 2003; **4**(3): 393–404.  
[Publisher Full Text](#)
  45. Hoelzel A, Dahlheim M, Stern S: **Low genetic variation among killer whales (*Orcinus orca*) in the eastern North Pacific and genetic differentiation between foraging specialists.** *J Hered*. 1998; **89**(2): 121–128.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  46. Hoelzel A, Potter C, Best P: **Genetic differentiation between parapatric 'nearshore' and 'offshore' populations of the bottlenose dolphin.** *Proc Biol Sci*. 1998; **265**(1402): 1177–1183.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  47. Hull J, Hull A, Sacks B, *et al.*: **Landscape characteristics influence morphological and genetic differentiation in a widespread raptor (*Buteo jamaicensis*).** *Mol Ecol*. 2008; **17**(3): 810–824.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  48. Ellegren H, Savolainen P, Rosén B: **The genetical history of an isolated population of the endangered grey wolf *Canis lupus*: a study of nuclear and mitochondrial polymorphisms.** *Philos Trans R Soc Lond B Biol Sci*. 1996; **351**(1348): 1661–1669.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  49. Prugnolle F, de Meeus T: **Inferring sex-biased dispersal from population genetic tools: a review.** *Heredity (Edinb)*. 2002; **88**(3): 161–165.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  50. Way JG: **Movements of transient coyotes, *Canis latrans*, in urbanized eastern Massachusetts.** *The Canadian Field-Naturalist*. 2007; **121**(4): 364–369.  
[Reference Source](#)
  51. Harrison DJ: **Dispersal characteristics of juvenile coyotes in Maine.** *J Wildl Manage*. Movements of transient coyotes, *Canis latrans*, in urban 1992; **56**(1): 128–138.  
[Reference Source](#)
  52. Zink RM, Barrowclough GF: **Mitochondrial DNA under siege in avian phylogeography.** *Mol Ecol*. 2008; **17**(9): 2107–2121.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  53. White DJ, Wolff JN, Pierson M, *et al.*: **Revealing the hidden complexities of mtDNA inheritance.** *Mol Ecol*. 2008; **17**(23): 4925–4942.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  54. Seddon JM, Parker HG, Ostrander EA, *et al.*: **SNPs in ecological and conservation studies: a test in the Scandinavian wolf population.** *Mol Ecol*. 2005; **14**(2): 503–511.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  55. Andersen DH, Fabbri E, Santini A, *et al.*: **Characterization of 59 canine single nucleotide polymorphisms in the Italian wolf (*Canis lupus*) population.** *Mol Ecol Notes*. 2006; **6**(4): 1184–1187.  
[Publisher Full Text](#)
  56. Sacks BN, Louie S: **Using the dog genome to find single nucleotide polymorphisms in red foxes and other distantly related members of the Canidae.** *Mol Ecol Res*. 2008; **8**(1): 35–49.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  57. Sacks BN, Moore M, Statham MJ, *et al.*: **A restricted hybrid zone between native and introduced red fox (*Vulpes vulpes*) populations suggests reproductive barriers and competitive exclusion.** *Mol Ecol*. 2011; **20**(2): 326–341.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  58. Nielsen R, Signorovitch J: **Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium.** *Theor Popul Biol*. 2003; **63**(3): 245–255.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  59. Sacks BN, Våge DI, Statham MJ: **A medium-throughput SNP assay for detecting genetic variation in coding and non-coding portions of the red fox genome.** *Conserv Genet Resour*. 2009; **1**(1): 459–463.  
[Publisher Full Text](#)
  60. Monzon J: **Data of genetic diversity in northeastern coyotes.** *figshare*. 2014.  
[Data Source](#)

## Current Referee Status:

---

### Referee Responses for Version 1



**Mark Weckel**

Center for Biodiversity & Conservation, American Museum of Natural History, New York, NY, USA

**Approved: 30 May 2014**

**Referee Report:** 30 May 2014

**doi:**[10.5256/f1000research.3819.r3945](https://doi.org/10.5256/f1000research.3819.r3945)

First, thank you for the opportunity to review Javier Monzón's article, "*First regional evaluation of nuclear genetic diversity and population structure in northeastern coyotes (Canis latrans).*"

I found this to be a well-written, engaging, and thoughtful manuscript with important contributions to the use of SNPs in studies of population structure and in elucidating the genetic structure of the northeastern coyote. The author adeptly targets a need for using multiple markers across multiple individuals for proper inference. The author furthermore supports the use of the canine SNP array in studies of coyotes, following the recommendation of [vonHoldt \*et al.\* \(2011\)](#), while also asserting the need for addressing ascertainment bias prior to the use of SNPs in non-model organisms.

I have no major reservations, rather a few minor comments regarding clarification of the methods along with some questions and considerations for future discussion and research.

#### Methods:

1. With such a large sample of high quality tissue available for the analysis, why were scats (n=6) included in the analysis? Were the scats important for adequate sampling across the study area?
2. Fourteen of vonHoldt *et al.*'s. (2011) coyote specimens were included in the initial selection of the 16 SNP panel used in subsequent analyses. Why were only 7 of vonHoldt *et al.*'s (2011) samples included in the analysis of diversity and structure?

#### Results:

1. Using the SNP panel, the author reports the observed and expected heterozygosity for the three putative population zones identified by [Kays \*et al.\* \(2010; Table 2\)](#). It would be interesting to see the  $H_o$  and  $H_e$  for the population clusters identify by the current STRUCTURE analysis (i.e.,  $K=2$  &  $K=3$ ).
2. What are the pairwise  $F_{st}$  values for two genetic clusters also supported by the STRUCTURE analysis?

#### Discussion:

1. There is ample evidence of genetic structure; however, I'm not sure that I would agree with the author's suggestion that there is "*some geographic structuring,*" at least based on Figure 2c. The author does a thorough job of listing the factors and processes that may be involved in explaining a lack of geographic structure, as well as the discrepancy between the current nuclear analysis of

population structure and the previous mtDNA analysis (Kays *et al.* 2010).

- A possible further explanation for the lack of geographic structuring could be that a SNP panel - with such high heterozygosity - may actually reflect older mutations that have had a longer time to move around the landscape via dispersal (see [Morin \*et al.\* 2004](#)). This might allow one to detect patterns of genetic structure without any obvious geographic patterns.
- The author also mentions, correctly, that based on the current state of knowledge, there is no sex-bias in coyote dispersal. I wonder if the author's data set could be used to investigate this by factoring in sex and rerunning the analyses presented in this manuscript.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

**Competing Interests:** No competing interests were disclosed.



**Linda Rutledge**

Biology Department, Trent University, Peterborough, ON, Canada

**Approved: 01 May 2014**

**Referee Report:** 01 May 2014

**doi:**[10.5256/f1000research.3819.r3944](https://doi.org/10.5256/f1000research.3819.r3944)

This is a very well-written, interesting article on utilizing highly-polymorphic SNPs to identify cryptic substructure in a population of eastern coyotes. The analyses are comprehensive (including Bayesian cluster analysis, principal components analysis, and  $F_{st}$  calculations) and the data do seem to support fine-scale population substructure. If it weren't for the extensive analyses I would find it very difficult to believe that 16 SNPs (even highly polymorphic ones) could find such fine-scale structure because it has been estimated that it take at least 15x more SNPs than microsatellites to identify structure in populations ([Haasl & Payseur 2011, \*Heredity\*](#)). Presumably the same clustering patterns would remain if the samples were analyzed at thousands of genome-wide SNPs that have not been selected for their polymorphism - i.e. presumably more, less polymorphic loci should show the same substructure.

The biological/ecological reason for the clustering remains unclear (but that is clearly beyond the scope of the paper and is really quite difficult to tease apart). The author provides plausible explanations for the structure (e.g. "*merging of the two colonization fronts*"; possible deer or human density patters), but I wonder if social structure (i.e. family/pack groups) may contribute to the genetic patterns found.

Overall, the manuscript reads very well and provides a rigorous analysis of a novel dataset that complements the authors other work.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

**Competing Interests:** No competing interests were disclosed.



**Astrid Stronen**

Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Aalborg,



Denmark

**Approved: 18 March 2014**

**Referee Report: 18 March 2014**

**doi:**[10.5256/f1000research.3819.r3946](https://doi.org/10.5256/f1000research.3819.r3946)

Thank you for the opportunity to review this manuscript. Javier Monzón presents a timely and relevant study of population structure in Northeastern coyotes, which complements earlier work in this area and highlights new and intriguing research questions. The author builds on previous genomic analyses of canids with a high-density SNP chip yet a limited sample of Northeastern coyotes, and a study of northeastern coyotes that used several hundred samples but a single genetic marker. The panel of selected SNPs is biased toward the study population of Northeastern coyotes, but the author makes this very clear and includes an evaluation of ascertainment bias across canid groups included in the initial study with the high-density SNP chip from which this panel of 16 loci was selected.

The article is well written, and the author has done a good job explaining the main limitations of the work and outlining why findings should be interpreted with caution. I have one suggestion on supplementary data that could be included for an additional perspective on genetic structure, and some minor comments on the manuscript text:

1. The author describes the SNP results relative to the areas of Ohio, the contact zone and the northeast zone from [Kays et al. 2010](#), and in Table 2 notes that “*most individuals genotyped at 16 nuclear SNPs represent a subset of those individuals sequenced*” for mtDNA. If individual mtDNA results are available for coyotes, it would be valuable to see the extent to which the mtDNA haplotypes classified as Great Lakes wolf (or eastern wolf), coyote and gray wolf from [Kays et al. 2010](#) (Fig 2) correspond with genetic structure based on SNPs. This would provide important information to help understand how similarities/discrepancies in genetic structure are distributed across space.
2. The relatively low genetic diversity for coyotes in Table 3 and the results from [vonHoldt et al. 2011](#) suggest a bias against coyotes, although coyotes may at least in some regions, have higher genetic diversity than wolves considering their range expansion and increase in population size. It might be worth noting that sequencing coyotes directly could provide a more in-depth view of genomic regions under selection. Although I agree with the author’s conclusion about the utility of using a closely related and well-studied species, alternate methods such as genotyping-by-sequencing could be recommended as possibilities for future research.
3. Introduction, 1<sup>st</sup> sentence “*Historically restricted to the open deserts and plains...*”. Though the coyote is widely regarded as a species of western and central North America, it is difficult to exclude the possibility that it could also have been present farther east prior to European colonization and landscape modifications.
4. Introduction, 3<sup>rd</sup> paragraph:
  - The sentence “*Population genetic structure plays considerable roles in evolution... the splitting of one species into two if the environments are markedly different*”. In addition to adaptation, it may be important also to mention genetic drift, and subsequent development of incompatibilities between isolated populations.
  - The sentence starting with “*The detection of genetic structure largely depends on the type and*

*number of molecular markers examined*”: Here you may want to mention explicitly the temporal resolution of markers with different mutation rates (as you have done in the Discussion).

- “*SNPs have become a popular and inexpensive tool*”: perhaps “*increasingly affordable*” or similar would be better here?

5. Discussion, 3<sup>rd</sup> paragraph, sentence: “*...cryptic genetic structure of Northeastern coyotes, especially as they have only inhabited the region for 30-80 years*”. The author notes earlier in the text that coyotes in the northeast zone may descend from a few founders, which likely arrived via a northern route through Ontario, whereas a second colonization front occurred farther south through Ohio. [Bierne et al. 2011](#) proposed that intrinsic (environment-independent) genetic incompatibilities may coincide with ecological boundaries, as these environmental transition zones often “trap” the tension zones resulting from e.g. assortative mating. Such processes might be influencing the contact zone between the southern and northern colonization fronts, and interact with the selective pressures associated with forested, agricultural and urban habitat, and deer density, noted in [Monzón 2012](#) (PhD Dissertation).

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

**Competing Interests:** No competing interests were disclosed.

---